



REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique  
Université Mouloud Mammeri de Tizi-Ouzou  
Faculté de Génie Électrique et d'Informatique  
Département d'Informatique  
Laboratoire LARI



# THESE DE DOCTORAT

(3<sup>ème</sup> cycle LMD)

**SPECIALITE**  
INFORMATIQUE

**OPTION**  
Réseaux, Mobilité et Systèmes Embarqués (RMSE)

---

## Système de monitoring comportemental de vaches laitières

---

Présentée par : ACHOUR Brahim

Devant le Jury Composé de :

M.	LALAM	Mustapha	Professeur	UMMTO	Président
M <sup>me</sup>	SAAD-BELKADI	Malika	Professeur	UMMTO	Encadreur
M.	RIAhLA	Mohamed Amine	MCA	UMBB	Examineur
M.	DAOUI	Mehammed	Professeur	UMMTO	Examineur
M.	FILALI	Idir	MCA	UMMTO	Examineur
M	LAGHROUCHE	Mourad	Professeur	UMMTO	Invité

## Résumé

L'augmentation du nombre d'animaux et la diminution de la main d'œuvre dans les élevages rendent le travail des éleveurs plus conséquent et la surveillance du bien-être et de l'état sanitaire de tous les animaux impossible. Cela, entraîne des pertes considérables. L'élevage de précision issu du développement de nouvelles technologies vient alors en aide. Il permet de soutenir les éleveurs dans le suivi du bien-être, de la santé, de la nutrition et de la reproduction des animaux en leur apportant diverses informations sur ces derniers et sur leur lieu d'élevage pour la prise de meilleures décisions. Dans le cadre de cette thèse, nous nous sommes intéressés aux systèmes de monitoring comportemental de vaches laitières. Ces systèmes peuvent être de bons indicateurs de leurs états de santé. En effet, les changements physiologiques associés à l'apparition de troubles métaboliques conduisent à une altération des activités et du comportement alimentaire des vaches. Le suivi de cette altération permet de fournir aux éleveurs des alertes précises et précoces. Dans cet ordre d'idées, nous avons développé deux systèmes de monitoring comportemental. Le premier est un système de suivi des comportements (activités) des vaches laitières en collectant des données à l'aide d'unités de mesure inertielle attachées à leurs dos. Les comportements suivis sont la position debout, la position couchée, se lever, se coucher, marcher et stationnaire. Le modèle de classification développé est basé sur des modèles de mélange fini univariés et multivariés ainsi que sur des arbres de décision. Le second système est dédié au monitoring du comportement alimentaire à base d'une approche d'analyse d'images. Différents classificateurs basés sur des réseaux de neurones à convolution sont utilisés pour surveiller le comportement alimentaire et effectuer l'identification des vaches laitières. Les résultats obtenus par ces deux systèmes sont intéressants et surpassent les résultats obtenus par d'autres travaux de recherche existants.

**Mots-clés:** Vache laitière, Élevage de précision, Classification du comportement, Modèle de mélange fini, réseau de neurones à convolution.

## Abstract

The increase in the number of animals and the decrease in the labor force in the farms make the work of the farmers more consequent and the monitoring of the welfare and health of all animals impossible. This leads to considerable losses. Precision Livestock Farming resulting from the development of new technologies comes to help. It supports farmers in monitoring the welfare, health, nutrition and reproduction of animals by providing them with various information on these animals and their breeding place for better decision making. As part of this thesis, we are interested in behavioral monitoring systems of dairy cows. These systems can be good indicators of their health states. Indeed, the physiological changes associated to the appearance of metabolic disorders lead to an alteration in the activities and feeding behavior of dairy cows. Monitoring this alteration provides farmers with precise and early alerts. In this vein, we have developed two behavioral monitoring systems. The first is a system for monitoring the behavior (activities) of dairy cows by collecting data using inertial measurement units attached to their backs. The followed behaviors are standing, lying, standing up, lying down, walking and stationary. The developed classification model is based on univariate and multivariate finite mixture models as well as on decision trees. The second system is dedicated to monitoring the feeding behavior based on an image analysis approach. Different classifiers based on convolutional neural networks are used to monitor feeding behavior and perform identification of dairy cows. The obtained results by these two systems are interesting and outperform those obtained by other existing research works.

**Key words:** Dairy Cow, Precision Livestock Farming, Behavior Classification, Finite Mixture Model, Convolutional Neuron Network.

## **Liste des publications**

### **Revues internationales de classe A+:**

Brahim Achour, Malika Belkadi, Rachida Aoudjit, Mourad Laghrouche. Unsupervised automated monitoring of dairy cows' behavior based on Inertial Measurement Unit attached to their back, Computers and Electronics in Agriculture, Volume 167, 2019, 105068, ISSN 0168-1699, <https://doi.org/10.1016/j.compag.2019.105068>

### **Revues internationales de classe A:**

Brahim Achour, Malika Belkadi, Idir Filali, Mourad Laghrouche, Mourad Lahdir, Image analysis for individual identification and feeding behaviour monitoring of dairy cows based on Convolutional Neural Networks (CNN), Biosystems Engineering, Volume 198, 2020, Pages 31-49, ISSN 1537-5110, <https://doi.org/10.1016/j.biosystemseng.2020.07.019>

### **Revues internationales de classe B:**

Brahim Achour, Malika Belkadi, Rachida Aoudjit, Mourad Laghrouche, Mustapha Lalam, Mehammed Daoui. Classification of dairy cows' behavior by energy-efficient sensor. Journal of Reliable Intelligent Environments. (soumis le 22 mars 2020)

### **Conférences internationales (IEEE) :**

Brahim Achour, Malika Belkadi, Rachida Aoudjit, Mustapha Lalam, Mehammed Daoui and Laghrouche Mourad. Dairy cows real time behavior monitoring by energy-efficient embedded sensor, Second International Conference on Embedded and Distributed Systems (EDiS' 2020).

## **Dédicaces**

Je dédie cette thèse à :

- L'esprit de mon père (que Dieu bénisse son âme);
- Ma mère, pour son amour, sa confiance son soutien moral et surtout ses sacrifices. . .
- Mes frères Ferhat et Hocine;
- Mon grand-père Amar et ma grand-mère Ouardia;
- Mes oncles Rachid et Farid et ma tante Ouzna;
- Toute ma famille et mes amis.

## Remerciements

Tout d'abord, mes remerciements sont adressés à Allah qui m'a donné la puissance et le courage pour achever ce travail.

Je tiens à remercier vivement Mr LALAM Mustapha, Professeur au Département d'Informatique de l'Université Mouloud Mammeri de Tizi-Ouzou et Directeur du Laboratoire de Recherche en Informatique (LARI) d'avoir mis à ma disposition tous les moyens du laboratoire et d'avoir accepté de présider le jury de soutenance.

Je tiens à exprimer ma profonde gratitude à Madame SAAD-BELKADI Malika, ma directrice de thèse, Professeur au Département d'Informatique de l'Université Mouloud Mammeri de Tizi-Ouzou, de m'avoir proposé ce sujet et pour sa grande qualité d'encadrement. Je la remercie pour son soutien, ses conseils, sa disponibilité et pour tout le temps qu'elle a consacré pour faire aboutir cette thèse.

Je remercie aussi Mr LAGHROUCHE Mourad, Professeur à l'Université Mouloud Mammeri de Tizi-Ouzou, pour son aide, ses conseils, et surtout pour son encouragement qui m'a beaucoup aidé pour achever ce travail.

J'adresse également mes remerciements à Mr RIAHLA Mohamed Amine, Maître de Conférences classe A à l'Université M'hamed Bougara Boumerdès, d'avoir accepté d'examiner cette thèse. Je remercie Mr DAOUI Mehammed, Professeur au Département d'Informatique de l'Université Mouloud Mammeri de Tizi-Ouzou, pour l'intérêt qu'il a porté à mon travail et pour le temps qu'il a consacré pour examiner cette thèse. Mes remerciements sont également adressés à Mr FILALI Idir, Maître de Conférences classe A au Département d'Informatique de l'Université Mouloud Mammeri de Tizi-Ouzou, d'avoir accepté de participer à ce jury et pour toutes les discussions fertiles au sein de LARI.

Je tiens à exprimer toute ma reconnaissance envers madame SELHI Yasmine gérante de la ferme laitière SOFLAIT, située dans la région de Draâ Ben Khedda, Wilaya de Tizi Ouzou, de m'avoir donné l'occasion de collecter des données et de réaliser des essais au sein de cette ferme. Je ne peux ne pas remercier tous les fonctionnaires de cette ferme.

J'exprime mes sincères remerciements à Madame AOUDJIT Rachida, Professeur au Département d'Informatique de l'Université Mouloud Mammeri de Tizi-Ouzou, à tous les doctorants et les membres du laboratoire LARI ainsi que tous ceux qui ont participé de près ou de loin à l'aboutissement de cette thèse.

J'exprime ma profonde reconnaissance à RAVI et FERHANI Farida d'avoir participé à la correction de mes articles écrits en anglais.

Je remercie chaleureusement ma mère et tous les membres de ma famille pour leur soutien affectif et pour tous leurs encouragements.



# Table des Matières

<b>Table des Matières</b>	<b>v</b>
<b>Liste des Illustrations</b>	<b>ix</b>
<b>Liste des Tableaux</b>	<b>xiii</b>
<b>Liste des abréviations</b>	<b>xvii</b>
<b>Introduction générale</b>	<b>xx</b>
<b>1 L'élevage de précision</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Historique de l'élevage de précision . . . . .	2
1.3 Définition de l'élevage de précision . . . . .	2
1.4 Le principe de base en élevage de précision . . . . .	3
1.5 Conception d'algorithmes en élevage de précision . . . . .	5
1.5.1 La première phase : Le calcul de la variable suivie . . . . .	5
1.5.2 La deuxième phase : la validation de l'algorithme . . . . .	6
1.5.3 Exemple d'un algorithme en élevage de précision . . . . .	6
1.6 Dispositifs électroniques utilisés en élevage de précision . . . . .	7
1.6.1 Les capteurs non invasifs . . . . .	9
1.6.2 Les Capteurs invasifs . . . . .	10
1.7 Utilisation possible des dispositifs électroniques . . . . .	12
1.7.1 L'identification . . . . .	12
1.7.2 La surveillance du comportement . . . . .	16
1.7.2.1 Suivi du comportement alimentaire . . . . .	16
1.7.2.2 Suivi de l'activité . . . . .	17
1.7.3 Surveillance de la santé animale . . . . .	19
1.7.4 La surveillance de la reproduction . . . . .	20

1.7.4.1	La détection du vêlage . . . . .	21
1.7.4.2	La détection de chaleurs . . . . .	22
1.7.5	La surveillance de la morphologie . . . . .	23
1.7.5.1	Suivi du poids vif . . . . .	24
1.7.5.2	Suivi de l'état corporel . . . . .	25
1.8	Impact de l'élevage de précision sur l'économie . . . . .	26
1.9	Impact de l'élevage de précision sur l'environnement . . . . .	27
1.10	Impact de l'élevage de précision sur les éleveurs . . . . .	28
1.11	Conclusion . . . . .	29
<b>2</b>	<b>Le suivi automatisé de la santé des vaches laitières</b>	<b>31</b>
2.1	Introduction . . . . .	31
2.2	Les maladies des bovins laitiers . . . . .	32
2.2.1	Les mammites . . . . .	33
2.2.2	La boiterie . . . . .	34
2.2.3	L'infertilité . . . . .	34
2.2.4	Les troubles métaboliques et digestifs . . . . .	35
2.3	Les systèmes de capteurs utilisés dans le suivi sanitaire . . . . .	35
2.4	Les mesures physiologiques utilisées dans le suivi sanitaire . . . . .	37
2.5	Le suivi comportemental pour la détection de maladies . . . . .	38
2.5.1	Modification comportementale associée à la boiterie . . . . .	38
2.5.1.1	Détection de la boiterie par le monitoring de l'activité . . . . .	38
2.5.1.2	Détection de la boiterie par le monitoring alimentaire . . . . .	39
2.5.2	Modification comportementale associée aux mammites . . . . .	41
2.5.3	Détection des troubles métaboliques et digestifs . . . . .	44
2.6	Conclusion . . . . .	44
<b>3</b>	<b>Exploration de données pour la classification des comportements de vaches</b>	<b>47</b>
3.1	Introduction . . . . .	47
3.2	Partie 1 : la fouille de données . . . . .	48
3.2.1	Le prétraitement de données . . . . .	49
3.2.1.1	Le nettoyage des données . . . . .	49
3.2.1.2	L'intégration de données . . . . .	50
3.2.1.3	La réduction de données . . . . .	50
3.2.1.4	La transformation de données . . . . .	51
3.2.2	L'exploration des données . . . . .	52
3.2.2.1	La classification . . . . .	52

3.2.2.2	Le regroupement (clustering) . . . . .	53
3.2.2.3	La régression . . . . .	53
3.2.2.4	L'évaluation du modèle d'apprentissage . . . . .	54
3.2.2.5	L'optimisation des performances de classification . . . . .	54
3.2.3	Les algorithmes d'apprentissage automatique . . . . .	54
3.2.3.1	Les arbres de décision . . . . .	55
3.2.3.2	Les réseaux de neurones . . . . .	55
3.2.3.3	Le modèle de mélange gaussien . . . . .	56
3.2.3.4	La régression linéaire . . . . .	57
3.2.3.5	Les machines à vecteurs de support . . . . .	57
3.3	Partie 2 : Fouille de données pour la classification comportementale de vaches laitières . . . . .	58
3.3.1	Classification comportementale par apprentissage automatique . . . . .	58
3.3.1.1	La construction de la base de données . . . . .	59
3.3.1.2	La préparation de données . . . . .	61
3.3.1.3	Le déroulement de l'apprentissage . . . . .	62
3.3.2	Classification comportementale par deep learning . . . . .	63
3.4	Conclusion . . . . .	65
<b>4</b>	<b>Monitoring comportemental de vaches laitières à l'aide d'unités de mesure iner- tielle fixées sur leurs dos</b> . . . . .	<b>67</b>
4.1	Introduction . . . . .	67
4.2	Matériel et méthodes . . . . .	68
4.2.1	Étape 01 : Développement, déploiement et collecte de données . . . . .	69
4.2.2	Étape 2 : Prétraitement et traitement des données brutes . . . . .	71
4.2.3	Étape 03 et 04 : Analyse statistique et construction de l'arbre de décision . . . . .	74
4.2.4	Étape 05 : Ajustement des seuils . . . . .	79
4.2.5	Étape 06 : Validation et optimisation des résultats . . . . .	80
4.2.5.1	Validation . . . . .	80
4.2.5.2	Optimisation des résultats et validation finale . . . . .	81
4.3	Résultats . . . . .	82
4.3.1	Analyse statistique . . . . .	82
4.3.2	Résultats d'ajustement de seuil . . . . .	84
4.3.2.1	Détermination du nombre de lois de mélange fini . . . . .	87
4.3.2.2	Détermination des seuils à l'aide des modèles de mélange . . . . .	88
4.3.3	Optimisation des résultats . . . . .	90
4.3.3.1	Par automates . . . . .	92

4.3.3.2	Fusion de données (modèle de mélange multivarié)	93
4.3.4	Construction de l'arbre de décision	97
4.4	Discussion	97
4.5	Conclusion	104
<b>5</b>	<b>Analyse d'images pour l'identification individuelle et le suivi du comportement alimentaire de vaches laitières</b>	<b>107</b>
5.1	Introduction	107
5.2	Matériel et méthode	109
5.2.1	Acquisition et sélection d'images	110
5.2.2	Choix des classes comportementales à suivre	115
5.2.3	Partitionnement de la base de données	115
5.2.4	Méthode de classification	116
5.2.5	Segmentation d'images	118
5.2.6	Identification des vaches	122
5.2.6.1	Approche 1 : Classification par réseau de neurones à convolution (apprentissage profond)	122
5.2.6.2	Approche 2 : Extraction de caractéristiques par un réseau de neurones à convolution et classification par machine à vecteurs de support (SVM).	122
5.2.6.3	Approche 3 : CNN pour regroupement puis Multi CNN pour classification	122
5.2.7	Disponibilité de l'aliment et identification de son type	123
5.3	Résultats	124
5.3.1	Acquisition et sélection d'images	124
5.3.2	Détection de présence et suivi du comportement alimentaire (CNN1 et CNN2)	124
5.3.3	La segmentation sémantique	126
5.3.4	Identification des vaches	131
5.3.5	Disponibilité de l'aliment et identification de son type	132
5.4	Comparaison des résultats	134
5.5	Conclusion	137
	<b>Conclusion générale</b>	<b>140</b>
	<b>Bibliographie</b>	<b>143</b>

# Liste des Illustrations

1.1	Schéma représentatif de la définition de l'élevage de précision . . . . .	4
1.2	Les étapes de base en élevage de précision . . . . .	5
1.3	Les différentes phases d'un algorithme en élevage de précision . . . . .	6
1.4	Etapes de développement d'un algorithme de détection automatique de la boiterie chez les vaches par analyse d'images. . . . .	8
1.5	Echelle d'étiquetage pour la détection de la boiterie (1) : vache non boiteuse avec une ligne du dos horizontale. (2) : vache légèrement boiteuse présentant une légère courbure du dos. (3) : vache modérément boiteuse. (4) : vache boiteuse avec une voussure marquée du dos. (5) : vache boiteuse très voussée. . . . .	9
1.6	Système de mesure de la composition de la respiration. Où : 1 à 4 et 6 désignent des moyens de fixation, 5 indique des ouvertures pour narines et 7 représente le trajet d'écoulement . . . . .	11
1.7	Exemple d'un enregistrement vidéo d'une vache quittant la salle de traite. . . . .	11
1.8	Boitier de positionnement GPS attaché à une vache. . . . .	11
1.9	Capsule utilisée pour surveiller à distance la température intraruménale d'une vache sur une longue période de temps. . . . .	12
1.10	Panorama et emplacement des capteurs disponibles sur le marché pour les bovins. . . . .	13
1.11	Différents emplacements des systèmes d'identification RFID . . . . .	15
1.12	Capteur de pression assurant le monitoring de la rumination. . . . .	17
1.13	Système de surveillance de la rumination par un capteur acoustique. . . . .	18
1.14	Capteur de température attaché à la queue permettant le monitoring de la fièvre. . . . .	21
1.15	Vache équipée de la ceinture Agribelt de Databel agrimonitor. . . . .	22
1.16	Principe de fonctionnement du système Heatime. . . . .	23
1.17	Système de mesure de distance entre la thoracique gauche et droite de la poitrine pour estimer le poids vif. . . . .	24
2.1	Systèmes de capteurs assurant le suivi sanitaire des vaches . . . . .	37

3.1	Processus de découverte de connaissances. . . . .	49
3.2	Architecture d'un système d'intégration de données . . . . .	51
3.3	Architecture d'arbre de décision binaire . . . . .	55
3.4	Architecture basique d'un réseau de neurones artificiels . . . . .	56
3.5	Hyperplan de base d'une machine à vecteurs de support linéaire . . . . .	58
3.6	Processus d'apprentissage automatique appliqué dans la classification comportementale des vaches . . . . .	59
3.7	Architecture basée sur les réseaux de neurones à convolution utilisée lors de la classification comportementale des vaches . . . . .	64
4.1	Enchaînement de la procédure de développement du système de classification comportementale. . . . .	69
4.2	Équipement d'acquisition de données . . . . .	71
4.3	Le schéma fonctionnel matériel du système d'acquisition de données. . . . .	73
4.4	Automate de modélisation des comportements de base d'une vache laitière. . . . .	74
4.5	Automate de modélisation des différents comportements d'une vache se trouvant dans l'état debout. . . . .	75
4.6	Arborescence de modélisation des comportements suivis. . . . .	76
4.7	Interface Matlab permettant de visualiser les signaux acquis. . . . .	77
4.8	Exemple montrant les étapes de la procédure de traitement des données en utilisant une fenêtre de segmentation de 5 secondes et une fréquence de 1 Hz. . . . .	77
4.9	Méthode de fusion de données. . . . .	83
4.10	Régression linéaire mixte sur les données de trois vaches de races différentes et de parités différentes pour la classification de diverses activités à l'aide de plusieurs variables. . . . .	85
4.11	Fonctions de densité spécifiant les classes comportementales suivies selon les variables choisies. . . . .	89
4.12	Modèle de mélange de trois lois normales de la variable médiane de l'axe-x de l'accéléromètre avec une segmentation de 1 seconde et une fréquence de 1 Hz (vache1). . . . .	91
4.13	Modèles de mélange consécutifs (A, B, C) de 3 lois normales (a) suivies de 3 lois normales (b) et de 2 lois normales (c) de la variable intervalle interquartile de l'axe-x de l'accéléromètre utilisée pour séparer la classe transition des autres classes. . . . .	91

4.14	Modèles de mélange consécutifs (A, B, C) utilisés pour la séparation des classes stationnaire et marche. (a) Mélange préliminaire de deux lois normales, (b) Mélange de deux lois normales qui déterminent le seuil D et (c) Mélange de deux lois normales qui déterminent le seuil E. . . . .	92
4.15	Validation de la transition debout/couchée par l'utilisation de l'espacement entre un mélange de deux lois normales des données des 4 minutes qui précèdent et qui succèdent la transition. . . . .	94
4.16	Classification multivariée des classes position debout, position couchée côté droit et position couchée côté gauche à l'aide des deux variables (médiane de l'axe-x et médiane de l'axe-y de l'accéléromètre). . . . .	95
4.17	Les quatre mélanges multivariés consécutifs séparant la classe comportementale transition debout/couchée à l'aide de deux variables (intervalle interquartile de l'axe-x et de l'axe-y de l'accéléromètre). . . . .	96
4.18	Architecture des arbres de décision résultants. . . . .	98
5.1	Procédure d'acquisition, de sélection et de classification d'images. . . . .	110
5.2	Zone et équipement d'acquisition d'images. . . . .	112
5.3	Organigramme de sélection d'images. . . . .	113
5.4	Images montrant les trois classes suivies. . . . .	115
5.5	Automate d'états finis déterministes qui modélise une période d'alimentation. . . . .	116
5.6	Architecture du réseau de neurones à convolution . . . . .	119
5.7	Schéma explicatif de la troisième approche. . . . .	123
5.8	Les différents aliments distribués dans la mangeoire. . . . .	124
5.9	Valeurs de la fonction de perte et d'exactitude durant l'entraînement et le test du CNN2 (a) et du CNN1 (b). . . . .	127
5.10	Caractéristiques extraites par la dernière couche de convolution du CNN2. . . . .	128
5.11	Résultat de la fonction de perte (Loss) lors du réentraînement du modèle Xception pour la segmentation des images. . . . .	129
5.12	Résultats de la segmentation d'images de huit différentes vaches en appliquant le modèle Xception réentraîné. . . . .	130
5.13	Caractéristiques extraites par la troisième et la quatrième couche de convolution du CNN4. . . . .	134



# Liste des Tableaux

1.1	Les outils de monitoring sanitaire des vaches boiteuses. . . . .	20
2.1	Fréquence des maladies en élevage . . . . .	33
2.2	Échelle de classification des vaches boiteuses . . . . .	34
2.3	Valeurs des quatre composantes principales correspondantes à chaque score de boiterie. . . . .	39
2.4	Valeurs des quatre composantes principales correspondantes à chaque score de boiterie. . . . .	40
2.5	Effet de la boiterie sur les variables suivies. . . . .	41
2.6	L'effet de lésions au pied et au jarret sur l'apport alimentaire. . . . .	41
2.7	Effet des mammites sur les différentes variables comportementales. . . . .	43
2.8	Impact des mammites sur : l'alimentation, la rumination, l'activité, le toilettage et la rotation de tête. . . . .	43
2.9	Impact des mammites sur : l'activité, l'abreuvement et le toilettage. . . . .	43
2.10	Les différences (selon les unités de mesure) entre : 5 jours avant le diagnostic clinique (CD), le jour du diagnostic clinique des différentes variables. . . . .	45
3.1	Synthèse des étapes d'acquisition, de prétraitement et de classification de données utilisées lors de la conception de systèmes de monitoring des comportements des vaches laitières par apprentissage automatique. . . . .	60
4.1	Description de la base de données . . . . .	72
4.2	Classes comportementales suivies . . . . .	78
4.3	Caractéristiques extraites de l'IMU et variables calculées pour chaque caractéristique . . . . .	79
4.4	critères d'évaluation de la validation de l'arbre de décision. . . . .	81
4.5	Régression linéaire qui montre l'influence des activités comportementales sur les variables suivies. . . . .	86

4.6	Résultats de classification des classes debout, couchée côté droit et couchée côté gauche à l'aide des deux modèles de mélange univarié et multivarié. . . .	93
4.7	Résultats de la validation de la classification de la classe transition debout/couchée. . . .	94
4.8	Résultats de la classification de la classe marche et stationnaire avec 4 méthodes. Où : Méthode 1 : un mélange univarié de deux fonctions de densité de probabilité (seuil D) Méthode 2 : un mélange multi variée de trois variables optimisant la méthode 1 (seuil D, G, H) Méthode 3 : un mélange univarié consécutif de deux fonctions de densité de probabilité suivis de deux autres fonctions de densité de probabilité (seuil E) Méthode 4 : un mélange multi variée de trois variables optimisant la méthode 3 (seuil E, G, H). . . . .	97
4.9	Comparaison entre le présent travail et d'autres publications pour les classes comportementales debout et couchée. . . . .	103
4.10	Comparaison entre le présent travail et autres publications pour la classe comportementale marche. . . . .	103
4.11	Comparaison entre le présent travail et autres publications pour la classe comportementale transition entre la position debout et couchée. . . . .	104
5.1	Nombre d'images utilisées dans l'entraînement et le test des CNNs . . . . .	118
5.2	Nombre d'images acquises pour chaque vache avant et après la sélection. . .	125
5.3	Matrice de confusion, exactitude et précision du CNN2. . . . .	126
5.4	Résultats obtenus lors du réentraînement du modèle Xception pour la segmentation des images . . . . .	129
5.5	Matrice de confusion de la première approche. . . . .	131
5.6	Taux d'exactitude des trois approches . . . . .	133
5.7	Comparaison des performances de classification pour la détection de la présence des vaches. . . . .	135
5.8	Comparaison des performances de classification, entre notre travail et différents travaux de recherche sur le suivi alimentaire. . . . .	135
5.9	Comparaison des performances de classification entre notre travail et différents travaux de recherche sur l'identification. . . . .	136
5.10	Comparaison des performances de la segmentation de notre méthode et celles des travaux de recherche existants. . . . .	138

# Liste des abréviations

AFD	Automate Fini Déterministe.
AGMM	Adaptive Gaussian Mixture Model ou modèle de mélange gaussien adaptatif.
BVD	Bovine Viral Diarrhea ou diarrhée virale bovine.
CNN	Convolutional Neural Network ou réseau de neurones à convolution.
DAC	Distributeur Automatique de Concentré.
DMI	Dry Matter Intake ou matière sèche ingérée.
EC	Electrical Conductivity ou conductivité électrique.
EdP	Elevage de Précision.
FK-NN	Fuzzy K-Nearest Neighbor ou K-plus-proches voisins flou.
FN	Faux Négatif.
FP	Faux Positif.
GES	Gaz à Effet de Serre.
GMM	Gaussian Mixture Model ou modèle de mélange gaussien.
GPS	Global Positioning System ou système de positionnement global.
HD	High Definition ou Haute Définition.
Hz	Hertz.
IBR	Infectious Bovine Rhinotracheitis ou rhinotracheite bovine infectieuse.
IMU	Inertial Measurement Unit ou unité de mesure inertielle.

---

K-NN	K-Nearest Neighbors ou K plus proches voisins.
MEMS	Micro Electro-Mechanical Systems ou systèmes micro-électromécaniques.
MIOU	Mean Intersection Over Union ou moyenne d'intersection sur l'union.
MLP	Multi-Layer Perceptron ou perceptron multicouche.
MMC	Modèles de Markov Cachés.
N2O	Nitrous Oxide ou protoxyde d'azote.
NEC	Note d'Etat Corporel.
NTIC	Nouvelles Technologies de l'Information et de la Communication.
ONUAA	Organisation des Nations Unies pour l'Alimentation et l'Agriculture.
P-valeur	Valeur de Probabilité.
PCA	Principle Component Analysis ou analyse en composantes principales.
PNN	Probabilistic Neural Network ou réseau de neurones probabilistes.
RCNN	Region-based Convolutional Neural Network ou réseau de neurones convolutionnels basé sur des régions.
RFID	Radio Frequency IDentification ou radio-identification.
ROI	Region Of Interest ou région d'intérêt.
RSE	Residual Standard Error ou l'erreur standard résiduelle.
RTLS	Real-Time Location System ou système de localisation en temps réel.
SB	Score de Boiterie.
SCC	Somatic Cell Count ou compteur de cellules somatiques.
SSIM	Structural SIMilarity ou similarité structurelle.
SVM	Support Vector Machines ou machine à vecteurs de support.
TIC	Technologies de l'Information et de la Communication.
USB	Universal Serial Bus ou bus universel en série.

UWB Ultra-WideBand ou bande ultra large.

VN Vrai Négatif.

VOS Velocity Of Sound ou vitesse du son.

VP Vrai Positif.



## Introduction générale

La prolifération des Nouvelles Technologies de l'Information et de la Communication (NTIC) nous ouvre de nouvelles perspectives, notamment dans les secteurs économiques et en particulier l'agriculture. On parle désormais d'agriculture de précision, d'élevage de précision et d'alimentation de précision. Le déploiement technologique dans les fermes intelligentes s'appuie sur la multiplication de nœuds capteurs ainsi que sur le développement de la télédétection, du télétraitement et de la géolocalisation.

L'élevage de précision est avant tout une nouvelle discipline de gestion de la ferme et du cheptel hébergé. Elle nécessite de réfléchir différemment et de prendre des décisions qui sont bien plus élaborées que les décisions habituellement prises. L'élevage de précision est principalement utilisé en élevage bovin laitier pour la détection des troubles sanitaires, la surveillance des événements liés à la reproduction, etc. Dans une même ferme, les vaches laitières ne sont pas nécessairement de même parité (nombre de portées), conformation (morphologie), race et stade de lactation. Pour cela, l'élaboration d'outils de diagnostic commence par la mesure de performances individuelles telles que la production laitière ou le comportement (alimentaire ou activité) pour ensuite les inclure dans des modèles prédictifs d'aide à la décision.

La collecte de variables comportementales par des capteurs et leur association aux modèles d'analyses prédictifs en élevage de précision offrent le potentiel d'améliorer le processus de suivi sanitaire. En effet, lorsqu'une maladie affecte une vache, son comportement montre des altérations. Ces dernières peuvent être détectées par l'éleveur en peu de temps si la taille du troupeau n'est pas importante. Néanmoins, garder un œil sur un grand troupeau peut être un défi majeur et les étapes de prévention ou de traitement de la maladie peuvent être entamées en retard.

La collecte de paramètres physiologiques tels que le sang, les fluides du rumen et le poids nécessite beaucoup de temps et de tests approfondis en laboratoire. Ces tests sont coûteux, stressants pour l'animal et amplifient le travail de l'éleveur. Pour cela, une technologie d'observation nouvelle et moderne basée sur des capteurs est nécessaire. Elle permet de collecter plus d'informations, de déterminer et de détecter tout changement dans le comportement

avant l'apparition de signes cliniques, permettant ainsi à l'éleveur d'être alerté, de manière précoce, pour prendre les mesures nécessaires le plus rapidement possible. L'utilisation de capteurs pour le suivi comportemental de vaches laitières permet également de prévenir et de limiter la propagation des maladies, de réduire l'intensité de travail et la pénibilité de certaines tâches répétitives (la distribution d'aliments, la traite), d'accroître les gains économiques, etc.

Dans le cadre de cette thèse, nous nous sommes intéressés au développement de systèmes de surveillance comportemental (activité et comportement alimentaire) de vaches laitières. Les résultats de ces systèmes peuvent être par la suite exploités pour la détection de troubles sanitaires. La collecte de données est réalisée dans la ferme commerciale SOFLAIT, située dans la région de Draa Ben Khedda, Wilaya de Tizi Ouzou.

Nous avons en premier lieu élaboré deux systèmes embarqués. Le premier est doté d'un capteur de mouvement dit unité de mesure inertielle (IMU pour Inertial Measurement Unit en anglais). Il est attaché aux dos des vaches et permet l'acquisition de données. Ces données sont ensuite traitées et classifiées en plusieurs comportements (activités). Le modèle de classification utilisé est basé sur des modèles de mélange univarié et multivarié ainsi que sur les arbres de décision. Le second système est doté de caméras installées dans les zones d'alimentation du bâtiment d'élevage. Il permet l'acquisition et la sélection en temps réel d'images qui sont utilisées pour l'identification individuelle et la surveillance du comportement alimentaire des vaches à base de réseaux de neurones à convolution. Les résultats obtenus par ces deux systèmes sont prometteurs et produisent des performances meilleures comparativement à ceux obtenus par d'autres travaux de recherche.

Cette thèse est organisée en cinq chapitres.

- Dans le premier chapitre, nous présentons une introduction générale au contexte de l'élevage de précision et le principe de fonctionnement des technologies actuellement utilisées dans les fermes intelligentes. Ensuite, nous présentons quelques équipements utilisés en élevage de précision tels que les outils d'identification, de surveillance du comportement (activités et alimentaire), de la santé, de la reproduction et de la morphologie. Puis, nous donnons l'impact de l'élevage de précision sur l'économie, l'environnement et le travail des éleveurs. Le but de ce chapitre est de mettre en relation tous les équipements déployés avec leurs valeurs ajoutées et comprendre leurs intérêts, avantages, limites et leurs fonctionnements.
- Dans le deuxième chapitre, nous donnons un aperçu sur le développement des NTICs dans le cadre du suivi sanitaire des vaches laitières. En premier lieu, nous définissons les maladies les plus courantes qui touchent les vaches laitières et nous décrivons par

la suite l'impact de ces maladies sur leurs comportements. L'ensemble des corrélations entre les modifications comportementales associées à l'apparition des troubles de santé est également donné. Le but est de comprendre comment une maladie impacte le comportement des vaches laitières, ce qui nous permet de sélectionner les comportements importants qui doivent être suivis.

- Dans le troisième chapitre, nous donnons une synthèse sur les méthodes d'exploration de données utilisées pour le suivi comportemental de vaches laitières. Premièrement, nous décrivons les principales étapes utilisées dans le data mining (fouille de données). Ensuite, nous montrons comment ces étapes sont appliquées dans le cadre du monitoring comportemental.
- Dans le quatrième chapitre, nous présentons notre première contribution, dans laquelle nous avons élaboré un système de détection automatisé des comportements des vaches laitières par IMU attachée à leurs dos. Il s'agit d'un système embarqué permettant de discriminer les activités suivantes : debout, couchée vers le côté droit, couchée vers le côté gauche, se lever, se coucher, marcher et stationnaire.
- Dans le cinquième chapitre, nous présentons la deuxième contribution. Il s'agit d'un système qui permet l'identification automatique et le monitoring du comportement alimentaire de vaches laitières. Il se base sur l'analyse d'images et l'utilisation de réseaux de neurones à convolution.

Nous terminons cette thèse par une conclusion générale et des perspectives.



# Chapitre 1

## L'élevage de précision

### 1.1 Introduction

Le secteur de l'élevage représente un tiers de la production agricole mondiale [1]. Il contribue considérablement aux conditions de vie et à la nutrition, notamment pour les individus pauvres et vulnérables au sein des pays en développement. Les moyens d'existence de 1,3 milliard de personnes reposent sur l'élevage [1]. L'élevage génère également des coproduits et constitue dans de nombreuses économies une réserve de richesse.

Selon les statistiques publiées par l'Organisation des Nations Unies pour l'Alimentation et l'Agriculture (ONUAA) [2], l'élevage occupe 70 % de toutes les terres cultivables et 30 % de la surface terrestre de la planète. En 2010, les produits d'origine animale tels que la viande, le lait et les œufs ont fourni 16 % de l'ensemble des calories et 31 % de protéines alimentaires à l'échelle mondiale. En 2013, les estimations du nombre de têtes à l'échelle mondiale faisaient état de 23 milliards de volailles, 1.6 milliard de bovins, 2 milliards d'ovins et de caprins et 1 milliard de porcs. L'élevage joue aussi un rôle important à travers la production de fumier, de cuir et de fibres.

Selon la ONUAA une action publique est indispensable pour réduire l'impact de la production animale sur l'environnement et garantir que cette filière contribue de manière décisive à la sécurité alimentaire et à la diminution de la pauvreté [2]. Donc le développement de méthodes à utiliser en élevage est d'une importance capitale à l'échelle mondiale.

Pour répondre aux attentes du marché, l'élevage a évolué d'un modèle familial vers un modèle productiviste. Néanmoins, ce type d'élevage présente des problèmes majeurs tels que :

- L'agrandissement de la taille des élevages.
- La diminution de leur nombre.
- La réduction de la main-d'œuvre présente dans les bâtisses d'élevages par rapport au

cheptel hébergé.

- Les conditions de travail ne correspondent plus aux attentes des éleveurs en matière de confort de vie, de loisirs et de rentabilité salariale.

Le travail d'un éleveur est de moins en moins rentable car le prix des matières premières utilisées en alimentation est toujours plus élevé. Une des solutions apportées par la science est l'automatisation et l'informatisation des élevages, que l'on appelle « élevage de précision ». Ce type d'élevage ouvre de réelles perspectives. Il permet essentiellement la surveillance de l'état de santé, de reproduction, de production et de morphologie de tous les animaux. Ce qui améliore la productivité et assure le bien-être du bétail.

## 1.2 Historique de l'élevage de précision

La notion de l'élevage de précision est plus ancienne d'une dizaine d'années que celle de l'agriculture de précision [3]. L'élevage de précision a fait son apparition sur le continent européen : La Grande Bretagne et la Belgique, puis adoptée en Allemagne, au Danemark, aux Pays-Bas et en Finlande. En 2002, avec le soutien de chercheurs Britanniques et Belges, il a été importé en Australie, puis plus tard en Amérique du Nord, Asie, Brésil et en Afrique du Sud [4].

L'élevage de précision a d'abord concerné les élevages avicoles et porcins, puis les élevages laitiers et caprins. Vers les années 1980 apparaissent les premiers robots de traite et les premiers DAC (Distributeur Automatique de Concentré) aux Pays-Bas. L'exploitation des données issues de capteurs est permise avec l'informatisation des élevages à partir des années 1990. Depuis ces années et surtout depuis les années 2000, de plus en plus de capteurs existent pour contrôler l'alimentation, la reproduction, l'état de santé des animaux, l'environnement des bâtiments, etc. Les outils utilisés en élevage de précision sont également de plus en plus performants [4].

## 1.3 Définition de l'élevage de précision

Plusieurs définitions sont proposées dans la littérature pour l'Élevage de Précision (EdP) dit Precision Livestock Farming (PLF) ou encore Smart Farming en anglais.

Bewley [5] définit l'élevage laitier de précision comme l'utilisation de technologies permettant de mesurer des indicateurs physiologiques, comportementaux ou de production sur les animaux pour améliorer les stratégies de gestion du troupeau et les performances de l'élevage qui peuvent être économiques, sociales ou environnementales.

Selon Berckmans [6], il s'agit du pilotage de l'élevage grâce à la surveillance et aux enregistrements de mesures automatisées et en temps réel de la production, de la reproduction, de la santé et du bien-être des animaux.

Selon [7], l'élevage de précision est un domaine interdisciplinaire intégrant des concepts d'informatique, d'économie, d'élevage, de nutrition animale et d'ingénierie. Il est souvent utilisé pour décrire de nombreuses technologies visant à améliorer les systèmes de gestion des fermes. L'élevage de précision est le résultat du fait que dans le passé, l'éleveur avait plus de temps pour faire des observations sur chacun de ses animaux et d'utiliser ses indicateurs personnels afin de repérer les anomalies et les changements. Alors qu'aujourd'hui, les éleveurs n'ont plus le temps pour les observations visuelles et auditives [4] à cause de l'augmentation de la taille des fermes ainsi que le cheptel de chaque exploitation.

Selon [3], l'élevage de précision signifie l'utilisation de la technologie moderne pour améliorer la surveillance du bétail et augmenter l'efficacité des processus de croissance des produits animaux (le lait, les œufs, la viande, etc.). L'élevage de précision crée des systèmes de gestion basés sur la surveillance et le contrôle continu en temps réel et automatisés de la santé, du bien-être des animaux, de la production, de la reproduction et de l'impact environnemental des animaux.

La définition ci-dessous est à notre avis plus globale et relativement consensuelle aux différentes filières animales : L'élevage de précision, c'est l'utilisation coordonnée de capteurs pour mesurer des paramètres comportementaux, physiologiques ou de production sur les animaux ou les caractéristiques du milieu d'élevage (température, hygrométrie, ventilation, etc.) et de Technologies de l'Information et de la Communication (TIC) pour échanger, stocker, transformer et restituer ces informations à l'éleveur afin de l'aider dans sa prise de décision en complément de ses observations et à piloter son élevage, en anticipant certaines situations qui pourraient devenir critiques pour un animal donné ou pour l'ensemble de l'élevage [6]. La Fig1.1 résume cette définition [8].

Notons que l'utilisation d'automatismes permettant de décharger l'éleveur de certaines tâches astreignantes (traite, alimentation, régulation de l'ambiance des bâtiments) peut être associée à l'élevage de précision. Si les automatismes sont couplés à l'utilisation de capteurs et de technologies de transfert d'informations. En effet, ces automatismes peuvent en général être déclenchés, régulés ou pilotés grâce aux capteurs qu'ils contiennent.

## 1.4 Le principe de base en élevage de précision

Le principe de base en élevage de précision est de reconnaître que l'animal est un « système CITD » : c'est à dire Complexe, Individuellement différent, variable dans le Temps et Dy-

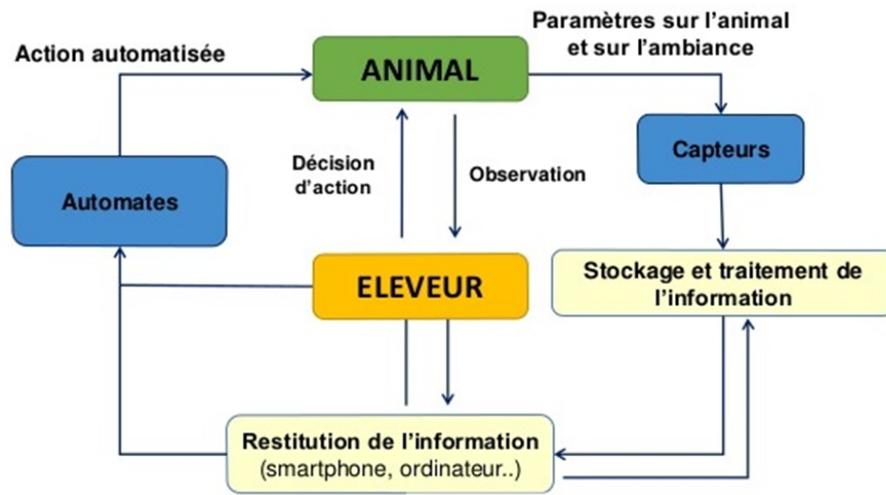


Fig 1.1: Schéma représentatif de la définition de l'élevage de précision

namique [9].

La complexité provient du fait que tout organisme en vie est toujours plus compliqué que tout système électromécanique. La complexité de transmission d'informations dans une seule cellule d'un organisme vivant par exemple est beaucoup plus élevée que dans la plupart des systèmes fabriqués par l'homme. La notion de «Individuellement différent» signifie que tous les organismes vivants sont individuellement différents dans leurs réponses. L'organisme vivant réagit différemment selon l'environnement qui l'entoure ou qui le stimule. Le stress fait partie des variables qui rendent la réaction de l'organisme différente d'un animal à un autre. Le caractère de la variabilité dans le Temps d'un organisme vivant signifie que la réponse de cet organisme à l'environnement ou à un facteur de stress peut être différente chaque fois qu'un événement se produit. Enfin, tous les organismes vivants sont des systèmes Dynamiques qui évoluent dans le temps.

L'élevage de précision part du principe que l'animal est la partie la plus cruciale dans le processus de production dans une ferme. L'un des initiateurs Belges de l'élevage de précision (Berckmans) précise que l'utilisation des technologies de précision en élevage implique trois conditions relatives aux informations fournies sur les animaux (Fig 1.2) [10]. Ces conditions sont :

- La première condition à remplir est que les variables animales doivent être mesurées et analysées en continu. Ces variables peuvent être très différentes, telles que le poids, l'activité, le comportement, l'apport alimentaire, la production sonore, les variables physiologiques (température corporelle, fréquence respiratoire, variables sanguines), etc.
- La deuxième condition pour assurer la surveillance des animaux est qu'à chaque instant une prédiction fiable doit être disponible sur la façon dont les variables animales varieront ou

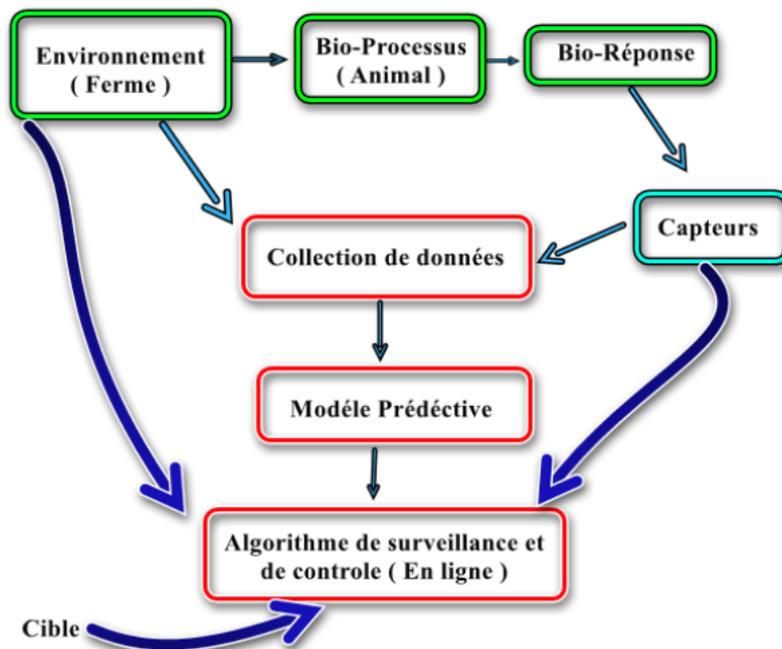


Fig 1.2: Les étapes de base en élevage de précision

comment l'animal répondra aux changements environnementaux.

- La troisième condition est que le modèle de prédiction et les mesures en ligne doivent être intégrés dans un algorithme d'analyse. Cet algorithme aura pour rôle de surveiller ou gérer les animaux automatiquement pour leur garantir une bonne santé et un bien-être.

## 1.5 Conception d'algorithmes en élevage de précision

La conception d'algorithmes en élevage de précision a pour but de déterminer l'état d'une variable cible, par exemple une maladie (l'animal est-il malade ou non). Cela nécessite deux phases: la première phase est le calcul d'une variable suivie et la seconde est la validation de l'algorithme. la Fig 1.3 illustre ces différentes phases.

### 1.5.1 La première phase : Le calcul de la variable suivie

La variable suivie peut donner un avertissement précoce de l'état de la variable ciblée par l'algorithme. Par exemple, le suivi du mouvement de la tête d'une vache (variable suivie) par un accéléromètre permet le monitoring du comportement alimentaire (variable cible). Le calcul de la variable suivie se fait en deux étapes :

- L'étiquetage et l'interprétation de données : cette étape a pour but de comprendre le

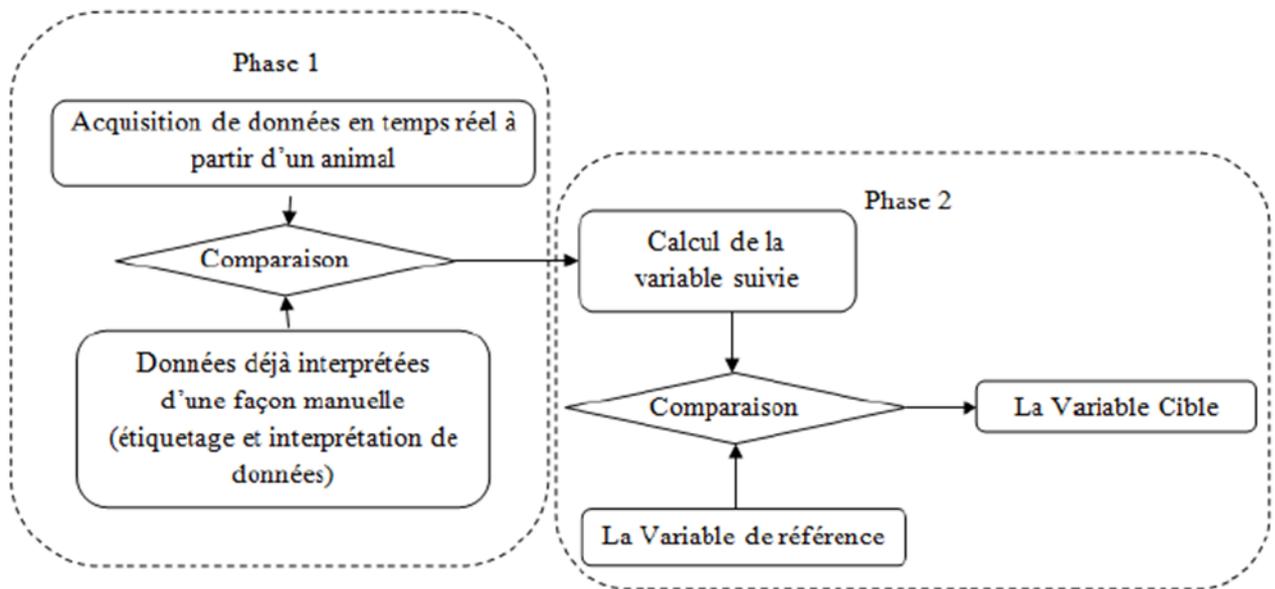


Fig 1.3: Les différentes phases d'un algorithme en élevage de précision

comportement dynamique de la variable suivie et de comprendre sa signification. L'étiquetage se fait par des analyses manuelles détaillées des données mesurées sur le terrain.

- La comparaison : ces données interprétées et annotées seront comparées à des données récupérées en temps réel à partir du terrain. Cette comparaison permet de calculer la variable suivie.

### 1.5.2 La deuxième phase : la validation de l'algorithme

L'étape de validation consiste à comparer la variable suivie à une variable de référence. La variable de référence est issue par l'utilisation d'une méthode démontrée scientifiquement exacte (l'analyse sanguine par exemple) [11]. Cette variable est fiable pour la quantification de la variable cible. La qualité d'une variable de référence à refléter la réalité détermine l'exactitude de l'algorithme. Son choix est d'une importance capitale.

### 1.5.3 Exemple d'un algorithme en élevage de précision

Cet exemple concerne le développement d'un système de surveillance en temps réel de la boiterie de vaches (Fig 1.4) [12].

La boiterie peut être définie comme une déviation de la démarche résultant d'une douleur ou d'une gêne occasionnée par des blessures ou d'une maladie aux pattes. La boiterie est le

problème numéro un en matière de bien-être animal. Dans la littérature, plus de 200 causes possibles ont été énumérées.

Dans cet exemple :

1. La variable cible : est une valeur indiquant le degré de boiterie de la vache.
2. La variable suivie : parmi les différentes variables calculées à partir de l'image de la vache la meilleure variable qui indiquerait la boiterie à un stade précoce est choisie comme la variable suivie.
3. L'étiquetage des données: un observateur humain sur la scène réalise l'annotation de boiterie. Une autre méthode consiste à effectuer un marquage visuel d'images vidéo par un être humain ( image par image) pour marquer les écarts par rapport à la marche normale (Fig 1.5).
4. La variable de référence : c'est le score d'un expert humain pour quantifier le degré de boiterie de la vache.

## 1.6 Dispositifs électroniques utilisés en élevage de précision

Plusieurs dispositifs électroniques (capteurs) sont utilisés en élevage de précision, ils sont diversifiés et sans cesse renouvelés. Ils permettent : (i) d'automatiser une mesure pouvant être réalisée manuellement (pesée de la quantité d'aliment ingérée) ou visuellement (chevauchements, vêlages), (ii) de mesurer de nouveaux paramètres non détectables par l'oeil humain ou difficilement mesurables (activité, rumination, température), etc.

Il s'agit notamment de l'utilisation de caméras, de microphones, de capteurs de mouvement (les accéléromètres 3D, les gyroscopes), de capteurs de température, de capteurs de contraction musculaire, etc. Ces capteurs sont placés suffisamment proches de l'animal pour qu'ils puissent remplacer les yeux et les oreilles des éleveurs pendant 7 jours par semaine, 24 heures par jour et 3600 secondes par heure.

Selon l'emplacement choisi pour ces capteurs, on distingue deux types : les capteurs attachés et non attachés [13]. Les capteurs attachés sont portés par l'animal à l'extérieur de son corps (comme l'accéléromètre) ou à l'intérieur de son corps (comme un bolus dans le rumen ou capteur sous la peau). Cependant, les capteurs non attachés sont des capteurs placés dans le bâtiment d'élevage (hors animal) à des endroits où les animaux passent à proximité pour faire des mesures ou sont intégrés dans un automate de traite, de pesée ou d'alimentation [14].

Selon toujours le critère d'emplacement nous pouvons séparer les capteurs en capteurs invasifs et non invasifs.

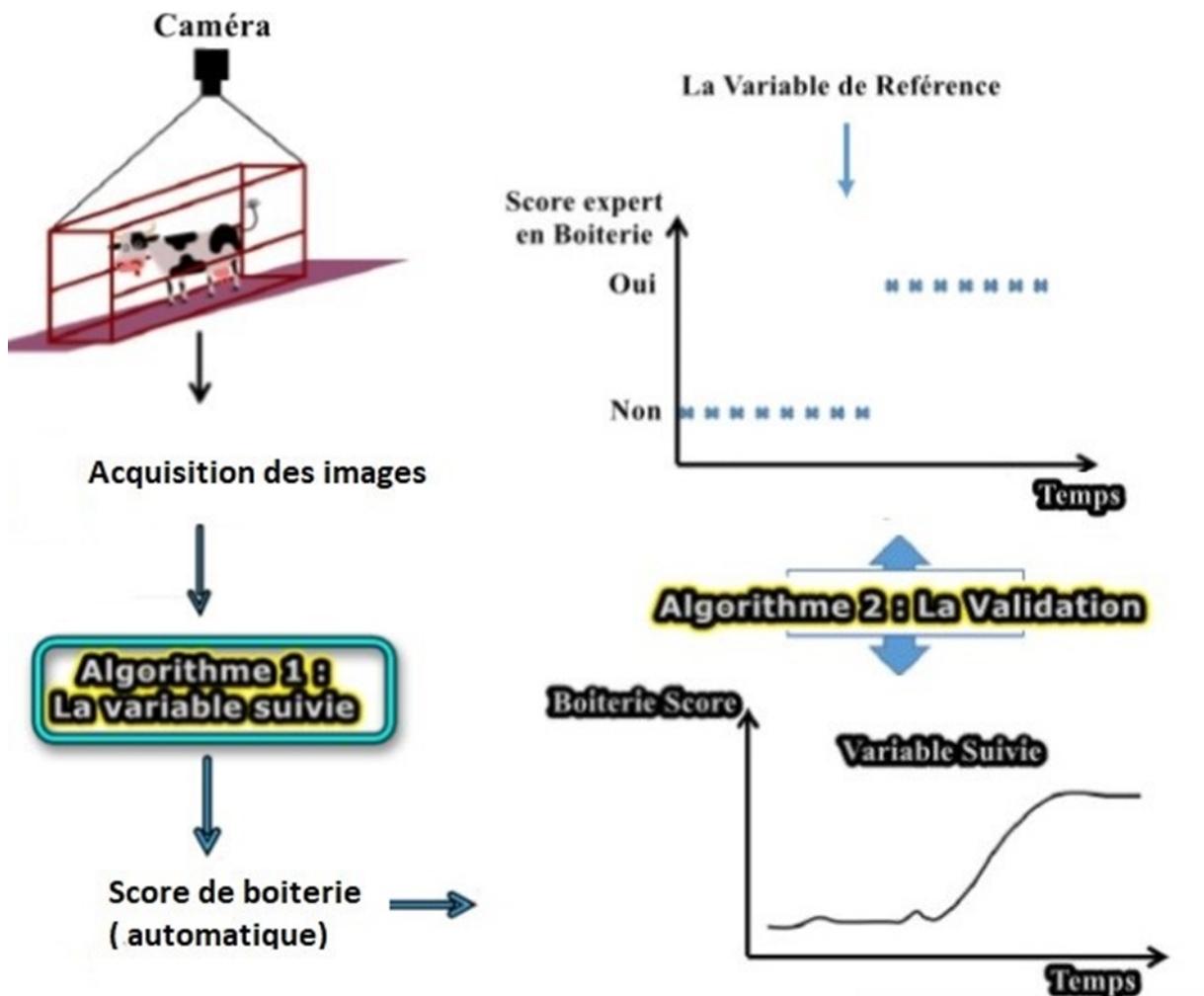


Fig 1.4: Etapes de développement d'un algorithme de détection automatique de la boiterie chez les vaches par analyse d'images.

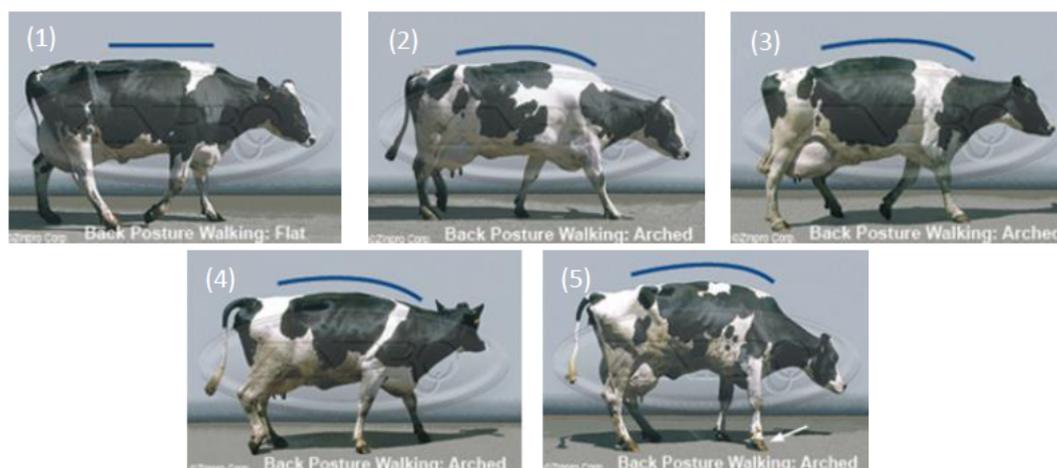


Fig 1.5: Echelle d'étiquetage pour la détection de la boiterie (1) : vache non boiteuse avec une ligne du dos horizontale. (2) : vache légèrement boiteuse présentant une légère courbure du dos. (3) : vache modérément boiteuse. (4) : vache boiteuse avec une voussure marquée du dos. (5) : vache boiteuse très voussée.

### 1.6.1 Les capteurs non invasifs

Se sont tous les capteurs qui ne sont pas placés à l'intérieur du corps de l'animal. Ils sont positionnés sur l'animal, intégrés dans un automate de traite, de pesée, d'alimentation ou placés dans les bâtiments d'élevage. Dans cette catégorie, on distingue des capteurs immobiles et des capteurs mobiles :

**a) Capteurs immobiles** Les animaux dans un bâtiment d'élevage ont généralement une routine quotidienne répétitive ; c'est-à-dire qu'ils se trouvent à des emplacements connus à des moments fixes (pendant la traite, l'alimentation, le repos, etc.). Par conséquent, les capteurs peuvent être placés à des endroits fixes où les animaux doivent régulièrement passer.

Un système de mesure de la composition respiratoire comme illustrer en Fig 1.6 peut être un exemple. Dans ce système, l'exhalation de l'animal passe de ses narines à travers des ouvertures en un ou plusieurs trajets d'écoulement. Ces trajets sont prévus pour le transfert de l'échantillon de l'appareil vers un moyen d'analyse [15]. Les caméras de surveillance sont un autre exemple de capteurs non invasifs immobiles. Lorsque la zone habituellement occupée par les vaches est dans le champ de vision d'une caméra, celle-ci peut continuellement fournir des informations sur les vaches dans le troupeau. La Fig 1.7 montre un exemple d'une image acquise lors d'une surveillance par caméra. L'analyse de cette image peut permettre la détection automatique de la boiterie chez les vaches [16].

En général, l'utilisation de capteurs immobiles dans l'élevage de précision est bénéfique

dans le sens où seulement un ou quelques capteurs sont nécessaires pour surveiller un troupeau. L'inconvénient est la limitation des ensembles de données disponibles du fait que ces capteurs sont peu nombreux, éloignés des animaux et permettent l'acquisition de données juste lorsque les vaches les traverses.

**b) Capteurs mobiles** Afin de surveiller un animal tout au long d'une journée et quel que soit l'endroit où il se trouve dans la ferme, le moyen le plus fiable est d'attacher des capteurs à cet animal (par exemple, par un collier de cou ou un ruban de cheville). Des capteurs typiques de ce type sont les capteurs de pression, les podomètres, les accéléromètres, les GPS, etc.

Récemment, des capteurs de positionnement GPS peu coûteux et sans infrastructure ont été utilisés pour identifier différents états de mouvement (manger, marcher, couché, debout, etc) des animaux. Ces GPS sont attachés aux colliers des animaux, comme le montre la Fig 1.8. Les performances de monitoring pourraient être augmentées de manière significative par une intégration de données de capteurs supplémentaires (par exemple, accéléromètres) [17].

Les accéléromètres sont actuellement des candidats plus prometteurs pour fournir des données fiables pour la surveillance des activités. Ils offrent un excellent compromis entre des conditions technologiques contradictoires comme le débit de données continuellement élevé et la minimisation de la consommation d'énergie.

L'utilisation d'accéléromètres dans le suivi des activités chez les vaches a été proposée il y a de nombreuses années. Les avantages récents de la conception des accéléromètres, notamment dans le domaine des accéléromètres à base de MEMS (systèmes micro-électromécanique), ont notablement amélioré la fiabilité des ensembles de données mesurées. Des résultats encore meilleurs sont réalisables lorsque l'on utilise plusieurs accéléromètres à différentes positions (par exemple pour différencier les mouvements de la tête de ceux du corps).

L'avantage des capteurs mobiles réside dans le fait qu'ils permettent une observation continue. Les inconvénients sont les difficultés liées à l'accès aux données du capteur ainsi qu'au risque de dommages des boîtiers de capteurs.

## 1.6.2 Les Capteurs invasifs

Des mesures de haute précision de certains paramètres physiologiques nécessitent le placement des capteurs à l'intérieur du corps de l'animal (par exemple dans le rumen ou sous la peau). Les capteurs typiques de ce type sont des thermomètres pour mesurer la température corporelle [18] ou bien les capteurs de mesure de la conductivité électrique et de la valeur du pH du fluide du rumen (Fig 1.9).

L'avantage des capteurs internes est la fiabilité des valeurs de mesure qui ne sont pas affectées par les conditions extérieures. Les inconvénients sont les difficultés de réutilisation

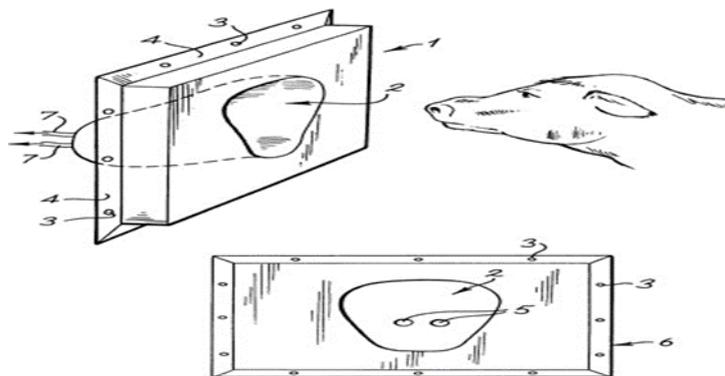


Fig 1.6: Système de mesure de la composition de la respiration. Où : 1 à 4 et 6 désignent des moyens de fixation, 5 indique des ouvertures pour narines et 7 représente le trajet d'écoulement



Fig 1.7: Exemple d'un enregistrement vidéo d'une vache quittant la salle de traite.



Fig 1.8: Boitier de positionnement GPS attaché à une vache.

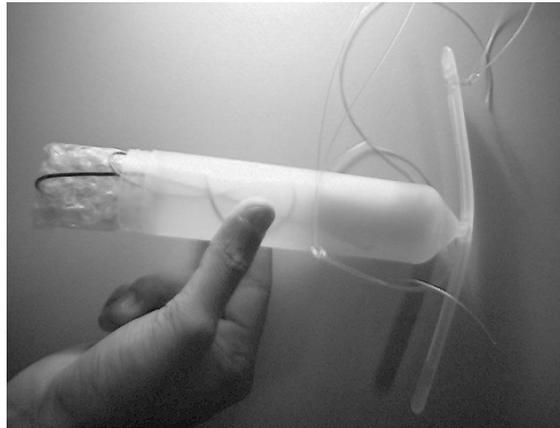


Fig 1.9: Capsule utilisée pour surveiller à distance la température intraruménale d'une vache sur une longue période de temps.

du capteur et le temps d'application limité des capteurs en raison de l'énergie requise.

## 1.7 Utilisation possible des dispositifs électroniques

Les différents dispositifs électroniques (capteurs) utilisés en élevage de précision mesurent différents paramètres et répondent à différents besoins d'allègement des charges de travail, d'identification individuelle des vaches, d'anticipation des problèmes sanitaires, de surveillance de l'activité, de l'alimentation, de la reproduction, de la morphologie, de la production, etc. Certains de ces capteurs mesurent un seul paramètre animal et permettent le suivi d'un seul événement. Par exemple, les capteurs de position de la queue sont utilisés uniquement pour le suivi des moments de vêlages. Cependant, d'autres capteurs mesurent plusieurs paramètres et permettent le suivi de plusieurs événements. Par exemple, les accéléromètres attachés aux colliers des vaches permettent de récolter des données d'activité et d'alimentation et assurent le suivi des événements reproductifs, sanitaires, etc. L'emplacement de ces capteurs doit être bien choisi en fonction des événements suivis. La Fig1.10. donne un panorama de capteurs disponibles sur le marché pour les bovins en spécifiant leurs emplacements, les paramètres mesurés et les événements suivis. Ci-dessous nous présentons quelques applications en élevage de précision et les principaux capteurs (outils) utilisés.

### 1.7.1 L'identification

L'identification est la première application de base qu'il faut assurer en élevage de précision. En effet, l'identification permet de gérer un grand nombre d'animaux, d'assurer la traçabilité des produits alimentaires d'origine animale [19], de fournir une meilleure preuve de propriété

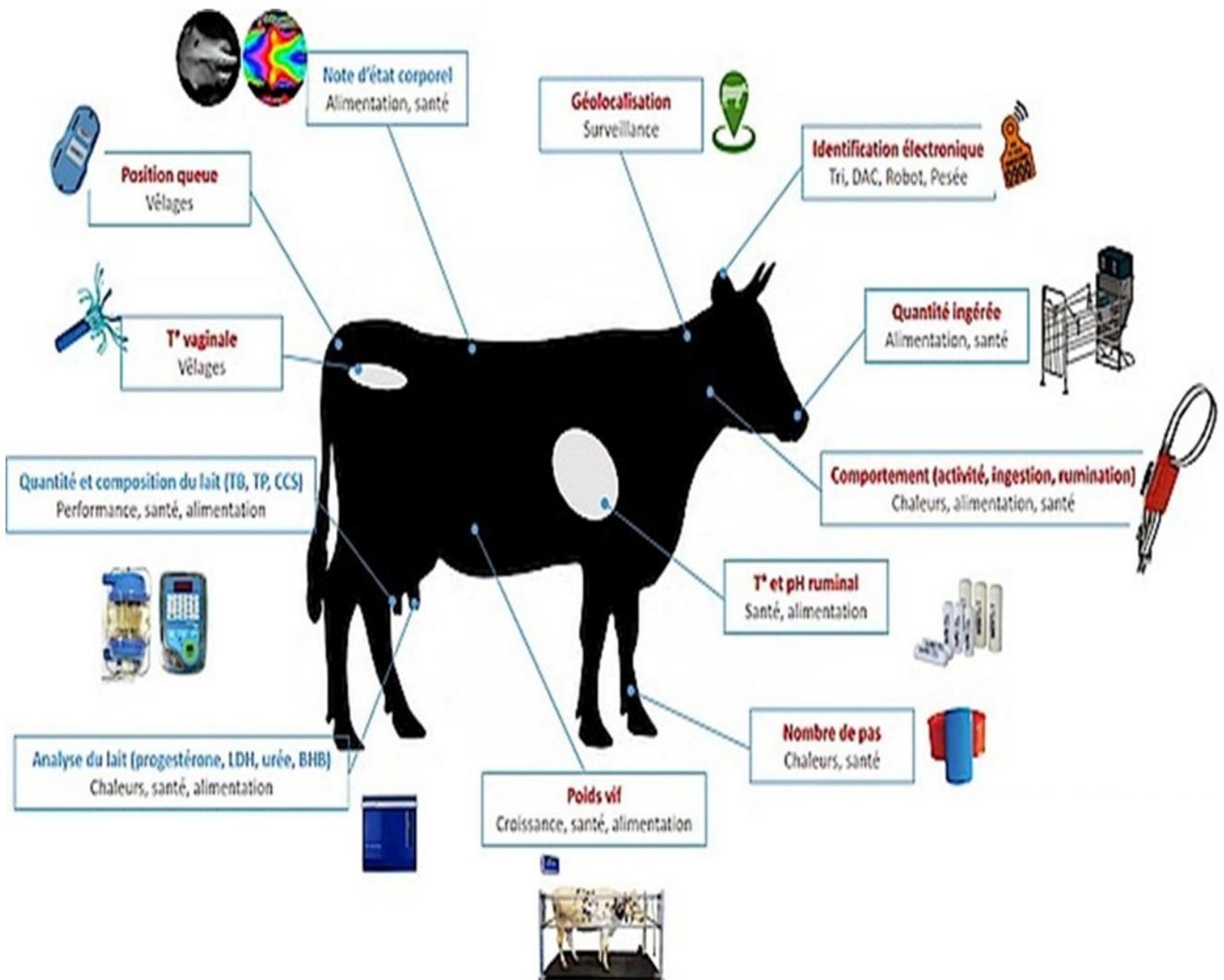


Fig 1.10: Panorama et emplacement des capteurs disponibles sur le marché pour les bovins.

afin de réduire les vols de bétail et améliorer la transparence des ventes. Plusieurs technologies d'identification sont utilisées pour assurer une identification individuelle et sans contact de vaches. Ces technologies peuvent être classées en trois groupes à savoir l'identification à liaison optique, l'identification à liaison RFID (radio fréquence) et l'identification biométrique [20].

Les tatouages et les marques auriculaires sont utilisés comme outils d'identification à liaison optique. En effet, ces outils permettent aux éleveurs d'identifier leurs vaches manuellement. Par exemple, lors de l'utilisation des marques auriculaires, chaque vache possède son propre code-barre et ce code est utilisé pour la différencier des autres vaches. Ces outils ne sont pas suffisamment fiables pour l'identification des animaux car ils sont très susceptibles d'être détruits ou arrachés [21].

L'identification par RFID a fait ses preuves en tant que solution courante pour l'identification individuelle dans les fermes [22]. Les systèmes RFID se composent principalement de deux dispositifs. Une étiquette (Tag) qui renferme une puce électronique qui contient des informations sur l'animal et un lecteur permettant de lire ces informations et qui communique avec le Tag par ondes radio [23]. Les étiquettes sont généralement attachées aux oreilles des animaux (Fig1.11-a), comme elles peuvent aussi être attachées aux pattes (Fig1.11-b), insérées dans un collier attaché au cou (Fig1.11-c) ou bien injectées dans le corps des vaches sous forme de bolus (Fig1.11-d). Le lecteur est souvent intégré dans des équipements tels que les robots de traite, les distributeurs d'aliments, les portes de tri, etc. L'identification RFID offre plusieurs atouts car elle assure essentiellement une meilleure lecture à distance et une vitesse de lecture plus rapide. Cependant, l'installation et l'utilisation des systèmes RFID peuvent causer du stress et affecter le bien-être des animaux. Par exemple, les RFID attachés aux oreilles doivent être fixés par une piqûre (Fig1.11-a). De plus, les vaches peuvent endommager les Tag RFID lors du toilettage par exemple. L'endommagement du dispositif ou sa perte peut entraîner un coût économique.

Les techniques d'identification biométrique reposent sur le fait que certaines caractéristiques biométriques sont uniques pour chaque vache et peuvent être utilisées comme empreintes d'identification [20, 21]. Par exemple, l'identification des vaches par le museau est basée sur le fait que les caractéristiques du museau de chaque vache sont uniques. Un autre exemple est que les systèmes d'identification basés sur la rétine utilisent le fait que le vaisseau rétinien de chaque vache est un identifiant unique. Une autre méthode basée sur la biométrie comprend la reconnaissance faciale des animaux à l'aide de caméras. Ces techniques sont apparues comme une tendance prometteuse pour remplacer l'identification RFID. Cependant, elles posent des défis en termes de précision de l'identification, de temps de traitement et d'opérabilité globale du système.



Fig 1.11: Différents emplacements des systèmes d'identification RFID

## 1.7.2 La surveillance du comportement

La surveillance des comportements du bétail est utile pour détecter les événements de reproduction et sanitaires et faciliter la gestion du troupeau. Cependant, la surveillance manuelle par des observations prend du temps, demande beaucoup de travail et parfois l'accès à l'animal n'est pas facile, comme en milieu montagneux. Pour ces raisons, les applications qui impliquent des capteurs pour surveiller automatiquement le comportement du bétail ont connu une croissance rapide ces dernières années [24]. Pour une application efficace en élevage, les capteurs doivent fournir des informations sur un large éventail de comportements [11]. Les activités de l'animal (debout, couché, marche, etc.) et les comportements alimentaires (manger, pâturage, etc.) sont tous deux essentiels pour détecter les problèmes de santé et faciliter la gestion des vaches laitières.

### 1.7.2.1 Suivi du comportement alimentaire

La surveillance du comportement alimentaire (mastication, rumination, alimentation) donne aux éleveurs plusieurs bénéfices tels qu'une détection précoce de trouble de santé, une sélection d'animaux plus efficaces, une rentabilité d'élevage, une optimisation de l'utilisation des ressources alimentaires et un meilleur pilotage de l'alimentation. Plusieurs outils ont été développés pour assurer le suivi du temps de rumination et d'ingestion des animaux et de signaler toute anomalie. Ces outils se basent sur des capteurs de pression, des accéléromètres, des microphones, etc.

- Capteur de pression

Un système de mesure automatique des mouvements de rumination par capteur de pression (Fig 1.12) chez les vaches a été développé par la station de recherche Agroscope Reckenholz Tänikon (ART). Ce système permet d'identifier trois types d'activité qui sont la rumination, l'alimentation, le pâturage [25]. Le capteur de pression est placé dans une enveloppe protectrice au-dessus du museau de la vache. Il est constitué d'une unité de mesure de pression et d'une unité d'enregistrement de données.

- Accéléromètre

Grâce à la miniaturisation des accéléromètres et à leur faible coût et leur capacité potentielle à enregistrer des données comportementales à haute résolution, plusieurs études ont utilisé des accéléromètres dans les sciences animales [26–29]. Selon ces études, plusieurs méthodes de classification peuvent être utilisées pour classer les activités des vaches afin de distinguer les activités à haute ou basse énergie. Les activités en haute énergie peuvent être utilisées pour le



Fig 1.12: Capteur de pression assurant le monitoring de la rumination.

suivi de l'alimentation des vaches. En effet, les capteurs basés sur un accéléromètre attaché au cou des vaches permettent d'enregistrer une accélération triaxiale en continu. L'orientation des trois axes d'accéléromètre change en fonction des comportements qu'effectue la vache. Un changement dans l'accélération de  $x$  correspond à un mouvement latéral vers la gauche ou vers la droite. Un changement dans l'accélération de l'axe  $y$  mesure les mouvements d'avancement ou de recul. Tandis que, les changements dans l'accélération de l'axe  $z$  mesurent la rotation du cou. La société Agis Automatisering commercialise un système nommé SensOor qui est constitué d'un accéléromètre tridimensionnel attaché à l'étiquette d'oreille (RFID). Dans SensOor l'accéléromètre surveille les mouvements de l'oreille et permet entre autre de surveiller l'alimentation et la rumination des vaches [30].

- Microphone

De nombreux systèmes automatisés ont été développés et commercialisés (Heatime-HR , Qwes-HR) pour la surveillance du comportement de rumination des bovins au pâturage à l'aide de microphone. Dans [31] un système utilisant un microphone acoustique a été développé. Ce système se compose d'une unité d'enregistrement (Fig. 1.13-a) et d'un microphone (Fig. 1.13-b). Il est monté sur un licol, l'enregistreur est placé derrière la tête de l'animal et le microphone est fixé à la sangle avant du licol près de sa bouche. Les données sonores sont enregistrées sur une carte mémoire SD, puis une analyse du signal acoustique pour mesurer des événements digestifs a été réalisée.

### 1.7.2.2 Suivi de l'activité

La surveillance de l'activité est réalisée en utilisant des dispositifs qui sont souvent attachés aux pattes ou à la tête des vaches. Des capteurs de mouvement (accéléromètre et inclinomètre) ou de localisation (GPS) sont utilisés pour détecter les activités quotidiennes des vaches (de-

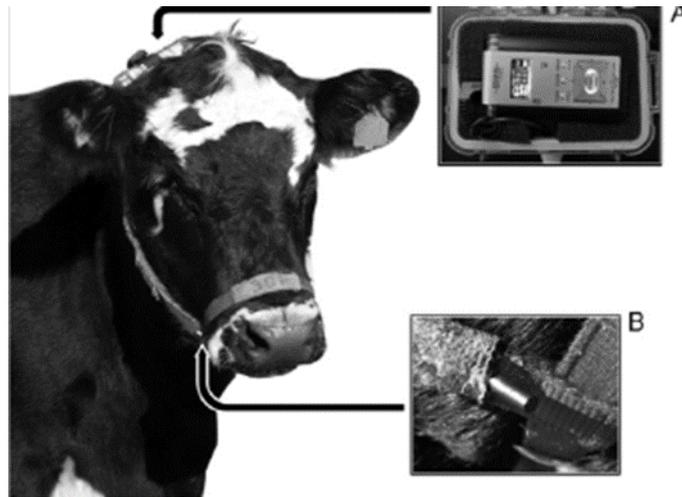


Fig 1.13: Système de surveillance de la rumination par un capteur acoustique.

bout, couchées, marcher, etc.). Le temps que les vaches laitières passent dans ces différentes activités est un indicateur important de leur état de santé et du bien-être. Par exemple, les temps de couchage élevés ( $>14,5$  heures/jour) sont un indicateur d'un risque accru de boiterie [32]. Selon [33], les vaches peuvent présenter jusqu'à 40 comportements individuels. Dans ce qui suit nous allons expliquer comment les différents dispositifs de mouvement ou de localisation sont utilisés pour assurer le suivi des activités des vaches.

- Capteur de mouvement

Des accéléromètres et les inclinomètres peuvent être fixés aux pattes des bovins pour mesurer l'orientation. Par exemple, lorsque la patte est horizontale, l'animal est très probablement couché. Ces dispositifs sont révélés utiles lors de leurs utilisations dans les fermes. Par exemple, les auteurs de [32] ont attaché des accéléromètres à plus de 2 000 vaches dans 43 fermes laitières et ont mesuré le temps pendant lequel les vaches se couchaient. Les accéléromètres attachés aux pattes peuvent également mesurer les caractéristiques d'accélération associées à chaque pas effectué par les vaches [34]. Des appareils commerciaux (par exemple IceTag® de Ice Robotics Inc, Edimbourg, Royaume-Uni) utilisant ce principe de fonctionnement sont largement utilisés dans les fermes. Actuellement, les Unités de Mesure Inertielle (IMU) sont largement utilisées pour collecter les paramètres qui permettent la classification des activités [35, 36]. Elles mesurent les accélérations linéaires et angulaires et estiment la trajectoire, ce qui élargit les données utilisées. Ces IMU sont composées d'un accéléromètre 3 axes, d'un gyroscope 3 axes et d'un magnétomètre 3 axes.

- Capteur de localisation

La surveillance du comportement à l'aide de Système de Positionnement Global (GPS) peut fournir des informations supplémentaires sur des comportements spécifiques du bétail. Les GPS sont des dispositifs largement utilisés pour diverses applications de localisation. La principale raison d'envisager le GPS pour la surveillance des vaches est que cet outil de positionnement est sans infrastructure locale. Cependant, lors de l'utilisation des autres alternatives comme les réseaux de capteurs, une infrastructure peut être nécessaire pour la communication. Les auteurs de [37] ont utilisé un GPS pour collecter des données permettant de classer le comportement du bétail en pâturage, repos ou marche. La fusion des données GPS avec des données d'accéléromètre permet d'augmenter le nombre de classes et les performances de la classification [38].

Les informations offertes par les outils de suivi des comportements et d'identification peuvent être utilisées directement par les éleveurs pour assurer le monitoring de leurs vaches [39]. En effet, les éleveurs peuvent consulter les variations quotidiennes de comportement de chaque vache ou consulter l'historique comportemental de ces vaches afin de prendre des décisions sur l'état des vaches. De plus, ces outils d'identification et de suivi du comportement peuvent être intégrés dans d'autres systèmes plus complexes pour créer des modèles prédictifs ou décisionnels permettant le suivi sanitaire, de reproduction et de morphologie [39].

### 1.7.3 Surveillance de la santé animale

La santé des vaches laitières est multifactorielle et complexe. De nouvelles maladies qui menacent la santé apparaissent chaque année. La détection des maladies à un stade précoce nécessite des tests de laboratoire longs et coûteux. Les capteurs ont le potentiel de résoudre ces problèmes en développant des outils de diagnostic pour la détection rapide des principales menaces sanitaires. L'utilisation de capteurs devient alors de plus en plus importante pour la gestion de la santé des vaches laitières. Au lieu de s'appuyer uniquement sur les sens et les connaissances des éleveurs, les capteurs sont un outil de soutien incontournable pour le diagnostic des maladies. Ces systèmes permettent de collecter automatiquement des données physiologiques (telles que la composition de la sueur et la température du corps), des données de l'environnement de l'exploitation (telles que l'humidité et la température ambiante), des données de production (telle que la quantité de lait produite) et des données comportementales (telles que les activités et l'alimentation), etc. Ces données collectées seront ensuite exploitées afin de détecter la présence de maladies et de fournir des alertes sanitaires aux éleveurs.

Il existe de nombreux types d'outils qui ont été déployés en pratique pour une surveillance efficace de la santé des animaux [14, 40]. Des capteurs sont utilisés pour assurer la détection automatisée des changements dans les paramètres liés à un événement de santé. Les principaux paramètres couramment utilisés dans le suivi sanitaire sont la température corporelle,

Symptôme physiologique	Changement comportemental	Capteur
-Déformation des onglons. -Erosion des talons.	Debout ou assis Moins de pâturage	Accéléromètre (cou, pieds)
	Voussure dorsale anormale	Caméra (emplacement externe)
	Répartition non uniforme du poids	Capteur de charge (sous pieds)

Tableau 1.1: Les outils de monitoring sanitaire des vaches boiteuses.

la température du rumen, la fréquence cardiaque, l'activité, l'alimentation, la rumination, le poids vif, le pH du fluide de rumen, la production laitière et la composition du lait.

Le tableau 1.1 donne un aperçu sur les différents capteurs utilisés dans la surveillance de la boiterie par exemple ainsi que les modifications comportementales et physiologiques associées à son apparition. Le principe est que les données de capteurs (accéléromètre, caméra, capteur de charge) sont récoltées en temps réel et injectées dans des modèles d'analyse prédictifs. Ces modèles permettent d'exploiter les données et suivre les modifications comportementales (position debout et couchée,...) afin de donner des prédictions sur l'état de la boiterie des vaches. Le même principe s'applique au suivi des autres maladies. Par exemple, La Fig 1.14 montre un système de collection de la température corporelle des vaches. Les données de ce système sont transmises à une borne où s'exécutent des modèles prédictifs permettant le suivi de la fièvre des vaches. Dans le chapitre 2 nous allons décrire comment les maladies les plus fréquentes impactent les activités et le comportement alimentaire des vaches.

#### 1.7.4 La surveillance de la reproduction

Une diminution des performances de reproduction au sein d'un élevage représente des pertes économiques directes (perte de production de lait, mortalité de veaux et les réformes subies) et indirectes importantes (inséminations supplémentaires, frais vétérinaires, travail supplémentaire). Plusieurs outils ont été développés pour le suivi de la reproduction afin de diminuer ces pertes et assurer le renouvellement des troupeaux. Ces outils permettent une maîtrise accrue de la reproduction, en tant qu'élément essentiel de l'élevage et ils permettent à l'éleveur de se décharger d'une partie de ses activités. La surveillance de la reproduction peut être assurée par la détection du vêlage ou des chaleurs.



Fig 1.14: Capteur de température attaché à la queue permettant le monitoring de la fièvre.

#### 1.7.4.1 La détection du vêlage

La prédiction du temps de vêlage (le vêlage est la mise bas chez les bovins) est un élément clé de l'élevage. Elle est utilisée pour décider s'il faut déplacer une vache à la maternité et décider si une surveillance humaine est nécessaire [1, 41]. Lors du vêlage, jusqu'à 51 % de bovins laitiers nécessitent une assistance humaine. Chez les bovins allaitants, près de 50 % de pertes de veaux se situent autour du vêlage (au vêlage et pendant les 2 jours qui suivent) [42].

La détection automatisée du vêlage assure la viabilité des veaux et réduit les risques des dystocies (difficulté en vêlage qui nécessite intervention humaine) [43, 44]. Plusieurs outils ont été développés, ils sont basés essentiellement sur la mesure des modifications physiques [45, 46] ou des changements dans les contractions musculaires. Plusieurs types de ces dispositifs sont actuellement sur le marché. Ils se basent sur des inclinomètres et des accéléromètres qui détectent la montée de la queue et les changements de comportements [47], des ceintures abdominales qui surveillent les contractions utérines [41] (Fig 1.15), des thermomètres intravaginaux détectant une chute de la température corporelle [48], des équipements destinés à détecter l'expulsion des veaux [49], des outils basés sur le temps de rumination [50, 51] et l'alternance de la position debout et couchée [52].



Fig 1.15: Vache équipée de la ceinture Agribelt de Databel agrimonitor.

#### 1.7.4.2 La détection de chaleurs

Après le vêlage, 45 jours sont nécessaires pour obtenir une involution utérine complète. Pendant cette période, les vaches doivent avoir manifesté une reprise de l'activité ovarienne. La mise à la reproduction peut s'envisager dès le 50<sup>ème</sup> jour et l'insémination fécondante doit avoir lieu avant le 90<sup>ème</sup> jour. Pour ce fait, l'éleveur n'a que 40 jours pour s'assurer que les vaches sont fécondées. La chaleur est le comportement particulier d'une femelle correspondant à la période appelée œstrus. Pendant ce temps, la vache accepte de s'accoupler et peut être fécondée. Pour déterminer le meilleur moment pour une insémination, il faut bien reconnaître les signes de la chaleur.

La détection des chaleurs est d'une importance capitale en élevage car elle influence le nombre d'inséminations et le temps qui s'écoule entre les vêlages [53]. Plusieurs recherches ont été faites pour l'automatiser en se basant sur les changements comportementaux et physiologiques des vaches en période de chaleur [54, 55].

Le changement du comportement des bovins laitiers en période de chaleur se manifeste par l'acceptation du chevauchement (une vache est considérée comme acceptant le chevauchement si elle reste statique au moins deux secondes, sans être bloquée, sous l'une de ses congénères qui la surmonte) et l'augmentation de l'activité physique. Pour ce fait, plusieurs outils de détection de chaleur se basent sur la surveillance du comportement tel que les détecteurs de chevauchement (HeatWatch 2 CowChips), les détecteurs d'activité basés sur l'utilisation de podomètres (Rescounter II, HeatSeeker 2, Podomètre Fullwod Packo) ou basés sur l'utilisation d'accéléromètres (Fig 1.16) (HeatPhone, Heat Box, Heatime) ont été développés [43, 56]. La vidéo surveillance est également utilisée pour permettre à l'éleveur une surveil-

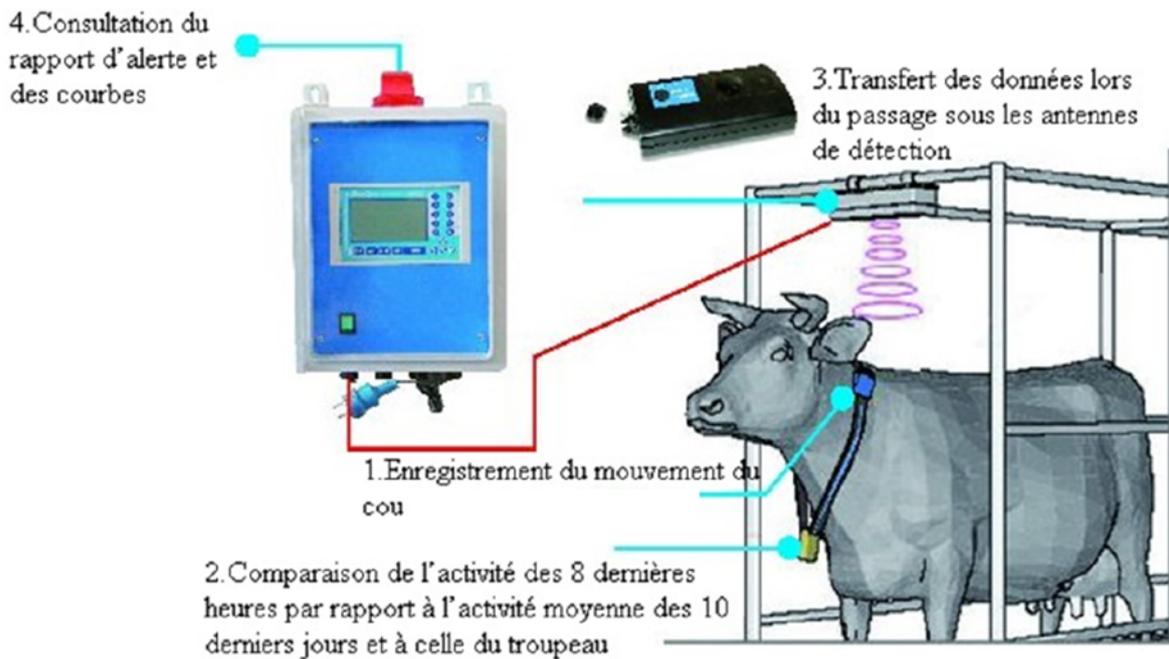


Fig 1.16: Principe de fonctionnement du système Heatime.

lance en continu de l'activité des vaches. Tous les comportements des vaches en chaleur peuvent être enregistrés en vidéo. Cette méthode ne permet pas encore une détection automatique des chaleurs, mais elle la facilite [57].

Les changements physiologiques sont aussi utilisés comme des indices pour la détection des chaleurs; une variation de la conductivité du lait indique une chaleur probable de l'animal. Le taux de progestérone du lait est un paramètre de référence pour encadrer la période des chaleurs. Herd Navigator est un outil d'analyse de lait, il permet une détection de vaches en chaleurs avec une précision de 95 % [58]. Les chaleurs peuvent aussi être détectées grâce à la température corporelle qui baisse 2 jours avant le début des chaleurs [57].

### 1.7.5 La surveillance de la morphologie

La morphologie a toujours été considérée comme un indicateur de la performance et de la bonne condition physique des animaux. Le suivi de la morphologie diffère en fonction des espèces animales. Par exemple pour les chevaux la morphométrie permet de qualifier la morphologie du cheval de manière très objective.

Pour le bétail, le suivi de la morphologie est désigné par le suivi du poids vif et de l'état corporel.



Fig 1.17: Système de mesure de distance entre la thoracique gauche et droite de la poitrine pour estimer le poids vif.

#### 1.7.5.1 Suivi du poids vif

Déterminer la valeur du bétail en industrie dépend de leur poids vif. Les éleveurs essaient de mesurer le poids des bovins en conformité avec la norme afin de les vendre à un bon prix. Le poids vif dépend de divers facteurs tels que la longueur du corps, la hauteur au garrot, la profondeur de la poitrine et l'âge. Le poids vif peut être mesuré de façon directe en utilisant des dispositifs de pesée installés sur les chemins que traversent quotidiennement les bovins, ou d'une façon indirecte par des techniques d'analyse de régression [53, 59, 60] et d'analyse d'images [61–63].

Les relations existantes entre le poids vif et certaines mesures corporelles telles que la hauteur du garrot, la longueur du corps et le tour de poitrine ont été démontrées par des techniques de régression [64]. La relation trouvée entre le poids vif et le tour de poitrine a atteint un taux de corrélation de 95,8% [59]. Grâce à ces corrélations, un système d'estimation du poids vif par une méthode de mesure indirecte a été développé (Fig 1.17). Ce système utilise la distance entre la thoracique gauche et droite de la poitrine pour estimer le poids vif en kg [60].

La mesure du poids vif par pesée nécessite la mise en place d'un lieu de pesage dans les fermes ou le placement d'une plate-forme de pesage dans les zones traversées par le bétail. Cependant, il peut y avoir des problèmes dans le fonctionnement des outils de pesage qui sont exposés aux conditions environnementales. Et parfois, le bétail pourrait être exposé à de petites blessures au cours de la pesée. Par conséquent, les éleveurs ne préfèrent pas allouer de l'espace, du personnel ou des ressources financières à des outils de pesage [61]. Grâce au développement des techniques de traitement d'images plusieurs systèmes ont été développés pour mesurer le poids vif à l'aide d'analyse d'images [62, 63]. Plusieurs compagnies utilisent

différentes technologies d'imagerie pour prédire le poids des vaches à l'aide de caméra 2D ou 3D (ClicRweight de ClicRTechnologies<sup>TM</sup>, optiSORT et optiSCAN et opticow de Hölscher-Leuschner et eYeScan de Fancom)

### 1.7.5.2 Suivi de l'état corporel

L'utilisation des réserves corporelles joue un rôle important dans la gestion de l'énergie des bovins. Ces réserves permettent aux vaches de produire beaucoup de lait en début de lactation et de manière continue. La mesure de l'état des réserves corporelles et plus précisément de leur variation est aussi un indicateur très intéressant du point de vue de l'alimentation, de la fertilité, de la santé et du bien-être animal. Pour les bovins laitiers, lorsque la production laitière atteint son pic, les réserves corporelles sont au plus bas, puis la baisse de production s'accompagne d'une reconstitution des réserves corporelles jusqu'au prochain vêlage. Cette utilisation accrue de l'énergie pour la production de lait peut cependant affecter les fonctions de santé et de reproduction. L'état corporel est étroitement lié à l'état des stocks de graisse. Les variations du poids vif (jusqu'à 18 %) ne reflètent qu'imparfaitement celles des réserves corporelles. L'estimation de l'état corporel est indispensable pour l'ajustement de l'alimentation et le suivi de la production, de la santé, de la nutrition et du bien-être des animaux.

L'estimation de l'état corporel en élevage se fait à l'aide de grilles de notation qui ont été mises en place selon l'espèce animale [65]. Plusieurs outils ont été développés dans le cadre de l'élevage de précision pour permettre une annotation automatique de l'état corporel. Ces outils se basent sur la technologie d'ultrason et de traitement d'images.

Les approches se basant sur la vitesse du son (dite Velocity Of Sound (VOS) en anglais) ont été développées pour évaluer les aspects de la composition corporelle des vaches [66]. VOS a un potentiel élevé en tant qu'outil de mesure de la composition des animaux vivants. Les principaux avantages de VOS sur les autres techniques d'ultrason sont : son objectivité et la livraison immédiate d'un résultat [67]. La mesure à l'aide des ultrasons des épaisseurs du gras dorsal présente un intérêt significatif dans la prédiction de la composition en gras et en muscles chez l'agneau [68].

Le développement des nouvelles technologies en imagerie offre de nouvelles perspectives intéressantes pour le calcul de la Note d'Etat Corporel (NEC) et le remplacement des observations visuelles par grille de référence (chère et peu sensible) par des méthodes d'analyse d'images 2D ou 3D [69–72]. L'idée est de mettre en relation un indicateur de la forme, majoritairement du bassin de la vache, avec la note de l'état corporel. Une étude [72] a évalué un système d'analyse d'images qui pouvait déterminer de manière objective le degré d'engraissement et le score d'état corporel chez les vaches laitières en utilisant des données de caméra 3D.

## 1.8 Impact de l'élevage de précision sur l'économie

Avec le besoin croissant de développement d'outils permettant d'aider l'éleveur dans la gestion de sa ferme et à prendre des décisions, l'analyse économique de ces outils s'avère être une phase cruciale du fait qu'un système à bas coût et avec une grande utilité a plus de chance d'être adopté dans plusieurs fermes [73]. La réussite du déploiement d'un système est liée à la présence d'informations sur son impact économique futur [74]. Actuellement, l'utilisation de la plupart des systèmes dans les fermes commerciales est lente. Cela est dû au fait que ces systèmes produisent de grandes quantités de données mais ils fournissent peu ou pas de conseils pour la prise de décision. Néanmoins, l'élevage de précision donne plusieurs atouts économique à la filière de l'élevage; ces impacts positifs sont :

1. Un rapport coût-bénéfice positif : les technologies d'élevage de précision permettent de récompenser le coût de l'investissement économique dû à l'achat de capteurs et d'outils. En effet, les profits liés aux pertes sanitaire et productif évitées grâce à ces technologies permettent de surpasser les coûts de l'investissement [75].
2. Une meilleure rentabilité et la réduction des coûts : les technologies d'élevage de précision permettent aux éleveurs de prendre des décisions plus rapides et mieux réfléchies, ce qui engendre une meilleure productivité et une meilleure rentabilité [76].
3. Une amélioration de l'efficacité de production et de la qualité des produits : les outils de monitoring déployés dans les fermes commerciales permettent essentiellement d'assurer l'augmentation de la taille du troupeau sans engendrer une augmentation de la main d'œuvre. Ces outils permettent aussi la simplification de la gestion du troupeau, un contrôle accru de la production [77] et de remplacer la main d'œuvre non qualifié [78].
4. Une meilleure commercialisation et traçabilité des produits de l'élevage : l'élevage de précision permet d'assurer l'échange d'informations tout au long de la chaîne alimentaire animale. Cela permet aux fournisseurs d'aliments pour animaux et d'intrants alimentaires d'améliorer la composition de leurs produits. La traçabilité permet aussi aux éleveurs la sélection du bon aliment (ou du bon fournisseur d'aliments). Les abattoirs peuvent utiliser le système comme base de coopération avec les exploitations pour produire et s'approvisionner en plus d'animaux selon les spécifications de poids et de conformation [73].
5. Une réduction du commerce illégal des produits de l'élevage : l'élevage de précision peut également contribuer à éviter le commerce illégal de bétail et de produits du bé-

tail. Etant donné que la contrebande d'animaux est un problème majeur (sanitaire et financier) dans plusieurs pays [73].

## 1.9 Impact de l'élevage de précision sur l'environnement

Plusieurs recherches [10, 79] ont estimé que les outils développés en élevage de précision peuvent être utilisés comme un moyen précieux pour évaluer et réduire les effets environnementaux des élevages. Ces outils permettent d'améliorer la qualité d'air [80] présent dans les élevages en réduisant les émissions d'ammoniac [81] et la présence de poussières [82].

L'alimentation est l'un des paramètres les plus importants dans l'élevage. Cependant, la mauvaise maîtrise du processus d'alimentation conduit à des conséquences négatives sur l'environnement telles que une augmentation des émissions de gaz à effet de serre (notamment le CH<sub>4</sub> [83]), une plus forte utilisation des ressources environnementales à cause de la suralimentation des animaux [84], le gaspillage ultérieur de nutriments dans l'environnement [85], etc. Les technologies d'élevage de précision permettent d'assurer le suivi de l'alimentation des vaches et l'étude des émissions liées à chaque aliment. Par exemple, l'amélioration de la qualité de l'ensilage permet de réduire de 11% l'émission de CH<sub>4</sub> [83] et l'utilisation de systèmes d'alimentation automatiques permet d'ajuster la quantité d'aliment servis et de répondre plus précisément aux besoins en nutriments de chaque animal [84]. De ce fait, la gestion et le contrôle continu de l'alimentation peuvent constituer la stratégie la plus appropriée pour réduire l'impact sur l'environnement.

L'impact environnemental peut être également réduit en limitant les émissions indésirables qui peuvent se produire lorsque les animaux sont confrontés à des problèmes de santé et de stress [86]. En effet, le suivi des événements sanitaires [87] et des troubles reproductifs [88] permettent non seulement de soulager et d'améliorer le processus du traitement mais aussi de réduire les effets environnementaux des troubles reproductifs et sanitaires. En effet, l'amélioration de la santé animale et la réduction de la mortalité animale peuvent réduire l'émission de plusieurs gaz tels que le protoxyde d'azote (N<sub>2</sub>O) et le CH<sub>4</sub> [89]. Dans [88], les auteurs ont montré qu'un meilleur contrôle sanitaire et reproductif pourrait réduire les émissions de Gaz à Effet de Serre (GES) issues d'élevage de 0,9 %. Selon [90] le monitoring de la boiterie peut entraîner une réduction de ses impacts environnementaux qui sont estimés à plus de 7 à 9 % pour le réchauffement climatique, l'acidification, l'eutrophisation et l'épuisement des combustibles fossiles. De plus, la diminution des taux de mammites cliniques (de 25 à 18%) et subcliniques (de 33 à 15%) permet une réduction d'épuisement des ressources abiotiques de 5,8% [91].

## 1.10 Impact de l'élevage de précision sur les éleveurs

L'élevage de précision a plusieurs impacts sur le travail des éleveurs. Ces impacts peuvent être positifs ou négatifs :

1. Le temps de travail : l'élevage de précision permet d'automatiser plusieurs tâches, ce qui conduit à une diminution dans le temps de travail de l'éleveur. Il permet aussi d'aider l'éleveur dans sa prise de décision et de lui fournir un meilleur contrôle de l'élevage à l'aide de détections précoces des problèmes sanitaires. Mais d'un autre point, l'éleveur doit consacrer plus de temps à gérer le matériel et d'autres problèmes en cas de panne ou de mauvais paramétrage du logiciel [92].

2. La productivité de l'élevage : l'élevage de précision permet d'augmenter la taille des troupeaux et de gérer plus d'activités sans augmenter la main d'œuvre. Mais il engendre aussi une grande baisse de la productivité en cas de dysfonctionnement.

3. L'astreinte : l'élevage de précision permet de réduire les tâches d'astreinte telles que la distribution de l'alimentation et la traite. Mais il engendre d'autres tâches comme la connexion permanente avec la ferme et la gestion des alarmes pendant la journée et la nuit.

4. La pénibilité physique du travail : l'éleveur est remplacé par des automates qui accomplissent le travail pénible et répétitif, ce qui offre un meilleur confort à l'éleveur. Mais la complexité de maîtriser les nouvelles technologies et l'augmentation de risques de pannes sont des facteurs qui peuvent engendrer une grande pénibilité physique chez les éleveurs.

5. La charge mentale des éleveurs : l'anticipation des décisions par le système d'élevage de précision permet de réduire les charges mentales des éleveurs. Mais, les alarmes qui peuvent surgir à tout moment causent d'autres problèmes mentaux.

6. L'adoption de nouvelles compétences : l'élevage de précision permet à l'éleveur d'acquérir de nouvelles connaissances en ce qui concerne la manipulation des nouvelles technologies (gestion et traitement de données, gestion d'automates) et permet aussi de remplacer la main-d'œuvre qualifiée. En effet, les outils PLF sont conçus de façon à faciliter leurs intégrations en ferme et à simplifier leurs usages par le personnel de la ferme. De plus, des formations d'initiation à la manipulation des outils sont souvent disponibles.

7. La relation homme animale : l'élevage de précision diminue le temps que consacre l'éleveur pour observer le comportement, la santé et le bien-être des animaux. Ce qui engendre une limitation des contacts entre l'homme et l'animal.

## 1.11 Conclusion

L'élevage de précision est une solution très importante pour les fermes, car il permet un monitoring continu des animaux dans le but d'accroître l'efficacité et la durabilité des élevages. Cependant, même si l'utilisation des systèmes d'élevage de précision donne des effets positifs évidents, des améliorations de la technologie et de la précision sont constamment nécessaires. Les systèmes d'identification et de monitoring de comportements (activité et alimentation), de santé, de morphologie, production et de reproduction sont progressivement déployés dans les fermes. Ces systèmes permettent une gestion continue et entièrement automatisée du cheptel et du bâtiment d'élevage. L'élevage de précision garantit aussi le bien-être des animaux et de l'éleveur.

Dans ce chapitre nous avons tout d'abord défini un ensemble de concepts relatifs à l'élevage de précision. Ensuite, nous avons donné un panorama de capteurs utilisés, de leurs emplacements (attachés aux différentes parties du corps de la vache ou placés à d'autres endroits dans la ferme) ainsi que les utilisations possibles de ces capteurs comme l'identification, le suivi d'activité, le suivi alimentaire, le suivi sanitaire, le suivi de la reproduction, etc. Après ces différents points, nous avons conclu que :

- L'identification individuelle est très importante et doit être envisagée avant toute autre application du monitoring de vaches laitières comme le suivi sanitaire, le suivi du comportement alimentaire, le tri sélectif, etc.
- Le suivi comportemental (activité ou alimentaire) est un bon indicateur de l'état de santé de vaches laitières. En effet, des comportements inhabituels permettent une détection précoce de troubles sanitaires avant même l'apparition de symptômes physiologiques.
- Les accéléromètres et les IMU sont des capteurs très fiables pour le suivi des activités animales. Cependant, le choix de l'emplacement de ces capteurs sur le corps de l'animal est d'une importance capitale. Dans le cadre de notre travail, nous avons exploré un nouvel emplacement qui est le dos des vaches pour assurer un suivi de leurs activités.
- Le comportement alimentaire a été largement considéré par des chercheurs. Néanmoins, ces chercheurs ont utilisé des capteurs attachés comme les accéléromètres ou les GPS qui peuvent être source de stress chez les vaches laitières. Pour remédier à ce facteur de stress, nous avons utilisé des caméras.

Dans le cadre de cette thèse nous utilisons des capteurs non invasifs (mobiles et immobiles) pour assurer le suivi comportemental (activité et alimentation) de vaches laitières. Les résultats de ce suivi comportemental peuvent être exploités pour la détection de manière précoce

des problèmes de santé chez les vaches. En effet, il existe des corrélations entre les altérations comportementales et les troubles sanitaires. Cependant, les vaches peuvent présenter jusqu'à 40 comportements individuels et les maladies n'influencent que quelques comportements majeurs. Dans cet ordre d'idées, le deuxième chapitre donne les comportements majeurs qui sont impactés par les maladies et qui doivent être suivis par les outils développés dans le cadre de notre thèse.

# Chapitre 2

## Le suivi automatisé de la santé des vaches laitières

### 2.1 Introduction

Les fermes laitières ont évolué et deviennent des entreprises à but lucratif. Les principales raisons de ce développement sont la concurrence et la pression du marché dans l'optique d'une réduction significative du coût du lait liquide. Ce changement a conduit à une mise à niveau majeure de la technologie et des équipements pour réduire les dépenses, les coûts de main-d'œuvre et hausser simultanément les bénéfices. D'une manière conventionnelle, un éleveur expérimenté s'occupe d'un nombre relativement peu élevé de bétails, alors que dans les fermes modernes et automatisées très peu d'ouvriers prennent en charge un grand nombre de bovins. Cela crée un besoin accru de développer des systèmes de monitoring de la santé des animaux. Ces systèmes permettent de prévoir la présence de problèmes de santé afin d'éviter leur propagation et de diminuer leur impact significatif sur la rentabilité de l'élevage [14].

En Algérie, l'élevage bovin représente 50 % de la valeur ajoutée agricole. Cependant, la prévalence des maladies est élevée [93, 94]. Les résultats d'une étude [93] effectuée dans la région montagneuse de Tizi-Ouzou réalisée sur 150 vaches laitières au niveau de 22 exploitations montrent une prévalence de 47 % de mammites subcliniques, ce qui influence la qualité du lait de la région jugée au-dessous des normes des industries laitières algériennes.

Toutes les maladies de vaches ont la capacité d'affecter la population humaine directement ou indirectement. Cela peut se produire, particulièrement, par la réduction de la qualité ou la quantité des produits (viande, lait) consommés par les humains. Ces maladies peuvent également engendrer une réduction de l'activité des bovins et peuvent avoir un impact important sur la reproduction [95]. Certaines maladies peuvent entraîner un risque important sur la longévité

et influencer la santé de tout le troupeau et peuvent causer des risques sanitaires qui peuvent être une grande préoccupation pour le commerce et la santé humaine.

La détection des vaches laitières malades à un stade précoce de la maladie par l'observation du comportement individuel de la vache est une pratique très importante dans les fermes laitières [96]. Cependant, en raison du nombre croissant de vaches présentes actuellement dans les élevages cette observation visuelle directe du comportement est devenue impraticable, ce qui implique un besoin croissant d'une détection automatisée des changements dans les comportements (activité et alimentation) pour remplacer l'observation et assurer la détection automatique et précoce des vaches malades.

Les programmes de surveillance de la santé varient considérablement entre les différentes exploitations. Ces programmes peuvent inclure l'évaluation de l'attitude, de l'appétit, de la locomotion, de la température rectale ainsi que la collecte de fluides corporels pour les tests de laboratoire. La réalisation de ces évaluations et tests de diagnostic pour un grand nombre de vaches peut nécessiter beaucoup de temps et de main-d'œuvre. En outre, le comportement des vaches est perturbé car les vaches sont examinées lorsqu'elles sont retenues dans des barrières autobloquantes [97]. À cet égard, la surveillance automatisée du comportement (activité et alimentation) des vaches par des capteurs peut aider à réduire les limites et les inconvénients de ces programmes de surveillance. Les données générées par les capteurs pourraient être utilisées seules ou avec des protocoles traditionnels de surveillance de la santé pour identifier les vaches présentant des troubles de santé. En outre, une surveillance continue du comportement (activité et alimentation) et des paramètres physiologiques peut permettre de détecter des changements subtils avant que des signes cliniques évidents ne s'affichent. Une détection antérieure de la maladie peut favoriser la réponse au traitement et empêcher la progression de cette maladie.

L'état de santé de vaches a un effet important sur la rentabilité de la ferme. Il affecte directement les intervalles de vêlage ce qui impacte le rendement laitier. La plus grande longévité est directement proportionnelle au bénéfice économique associé à chaque vache, car elle permet d'augmenter la quantité de lait produite par une vache. De plus, avec une grande longévité, l'investissement dans le renouvellement du troupeau est plus petit [13].

Dans ce qui suit nous décrivons les maladies principales de vaches laitières.

## 2.2 Les maladies des bovins laitiers

La mammite et la boiterie [98] sont les maladies les plus fréquentes chez les bovins laitiers comme l'illustre le tableau 2.1. D'autres maladies touchent également les bovins avec une fréquence non négligeable comme les troubles de la reproduction et les maladies métaboliques

Maladie	Pourcentage de vaches Infectées (%)
Mammites	16,50 %
Boiterie	14,00 %
Infertilité	12,90 %
rétention placentaire	7,80 %
Fièvre du lait	4,60 %
Abomasum déplacé	3,50%
Respiratoire	3,30 %

Tableau 2.1: Fréquence des maladies en élevage

telles que la maladie de l’ovaire kystique, l’abomasum déplacé, le placenta retenu, la cétose et la fièvre du lait.

### 2.2.1 Les mammites

La mammite est considérée comme l’une des maladies les plus fréquentes et les plus coûteuses dans l’industrie laitière [98]. Si elle n’est pas traitée, elle peut entraîner une dégradation du bien-être, de la santé, de la productivité, de la qualité du lait et entraîner l’abattage voire la mort des vaches atteintes.

La mammite est une inflammation de la mamelle d’origine infectieuse ou toxique. Les pathogènes qui sont responsables de la mammite sont principalement les bactéries staphylococcus aureus, streptococcus uberis et escherichia coli. La mammite peut se classer selon son aspect clinique en mammite subclinique, mammite aiguë, mammite aiguë et gangréneuse et mammite chronique ou récurrente.

Les principales conséquences des mammites sont :

- La mise à la réforme : La mammite représente l’un des trois principaux motifs de mise à la réforme des vaches.
- L’augmentation du risque d’avoir d’autres maladies : Les vaches atteintes de mammite et de baisse d’appétit sont plus susceptibles d’avoir une cétose et une hypocalcémie.
- Troubles de la reproduction : Les vaches souffrant de mammite quarante-cinq jours après l’insémination artificielle sont pratiquement trois fois plus susceptibles d’avoir un avortement.
- Perte de production laitière : La perte de lait suite à la présence d’une mammite clinique a été évaluée à 1200 litres de lait pour la période restante de la lactation.

Etat	Degré	Description	Critères d'évaluation
Non boiteuse	SB 1	Normale	Marche rapidement et avec confiance en faisant de grands pas avec une ligne du dos horizontale.
	SB 2	Boiterie légère	Marche plus lentement, à plus petits pas avec une ligne du dos voûtée.
boiteuse	SB 3	Boiterie modérée	Marche lentement faisant délibérément de petits pas avec une ligne du dos voûtée.
	SB 4	Boiterie sévère	Se déplace lentement, faisant plusieurs arrêts pour reposer le membre affecté. Répartit le poids inégalement. Se tient debout et marche avec la ligne du dos très voûtée.

Où SB est un Score de Boiterie ( SB 1 : vache normale, SB 2 : légèrement boiteuse, SB 3 : modérément boiteuse, SB 4 : boiteuse )

Tableau 2.2: Échelle de classification des vaches boiteuses

- Surcharge de travail : Quatre heures sont consacrées aux soins d'une vache présentant une mammite clinique sévère.

### 2.2.2 La boiterie

La boiterie est une affection pénible qui amène la vache à marcher d'une manière différente afin de diminuer le poids de ses membres blessés. Elle est un problème du bien-être en raison de sa forte prévalence dans les troupeaux à travers le monde. Les vaches boiteuses peuvent avoir de l'inconfort et une douleur de longue durée [99]. La boiterie en elle-même n'est pas une maladie mais un symptôme causée par diverses maladies (Pathologies infectieuses du pied). La boiterie conduit à une diminution de la mobilité, à une réduction de la quantité d'aliments consommés, à une baisse de la production et à des problèmes de reproduction [100]. La prévention de la boiterie améliore la productivité, abaisse les coûts de traitement, atténue le stress des animaux et des éleveurs et préserve la santé des vaches. Donc détecter et traiter les vaches boiteuses est primordial pour les éleveurs.

Des capteurs qui mesurent les comportements associés à la boiterie chez les vaches peuvent aider en alertant l'éleveur des vaches qui ont besoin de traitement [101, 102].

Selon la sévérité de la boiterie, les vaches sont classées en 4 états comme le montre le tableau 2.2.

### 2.2.3 L'infertilité

L'infertilité se manifeste par le besoin de faire plus de deux inséminations pour provoquer ou non une gestation. Au niveau du troupeau, l'infertilité peut avoir lieu par des causes infectieuses et toxiques telles que : Campylobacteriosis, Rhinotrachéite bovine infectieuse (IBR),

virus de la diarrhée virale bovine (BVD), Neospora, Anaplasmose et toxicité au nitrate ou par des causes non infectieuses telles que : la nutrition, le stress et la génétique. Les maladies reproductives ont tendance à se développer progressivement et peuvent être difficiles à identifier jusqu'à ce qu'elles soient bien établies dans le troupeau. Les symptômes des maladies reproductives incluent des taux de fécondité médiocres, des avortements, la mortalité des veaux, etc.

#### 2.2.4 Les troubles métaboliques et digestifs

Les troubles métaboliques et digestifs commençant deux semaines avant le vêlage sont un ensemble de maladies qui altère la santé des vaches pendant une à neuf semaines après le vêlage. Elles peuvent influencer la performance de reproduction des semaines à quelques mois plus tard [103].

Les troubles métaboliques et digestifs tels que la cétose, l'abomasum déplacé et l'indigestion sont préjudiciables au bien-être de la vache et à la rentabilité de la ferme. En effet, ils entraînent des pertes dans la production laitière, augmentent le risque d'abattage et de décès, augmentent les coûts de traitement et nuisent à la performance de la reproduction.

La période de transition s'étend de trois semaines avant à trois semaines après le vêlage. Elle est une période difficile pour la vache laitière. Au début de l'allaitement, l'apport énergétique est habituellement inférieur aux besoins énergétiques de la production. Il en résulte un bilan énergétique négatif. Ce bilan énergétique négatif est associé à un risque accru de troubles de la santé tels que la fièvre du lait, la cétose et l'abomasum déplacé qui sont souvent liés entre eux.

Ces maladies ont de multiples impacts sur l'économie des élevages laitiers, sur les paramètres physiologiques des vaches laitières aussi que sur le comportement des vaches.

### 2.3 Les systèmes de capteurs utilisés dans le suivi sanitaire

Les auteurs de [11] présentent une synthèse sur les systèmes de capteurs développés dans le cadre de suivi sanitaire en élevage de précision. Ces systèmes de capteurs développés pour le suivi sanitaire de vaches peuvent être décrits par quatre différents niveaux comme le montre la Fig 2.1 [11].

1. Niveau technique : dans ce premier niveau, des systèmes de capteurs sont mis au point pour détecter automatiquement les comportements (activité, alimentation) ou mesurer les paramètres physiologiques (température corporelle, poids, etc.). A ce stade, les comportements détectés peuvent être utilisés seuls de sorte à permettre aux éleveurs de dé-

cider eux-mêmes comment les données comportementales doivent être interprétées. Par exemple, un éleveur pourrait juger la diminution du temps d'alimentation d'une vache et en conclure qu'il doit examiner la vache, la réformer, etc. Comme ils peuvent être intégrés dans des systèmes plus sophistiqués qui assurent des interprétations automatiques et offrent des conseils pour aider l'éleveur dans sa prise de décision.

2. Les interprétations : elles résument les changements dans les données des capteurs pour produire des renseignements sur l'état de la vache. Les interprétations permettent de faire des corrélations entre les données des capteurs et de l'état sanitaire des vaches.
3. L'intégration de l'information : cette étape est réalisée lorsque l'information des capteurs est complétée par d'autres renseignements (les chaleurs, l'économie, etc.). Cette intégration permet de produire des conseils (le traitement, la mise en réforme, etc.).
4. La prise de décision : dans ce niveau, en se basant sur les conseils offerts par le système de suivi sanitaire, l'éleveur prend une décision finale. De même, le système peut prendre la décision de façon autonome.

Les systèmes développés dans le cadre de cette thèse appartiennent au Niveau technique. En effet, nous avons conçu deux systèmes permettant de détecter automatiquement les activités et le comportement alimentaire de vaches laitières. Cependant, les vaches peuvent présenter jusqu'à 40 comportements différents. Les maladies influencent d'une manière significative qu'un sous-ensemble de ces comportements. Pour cela, nous allons présenter dans la suite de ce chapitre les résultats des études menées dans le niveau 2 (les interprétations) afin de trouver les corrélations les plus importantes entre les maladies et les comportements. Ce qui nous sert à définir les comportements qui doivent être détectés automatiquement par nos systèmes

L'élaboration de systèmes assurant le suivi sanitaire de vaches nécessite l'acquisition de données comportementales et physiologiques. Les données physiologiques permettent une détection précise des maladies [11]. Le suivi comportemental est utilisé pour assurer la détection précoce des maladies avant l'apparition des symptômes physiologiques. En général, les outils physiologiques sont utilisés comme méthode de référence lors de la validation des performances des outils comportementaux. Dans ce qui suit nous allons donner brièvement les mesures utilisées dans le cadre du suivi physiologique des maladies. Ensuite, nous allons décrire les relations existantes entre les comportements et les maladies.

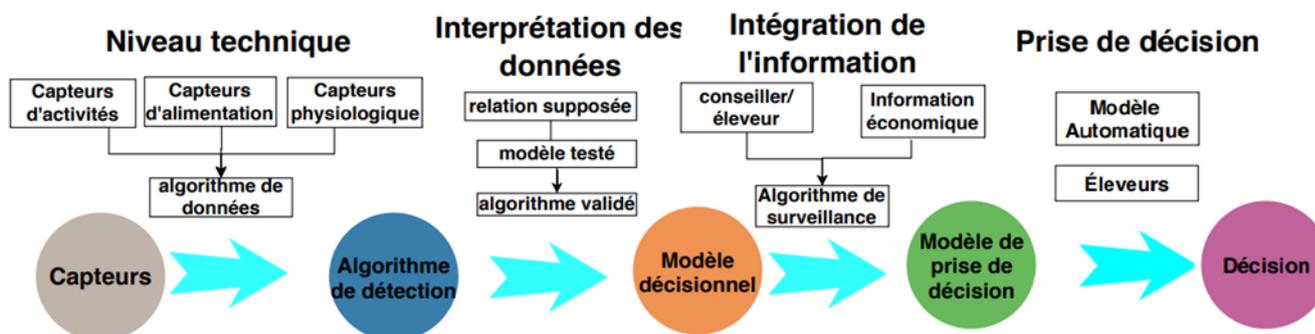


Fig 2.1: Systèmes de capteurs assurant le suivi sanitaire des vaches

## 2.4 Les mesures physiologiques utilisées dans le suivi sanitaire

Les changements physiologiques sont des signes qui montrent que la maladie commence à se répandre. Pour la mammite, par exemple, ces signes sont : la température corporelle élevée et le gonflement de la mamelle qui commence à être visible. Ci-dessous nous donnons les différentes mesures physiologiques qui peuvent être utilisées pour détecter chacune des maladies [11, 102].

- Mammite : la mesure de conductivité électrique (EC) du lait, la détection de la couleur du lait, la numération des cellules somatiques du lait (SCC capteur) et le bolus reticulum
- Boiterie : la répartition du poids entre les jambes, la plaque de distribution de force et les biomarqueurs sanguins.
- Infertilité : le taux de progestérone, la température corporelle et la vocalisation.
- Troubles métaboliques et digestifs : Le pH du fluide de rumen, la température du rumen, le pourcentage de graisse du lait, le niveau de cétone dans le lait et la température corporelle.

Les mesures physiologiques sont souvent utilisées comme méthode de validation des algorithmes développés dans le cadre du suivi comportemental des maladies. Dans ce qui suit nous décrivons les relations existantes entre les maladies et les changements comportementaux.

## 2.5 Le suivi comportemental pour la détection de maladies

Les maladies entraînent des changements dans la motivation comportementale et conduisent à une diminution de l'activité et à un manque d'appétit. Ce changement comportemental est une réponse de l'organisme de l'animal pour améliorer sa résistance à la maladie et faciliter le rétablissement. Généralement, les agents pathogènes envahissants initient des processus exigeant de l'énergie du corps. Ces processus ont pour objectif d'augmenter l'efficacité du système immunitaire pour surmonter l'infection. Par conséquent, l'énergie passera des activités comportementales telles que l'alimentation, le contact social ou le toilettage vers les besoins immunitaires. Les animaux malades pourraient avoir un ensemble différent de priorités comportementales [104]. Le faible appétit a été aussi interprété comme un moyen de réduire l'ingestion de certains micronutriments nécessaires à la multiplication des agents pathogènes [105]. Ces changements dans le comportement alimentaire et de l'activité sont souvent utilisés pour prédire les maladies chez les vaches laitières.

### 2.5.1 Modification comportementale associée à la boiterie

La boiterie dans les troupeaux laitiers est traditionnellement détectée par inspection visuelle, ce qui est long et subjectif. Comparées à des vaches en bonne santé, les vaches boiteuses passent souvent plus de temps à se coucher, marchent moins et changent de comportement alimentaire. Dans la littérature, plusieurs paramètres comportementaux ou de production, d'analyse de démarche ou de la posture ont été utilisés afin d'assurer la détection automatique de la boiterie. Dans [106], une synthèse a été faite pour regrouper une liste globale des variables utilisées dans 244 publications scientifiques qui abordent la détection de la boiterie par un suivi comportemental.

#### 2.5.1.1 Détection de la boiterie par le monitoring de l'activité

L'utilisation de capteurs mesurant l'activité est en augmentation dans les fermes modernes. Ainsi, certaines études ont utilisé des accéléromètres pour détecter automatiquement la boiterie [32, 107, 108]. Dans [107], une étude a été menée pour étudier les effets du degré de boiterie de plusieurs vaches sur les variables de comportement mesurées par accéléromètre. Sur quatre fermes commerciales, les données proviennent d'accéléromètres installés sur les pattes postérieures de 348 vaches holsteins. Le score de boiterie a été évalué sur la base de huit variables : la durée d'allongement par jour, la durée de la position debout, la durée de marche, le nombre total de pas, la fréquence de pas, l'accélération totale pendant le couchage (transition debout/couchée), l'accélération totale pendant la transition couchée/debout et l'accélération

SB 1	SB 2	SB 3	SB 4	Remarque
Durée de marche (min/jour)				
39,1	37,3	36,9	33,4	Diminue de 6 min entre SB4 et SB1
Accélération totale pendant la marche (g/jour)				
3327	3040	2827	2654	Diminue entre SB1, 2, 3, 4 vache malade Marche plus lentement que les saines
Accélération totale pendant la transition couchée debout				
1687	1482	1217	1446	Diminue entre SB2 et SB3
Durée de la position couchée (min/jour)				
684	689	673	724	Augmente chez SB4 de 40 minutes

Tableau 2.3: Valeurs des quatre composantes principales correspondantes à chaque score de boiterie.

totale pendant la marche. L'analyse a été faite en utilisant la méthode d'analyse de composantes principales (PCA). Les résultats les plus importants sont montrés dans le tableau 2.3. Cette publication a montré clairement que l'activité d'une vache semble être un bon indicateur de boiterie.

Dans [108], l'objectif était d'évaluer l'efficacité de la mesure du comportement de couchage par podomètre pour la détection de la boiterie. Un total de 30 vaches laitières holsteins en lactation ont été sélectionnées en fonction de leurs scores de locomotion. Les variables suivies sont : la durée (minimale, maximale, moyenne) d'une période de la position couchée et le nombre de périodes de la position couchée. Cette étude a prédit la boiterie chez les vaches avec une précision de 76 % en se basant sur les déviations du comportement.

Dans [32], l'objectif était d'évaluer l'association entre le comportement de couchage (transition debout/couchée) et la boiterie puis de déterminer si les mesures du comportement de couchage pouvaient être utilisées pour diagnostiquer la boiterie. Un accéléromètre a été utilisé pour l'acquisition de données. Les variables calculées à partir de cet accéléromètre sont : (1) la durée de la position couchée quotidienne (min/jour), (2) la fréquence des périodes de couchage par jour (n/jour), (3) la durée moyenne de la période couchée (min/période) et (4) l'écart-type de la durée de la période couchée (min/période). Les résultats sont montrés dans le tableau 2.4. Les seuils ont fourni une grande spécificité pour distinguer correctement les vaches qui n'étaient pas boiteuses. Cette étude montre que le comportement de couchage peut être utilisé pour la détection des vaches boiteuses.

### 2.5.1.2 Détection de la boiterie par le monitoring alimentaire

Les changements de comportement alimentaire peuvent être utiles pour identifier les problèmes de santé chez les bovins. Les vaches boiteuses ont une mobilité restreinte ce qui in-

Degré	Score de boiterie $\geq 3$		Score de boiterie = 4	
	Non boiteuse	Boiteuse	Non boiteuse	Boiteuse
Temps couché (h/jours)	11,2	11,6	11,2	12,8
Fréquences de période couchée (n/période)	8,1	8	8,1	8
Durée de période couchée (min/période)	80,2	83,1	80,3	95,3
Écart-type de durée de période couché (min/période)	44,0	46,9	44,4	50,7

Tableau 2.4: Valeurs des quatre composantes principales correspondantes à chaque score de boiterie.

fluence leurs habitudes quotidiennes et réduit leurs déplacements. Plusieurs études se sont intéressées au suivi du comportement alimentaire pour détecter les vaches boiteuses [95, 109, 110]. Dans [95], une étude a été faite pour identifier des changements dans les comportements alimentaires à court terme associés à l'apparition des troubles de santé (cétose, boiterie). Cette étude a montré que les troubles de la locomotion n'ont engendré qu'une petite diminution quotidienne de l'apport alimentaire avant le diagnostic. Cette diminution était probablement causée par la diminution rapide de la consommation d'aliments au cours des deux derniers jours avant le diagnostic. L'effet des troubles de la locomotion était proportionnellement beaucoup plus élevé sur le temps d'alimentation et le taux d'alimentation comme le montre le tableau 2.5. Les animaux ont probablement répondu à la douleur en se nourrissant progressivement à un rythme plus rapide et ont continué à consommer des quantités similaires de nourriture pendant des périodes plus courtes. En effet, le taux d'alimentation a montré une augmentation très constante, jusqu'à 2 ou 3 fois, avec l'apparition de problèmes de locomotion. La diminution, également, considérable du temps d'alimentation quotidienne a montré que se pourrait être un bon indicateur de ce type de maladie. Il semble probable que le temps d'alimentation soit une meilleure caractéristique de l'identification précoce de la boiterie que le rendement du lait car l'apport alimentaire et le rendement en lait n'étaient que légèrement réduits.

Dans [109], le but était de quantifier la diminution de l'apport alimentaire et les pertes de production de lait liées aux troubles de santé chez les vaches laitières (mammite, boiterie, etc.). Des enregistrements individuels quotidiens de l'apport alimentaire, de la production laitière et des troubles de santé étaient disponibles dans une ferme expérimentale (de 551 vaches holsteins). Les résultats obtenus par cette étude sont montrés dans le tableau 2.6. D'après

Variabes	Modification
Alimentation (kg de matière fraîche/J)	-1,57
Temps d'alimentation (min/J)	-19,10
Taux d'alimentation (g de matière fraîche/min)	21,61
Repas (n/j)	-0,226
Visites (n/j)	-10,50

Tableau 2.5: Effet de la boiterie sur les variables suivies.

Maladies	Jours d'arrivée	Jours de récupération	Effet initial (kg)	Effet total (kg)
Lésions au jarret	-5	>126	2,7	48,1
Lésions au pied	-4	56	6,4	27,8

Où :

Jours d'arrivée exprime le nombre de jours où la diminution de l'apport alimentaire a commencé par rapport au jour du diagnostic de la maladie (jour 0).

Jours de récupération exprime le nombre de jours entre le début de la maladie et le rétablissement.

Effet initial exprime les pertes cumulées de l'apport alimentaire depuis le jour de l'arrivée de la maladie jusqu'au jour du diagnostic.

Effet total exprime les pertes cumulées de l'apport alimentaire depuis le jour de l'arrivée de la maladie jusqu'au jour de la guérison.

Tableau 2.6: L'effet de lésions au pied et au jarret sur l'apport alimentaire.

cette étude, l'apport alimentaire et la production laitière ne semblent pas un bon indicateur pour la détection précoce des problèmes de boiterie, mais semble être un bon indicateur pour suivre la récupération et l'état de la vache après le diagnostic.

Dans [110], l'objectif était de quantifier l'effet de la boiterie sur l'alimentation et la rumination. Les vaches boiteuses différaient significativement des vaches non boiteuses dans trois aspects de leur comportement alimentaire : le temps, la fréquence et le taux d'alimentation. La boiterie réduisait le temps d'alimentation quotidienne de 46 % et la fréquence d'alimentation de 44 %. L'impact de boiterie sur le temps d'alimentation s'exprime par le fait que les vaches boiteuses mangent plus rapidement.

### 2.5.2 Modification comportementale associée aux mammites

Les vaches laitières atteintes de mammite montrent des changements dans leurs comportements. Les changements dans l'activité et le comportement alimentaire ont été rapportés par plusieurs études [104, 105, 111].

Le but dans [111] était d'évaluer l'impact des mammites sur le comportement alimentaire et l'activité durant la traite chez 60 vaches laitières. Les résultats sont présentés dans le tableau

2.7. D'après cette étude, dans le cas de mammites expérimentales ou de mammites cliniques aiguës, la diminution du comportement de couchage peut être considérée comme une des réponses à une infection de la mamelle. De plus, les résultats montrent que le comportement alimentaire pourrait être un indicateur précoce de la mammite car il diminue de 7 à 10 % pendant plus de 4 jours avant le diagnostic et le traitement de la mammite. La diminution de l'apport alimentaire persiste après le traitement par antibiotique, il peut être un indicateur de l'état de récupération après la période de traitement.

Dans [104], l'objectif était d'examiner le comportement des vaches laitières dans la période allant de 2 jours avant (J-2 et J-1) à 3 jours (J 0, 1 et 2) après une provocation expérimentale de mammite avec *Escherichia coli*. Les effets de la mammite induite par l'expérience sur le comportement ont été examinés chez 20 vaches holsteins Friesian. Les résultats sont présentés dans le tableau 2.8. Dans les premières 24 h après l'infection, les vaches laitières ont affiché une diminution de l'activité de marche et ont passé moins de temps à ruminer et à s'alimenter. Le temps d'alimentation a diminué seulement quelques heures après l'injection de *Escherichia coli* et il est resté abaissé au cours des 24 premières heures. La quantité de matière sèche ingérée (DMI) a diminué au cours de cette période. Les vaches malades ont eu une fréquence plus faible de comportement d'auto-toilettage et de rotation de tête vers mamelle par rapport aux jours précédents l'infection et après la récupération. De plus, les réponses comportementales étaient en corrélation avec les symptômes cliniques de la mammite tels que le rendement du lait et la température rectale.

Dans [105], l'objectif était d'évaluer le comportement et les signes cliniques des vaches avant et après la mammite clinique expérimentale induite par l'endotoxine d'*Escherichia coli*. Les vaches utilisées dans les expériences étaient des races finlandaises Ayrshire ou Holstein Friesian. Le comportement des vaches a été enregistré en continu pendant 48 h, en commençant 24 heures avant l'induction de la mammite. Les résultats sont présentés dans le tableau 2.9. D'après cette étude, la réduction de la consommation d'aliments semble être un bon indicateur de tous les troubles métaboliques. De plus, les vaches ont passé moins de temps couchées le jour de l'induction que les journées précédentes. Les vaches ont également passé plus de temps à manger de l'ensilage pendant la journée d'induction et elles ont aussi fait plus de pas que les journées précédentes. Les vaches ont passé moins de temps couchées et ruminées et ont moins bu lorsqu'elles étaient malades. Les vaches ont augmenté le temps debout et évitaient de se coucher sur le côté de la mamelle enflammée. Les résultats de cette étude ont confirmé le potentiel du monitoring comportemental pour la détection des vaches souffrant de mammite.

Variabes	Vaches en mammites	Vaches saines
Apport alimentaire (kg/j)	41	43,5
Temps de position couchée (min/j)	732	756
Taux d'alimentation (Kg/min)	0,22	0,26
Fréquence de levage de patte (levé/traite)	2,8	2,6
Nombre de coups de pied (coup/traite)	1,7	0
Période d'allongement (Période/j)	2,4	2,2
Nombre de pas (pas/jour)	1380	1284

Tableau 2.7: Effet des mammites sur les différentes variables comportementales.

Variable	Modification	Durée
Temps d'alimentation	Diminue	J0 < à {j-2, j-1, j1, j2}
Temps de rumination	Diminue	J0 < à {j-2, j-1, j2}
Position couchée	Augmente	J2 > à {j-2, j-1, j0, j1}
Fréquence de rotation (tête vers mamelle)	Diminue	J0, J1 < à {j-2, j-1, j2}
Fréquence autotoiletage	Diminue	J0 < à {j-2, j-1, j1, j2}
Ralentissement	Augmente	J0 > à {j-2, j-1, j1, j2}
Quantité matière sèche ingérée (DMI)	Diminue	12 h à 36h après infection

Tableau 2.8: Impact des mammites sur : l'alimentation, la rumination, l'activité, le toilettage et la rotation de tête.

Comportement	Durée et fréquence journalière totale	
	contrôle	induction
Position debout (Durée)	638	759
Position couchée (Durée)	800	670
Coucher sur le quartier de contrôle (Durée)	381	326
Coucher sur le quartier induit (Durée)	418	353
Mange fourrage (Durée)	221	252
Boire (Durée)	13	14
Rumination (Durée)	415	394
Fréquence de marche	1160	1413

Tableau 2.9: Impact des mammites sur : l'activité, l'abreuvement et le toilettage.

### 2.5.3 Détection des troubles métaboliques et digestifs

L'apparition de troubles métaboliques entraîne plusieurs variations d'activité et de comportement alimentaire [97, 112, 113]. L'objectif de [97] était d'évaluer la performance d'un système automatisé de surveillance de la santé pour identifier les vaches atteintes de troubles métaboliques et digestifs, en utilisant un système d'alerte qui combine la rumination et l'activité physique. Les troubles de santé surveillés comprenaient : (1) rétention placentaire, (2) la fièvre du lait, (3) la cétose, (4) l'abomasum déplacé (DA), (5) indigestion, (6) métrite : aqueuse, rose/marron, (7) mastite clinique, (8) pneumonie et (9) boiterie. Le tableau 2.10 montre les différences obtenues entre 5 jours avant le diagnostic clinique et le jour du diagnostic clinique pour le temps quotidien de rumination, l'activité quotidienne et la production de lait. D'après cette étude, la sensibilité du suivi comportemental pour détecter les vaches atteintes de tous les troubles métaboliques était supérieure à 90 %. Donc les vaches malades peuvent être identifiées à un stade précoce.

L'étude [112] vise à déterminer si les changements dans les variables suivantes : (1) le rendement laitier, (2) l'apport alimentaire, (3) visites à la zone d'alimentation, (4) la rumination, (5) l'activité et (6) le poids corporel, peuvent être utilisées pour la détection précoce de trois troubles de métabolisme (la fièvre du lait, l'abomasum déplacé et la cétose). D'après cette étude, les prédictions les plus performantes sont liées à l'activité, la rumination, la réduction du poids au début de lactation et les alertes liées à l'alimentation (visite à la zone d'alimentation).

L'objectif dans [113] était de comparer les comportements des vaches laitières avec et sans cétose pendant les jours autour du vêlage pour déterminer si des changements dans le comportement alimentaire et de l'activité pourraient être utiles dans l'identification précoce des vaches malades. Les comportements alimentaire et les activités ont été mesurés à partir de 7 jours avant le vêlage jusqu'à 21 jours après le vêlage. D'après cette étude, le temps de la position debout était plus long pour les vaches en cétose que celles saines pendant la semaine avant le vêlage et le jour 0. Mais il n'était pas différent pendant les 21 jours après le vêlage. La durée moyenne des périodes de la position debout était plus longue pour les vaches en cétose dans le jour de vêlage comparativement aux vaches saines.

## 2.6 Conclusion

Le suivi automatique des marques sanitaires est réalisé grâce au développement d'outils qui doivent être exploités par les éleveurs pour prendre des décisions en réponse aux différentes maladies. En effet, lorsqu'une vache est malade, plusieurs altérations comportementales se manifestent. Ces modifications provoquent souvent une baisse d'activité de sorte que la vache

Variables	Vache saine	Vache atteinte de troubles
Rumination (min/j)	-5,5	-152,6
Activité (unité arbitraire/j)	+10,2	-72,5
Production de lait (kg/j)	+3,3	-7,9

Tableau 2.10: Les différences (selon les unités de mesure) entre : 5 jours avant le diagnostic clinique (CD), le jour du diagnostic clinique des différentes variables.

passé plus de temps en période de repos. Le comportement alimentaire est aussi impacté par une diminution remarquable dans le taux et la fréquence d'alimentation.

Dans ce chapitre, nous avons présenté les principales maladies qui infectent les vaches laitières en se focalisant sur l'impact comportemental de ces maladies. Nous avons montré que le monitoring comportemental offre un potentiel considérable pour améliorer les processus de surveillance automatique du bien-être et de la santé des vaches. Les systèmes de monitoring comportemental permettent de détecter tout changement avant l'apparition des signes cliniques permettant ainsi à l'éleveur de prendre les mesures nécessaires le plus rapidement possible.

La synthèse présentée dans ce chapitre a essentiellement montré les comportements majeurs qui sont le plus communément associés aux événements sanitaires. D'après cette synthèse, les activités majeures sont : le temps de la position debout, le temps de la position couchée et la marche. Les paramètres du comportement alimentaire les plus souvent suivis sont : le temps d'alimentation et le taux d'alimentation. Donc, l'ensemble de ces comportements alimentaires et activités vont être détectés automatiquement par les systèmes développés dans le cadre de notre thèse.

Dans le chapitre suivant, nous présenterons une synthèse qui porte sur le développement de systèmes (Niveau technique) permettant la détection automatique en temps réel de données comportementales.



# Chapitre 3

## Exploration de données pour la classification des comportements de vaches

### 3.1 Introduction

Le déploiement des Nouvelles Technologies de l'Information et de la Communication (NTIC) et leur intégration aux sciences biologiques, chimiques, physiques, écologiques et économiques permettent de renforcer les tâches de gestion des fermes laitières et le processus de prise de décision.

Le comportement des vaches est complexe et imprévisible, il doit être bien compris. Les technologies modernes qui surveillent en permanence divers aspects (activité, alimentation, morphologie, etc.) produisent de larges quantités de données afin de contribuer à la bonne compréhension de tous les comportements. L'analyse de ces données hétérogènes permet l'extraction de l'information utile servant à l'amélioration de l'état de santé et de la productivité.

La collecte des variables comportementales nécessite l'utilisation de réseaux de capteurs sans fil. Cette collecte produit une quantité massive de données. Le volume, la variété et la rapidité d'acquisition de ces données sont à l'origine de la notion du « Big data » [114]. Le Big data apporte des solutions qui permettent de stocker et de traiter différents types de données générées. La fouille de données (Data Mining) permet d'approfondir l'analyse sur ces données pour extraire de l'information ou effectuer des prédictions. Le Data Mining utilise des techniques d'apprentissage automatique pour prédire des résultats. L'apprentissage automatique offre un grand potentiel d'analyse de données volumineuses. Il est un sous-domaine de

l'intelligence artificielle qui fournit une stratégie modulaire et évolutive. L'apprentissage automatique donne aux machines la capacité d'apprendre sans être strictement programmées, permettant de créer de nouvelles perspectives d'extraction de l'information et de sa compréhension.

Ce chapitre est divisé en deux parties. La première partie sert à donner les différentes étapes utilisées dans le processus de fouille de données. La deuxième partie permet de présenter les approches de fouille de données utilisées pour assurer le suivi comportemental (activité et alimentation) de vaches laitières.

## 3.2 Partie 1 : la fouille de données

La fouille de données est un processus automatique ou semi-automatique de découverte d'informations dans une large échelle de données [115]. Elle utilise un ensemble de méthodes de régression, de classification, de regroupement et de prétraitement dans le but de rechercher, d'identifier, de valider l'information et de faire des prévisions [115]. En effet, l'informatisation des secteurs industriels entraîne une explosion de la taille des données [116]. Les méthodes de fouille de données sont utilisées dans le but de parcourir les bases de données pour extraire des informations jusque-là inconnues et pour générer des prévisions [117]. En fonction de certaines mesures, l'ensemble de données est analysé pour engendrer une information précieuse. Par exemple, en élevage la sélection animale est définie comme une opération décisionnelle permettant de regrouper les animaux en deux partitions : les animaux à maintenir en ferme pour le renouvellement et ceux à vendre [118]. Plusieurs critères interviennent dans cette sélection et ils sont liés à l'historique des vaches. Des informations comme la parité et l'état de santé sont utilisées par des éleveurs expérimentés pour établir cette décision. La bonne sélection a des avantages et assure le succès de renouvellement du troupeau. La fouille de données permet de ressortir les paramètres les plus importants utilisés par les experts et donc propager les compétences entre les éleveurs. La prise en considération de ces paramètres est importante car elle permet d'automatiser le processus de prise de décision.

La fouille de données ou «la découverte de connaissances à partir de données » est composée d'une séquence de deux étapes : le prétraitement de données et l'exploration de l'information comme le montre la Fig 3.1. Dans l'étape du prétraitement, les données sont préparées pour l'extraction de l'information et dans l'étape d'exploration de l'information, les données sont utilisées pour produire une base de connaissances.

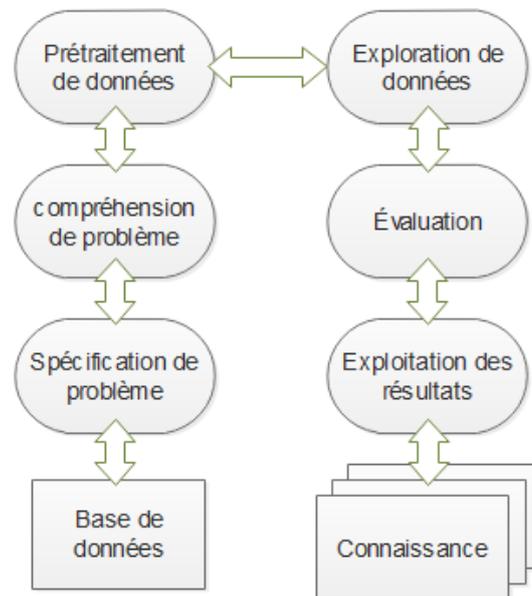


Fig 3.1: Processus de découverte de connaissances.

### 3.2.1 Le prétraitement de données

La préparation de données est une phase importante du processus de découverte de connaissances. Les données collectées sont de multiples sources, hétérogènes, bruitées, pas assez contrôlées, avec des combinaisons incohérentes et avec des manques [119]. Le prétraitement a pour but d'améliorer la qualité des données en utilisant des techniques de nettoyage, de réduction, d'intégration et de transformation de données. Le nettoyage de données est utilisé pour améliorer la qualité de données issues d'une seule source. Tandis que, l'intégration de données permet d'améliorer la qualité de données issues de plusieurs sources. La réduction et la transformation de données visent à sélectionner ou à ressortir les données pertinentes qui doivent être utilisées par le modèle d'exploration de données. Le prétraitement de données a également pour rôle d'adapter l'ensemble des données aux exigences des algorithmes d'exploration afin d'améliorer les performances des modèles de découverte de connaissances [114].

#### 3.2.1.1 Le nettoyage des données

Le nettoyage de données est un processus de filtrage de données inutiles. Son rôle est de compléter les valeurs manquantes et de filtrer les données bruitées. Il permet d'identifier et de supprimer les valeurs aberrantes, de résoudre les incohérences et de détecter les anomalies [120]. Cette étape nécessite généralement une intervention humaine.

Les valeurs manquantes peuvent être nettoyées ou corrigées d'une façon manuelle ou automatique [121]. Dans l'approche manuelle, lors de l'étiquetage de données, les valeurs manquantes ou aberrantes peuvent être supprimées ou être complétées en les classant comme étant « Inconnues ». Dans l'approche automatique, les données manquantes peuvent être complétées en utilisant la mesure de la tendance centrale de l'attribut ou bien en utilisant la valeur la plus probable pour remplacer la valeur manquante. Cette valeur peut être déterminée à l'aide d'une régression par exemple.

La qualité des modèles de découverte de connaissances est liée aux données d'entraînement utilisées lors de l'apprentissage. Lorsque la collection est imparfaite, les données sont inévitablement bruitées. Le bruit est une erreur aléatoire ou une variance dans une variable mesurée [122]. Une solution précise visant à filtrer le bruit est difficile à obtenir. Néanmoins, les méthodes de visualisation de données peuvent être utilisées pour identifier les valeurs aberrantes qui peuvent représenter du bruit. A cause de cette difficulté à nettoyer les données bruitées, la phase de collecte de données permettant de concevoir les bases de données utilisées lors de l'élaboration de notre système de monitoring des activités (chapitre 4) est réalisée de façon à éviter toute sorte de bruit. En effet, le capteur d'acquisition de données a été programmé pour signaler le dysfonctionnement du matériel ce qui permet d'éviter les bogues matériels qui sont la principale cause du bruit.

### 3.2.1.2 L'intégration de données

L'intégration de données est le processus qui consiste à combiner des données provenant de différentes sources. Elle utilise un ensemble d'approches permettant de prétraiter des données issues de sources variées et contenant des attributs corrélés [123]. L'objectif est de supprimer les données redondantes et incohérentes. Ces approches fournissent une interface structurée et unifiée entre des données hétérogènes et les modèles d'exploration. La Fig 3.2 montre un exemple de module d'intégration de données de différentes sources utilisées dans la détection d'incendies [124]. Le module d'intégration fournit deux interfaces. Une interface inférieure pour les données, elle permet de réaliser des modélisations, de la fusion et du filtrage. Une interface supérieure pour le modèle d'exploration qui permet d'unifier les données.

### 3.2.1.3 La réduction de données

L'analyse de grand volume de données nécessite un temps élevé. Ces données sont souvent complexes et contiennent un nombre élevé d'attributs qui peuvent être corrélés. Les techniques de réduction de données sont appliquées pour réduire de manière dynamique cette complexité de sorte à supprimer ou à regrouper les instances redondantes [125]. Elles per-

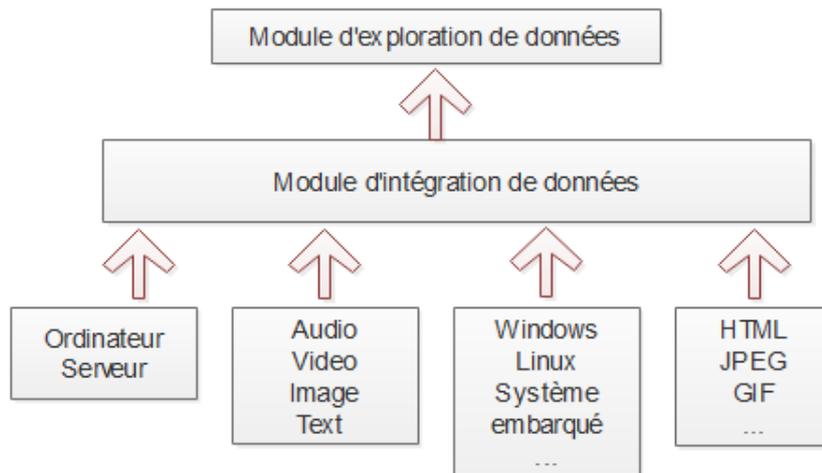


Fig 3.2: Architecture d'un système d'intégration de données

mettent d'obtenir une représentation réduite de l'échelle des données tout en maintenant la signification et la structure des données d'origine. Les données pertinentes seront extraites de la base de données et exploitées par des algorithmes d'apprentissage. Les données réduites permettent un entraînement rapide et précis, une meilleure compréhension des résultats et une mise à jour des modèles existants [126].

La sélection des caractéristiques vise à supprimer les attributs inutiles et répétitifs afin de ne garder que les données utiles [127]. Cette sélection peut se faire en parallèle avec l'algorithme d'apprentissage ou indépendamment à l'aide de filtre.

La réduction d'instances est utilisée pour sélectionner ou générer les instances les plus représentatives. En effet, les bases de données peuvent contenir des instances non informatives, bruyantes ou redondantes. Elles occupent un large volume et augmentent le temps d'exécution. Les algorithmes de sélection doivent décider quelle instance est importante. La génération de nouvelle instance permet de réduire le bruit [128]. Lors de la conception de nos systèmes de suivi comportemental (chapitre 4 et 5), des techniques de réduction de données seront utilisées pour sélectionner les données pertinentes qui doivent être utilisées par les algorithmes d'apprentissage.

#### 3.2.1.4 La transformation de données

La transformation de données peut améliorer la précision et l'efficacité des algorithmes d'apprentissage [129]. La construction d'attributs, l'agrégation, la normalisation et la discrétisation sont les techniques utilisées. Ces techniques peuvent être exécutées indépendamment les unes des autres. La normalisation permet d'attribuer une même échelle (plage de valeurs allant de

0,0 à 1,0) à tous les attributs. Dans la discrétisation, les données quantitatives sont transformées en données qualitatives. L'agrégation permet de résumer les données et cela en combinant deux ou plusieurs attributs en un seul attribut. La construction d'attributs aide le processus d'exploration en formant de nouveaux attributs. Pour cela, cette technique sera utilisée pour former de nouveaux attributs lors de la conception du système de détection d'activités (chapitre 4).

## 3.2.2 L'exploration des données

Après l'étape de préparation de données, des méthodes d'exploration de données sont appliquées pour produire la base de connaissances. Elles se basent sur des approches de classification, de regroupement et de régression. Les modèles élaborés à partir de ces approches doivent être évalués pour tester leurs aptitudes à représenter les connaissances. De plus, des approches d'optimisation peuvent être utilisées pour combiner les différentes approches et obtenir de meilleures performances lors de l'évaluation.

### 3.2.2.1 La classification

La classification de données vise à apprendre des associations entre des caractéristiques et une variable cible d'intérêt. Le concept est qu'à partir d'un ensemble de données préalablement libellées, la méthode de classification apprend à déterminer la classe d'appartenance de nouvelles données [130]. La classification de données est un processus composé de deux étapes (l'apprentissage et le test). Lors de l'apprentissage, l'entraînement du modèle de découverte de connaissances est réalisé à l'aide d'instances libellées. Puis les performances du modèle sont évaluées dans la phase de test. Les algorithmes de classification sont utilisés dans l'apprentissage automatique, la reconnaissance de formes, les statistiques, la prévision, etc. La classification est une méthode d'apprentissage supervisé car l'appartenance de l'ensemble de données est préalablement connue. Plusieurs méthodes de classification de données ont été développées telles que les réseaux de neurones, les méthodes probabilistes, les arbres de décision et machine à vecteurs de support (SVM) [131]. Les performances de classification dépendent du choix de la méthode de classification et du choix réaliser lors de l'acquisition et du prétraitement des données. Dans le cadre de la conception de nos systèmes de suivi des activités et du monitoring alimentaire plusieurs méthodes de classification seront utilisées. Ces méthodes sont introduites en section 3.2.3.

### 3.2.2.2 Le regroupement (clustering)

Le regroupement est un processus qui vise à regrouper un ensemble de données en plusieurs groupes (clusters). Il se base sur l'apprentissage non supervisé; il est appliqué lorsqu'il n'y a pas de classe à prédire. Le regroupement fonctionne de sorte que les objets vont être divisés en groupes selon leurs natures [132]. Comparé à la classification supervisée, le regroupement est une tâche plus difficile du fait que les groupes ne sont pas préalablement définis; la décision d'appartenance d'une nouvelle caractéristique à un groupe est difficile et le libellé de classe de chaque cluster est inconnu. Les objets du même cluster présentent une similarité élevée et doivent être différents des objets d'autres clusters [133].

Le regroupement en tant qu'outil d'exploration de données est utilisé dans de nombreux domaines d'application tels que la biologie et la science médicale [134]. Il permet une meilleure visualisation et interprétation des données. Par exemple, les vaches malades présentent des symptômes similaires et peuvent être groupées ensemble dans un groupe qui porte le nom de leur maladie. Cependant, les vaches saines ne présentent aucun de ces symptômes donc elles ne doivent pas être groupées dans le même groupe.

Des méthodes de densité, de partitionnement et hiérarchiques sont utilisées dans le regroupement. Les regroupements peuvent générer différents groupes sur le même ensemble de données. Ils peuvent être classés de manière générale en trois catégories : dur (hard clustering), flou (fuzzy clustering) et doux (soft clustering). La mise en cluster est utile car elle peut mener à la découverte de groupes précédemment inconnus dans les données [135]. Les algorithmes de base sont : les k-moyennes floues et les k-moyennes floues intuitionnistes. Ils ont été largement améliorés pour qu'ils puissent être utilisés dans le cas du regroupement de données volumineuses.

### 3.2.2.3 La régression

La régression constitue un modèle d'apprentissage supervisé qui vise à fournir la prédiction d'une variable de sortie en fonction des variables d'entrées connues. Les algorithmes les plus connus incluent la régression linéaire, la régression quadratique et la régression logistique [136]. La mesure du degré de relation linéaire entre deux variables est appelée régression linéaire simple [137]. L'analyse de régression est une méthodologie statistique qui est le plus souvent utilisée pour la prédiction numérique. La régression comprend également l'identification des tendances de la distribution sur la base de données disponibles. Les modèles de régression ont des fonctions à valeurs continues comparativement à la classification qui prédit des étiquettes catégorielles. Lors de la conception de système de suivi des activités des vaches (chapitre 4), nous allons utiliser la régression linéaire afin de sélectionner les variables

qui ont le plus d'aptitude à classer les différentes activités.

#### **3.2.2.4 L'évaluation du modèle d'apprentissage**

Un système d'exploration de données a le potentiel de générer des milliers, voire des millions de modèles ou règles. Un modèle est intéressant s'il valide une hypothèse que l'utilisateur cherche à confirmer. L'évaluation de modèles permet d'identifier les modèles vraiment intéressants représentant des connaissances [138]. L'évaluation permet aussi de comparer les performances de plusieurs classificateurs et de choisir celui qui convient mieux pour créer le modèle de découverte de connaissances [139]. Selon le domaine d'utilisation de l'apprentissage, différentes mesures de performances sont utilisées. L'exactitude, la sensibilité, la spécificité et la précision sont des performances couramment utilisées lors de l'évaluation des modèles générés. Les performances de validation sont calculées en utilisant quatre mesures d'évaluation: le vrai positif, le vrai négatif, le faux positif et le faux négatif. Les deux premiers désignent une classification correcte par contre les deux autres désignent une erreur de classification. Les données utilisées en phase de validation doivent être différentes de celles utilisées en phase d'entraînement. Donc l'ensemble de données doit être divisé en deux groupes, les données d'entraînement et celles de validation [140].

#### **3.2.2.5 L'optimisation des performances de classification**

La présence de données de chaque classe dans l'ensemble de toutes les données est souvent non équitable. Cela engendre un déséquilibre dans la répartition des volumes de données associées à chaque classe. Afin d'améliorer les performances des modèles de classification et éviter ce déséquilibre, des techniques sont utilisées. Ces techniques sont : le sur échantillonnage, le sous-échantillonnage, le déplacement du seuil, et les théories des d'ensembles [141, 142]. De plus, des approches hybrides qui se basent sur la combinaison de plusieurs classificateurs peuvent être utilisées pour améliorer les performances des modèles d'apprentissage [143, 144]. Une approche d'optimisation qui se base sur la combinaison de plusieurs classificateurs sera présentée dans les chapitres 4 et 5.

### **3.2.3 Les algorithmes d'apprentissage automatique**

Dans cette section, nous allons présenter cinq algorithmes d'apprentissage les plus couramment utilisés [145]. Ces algorithmes se basent sur : une structure hiérarchique comme les arbres de décision, des modèles probabilistes comme les modèles de mélanges, des algorithmes bioinspirés comme les réseaux de neurones, un modèle de régression comme la régression

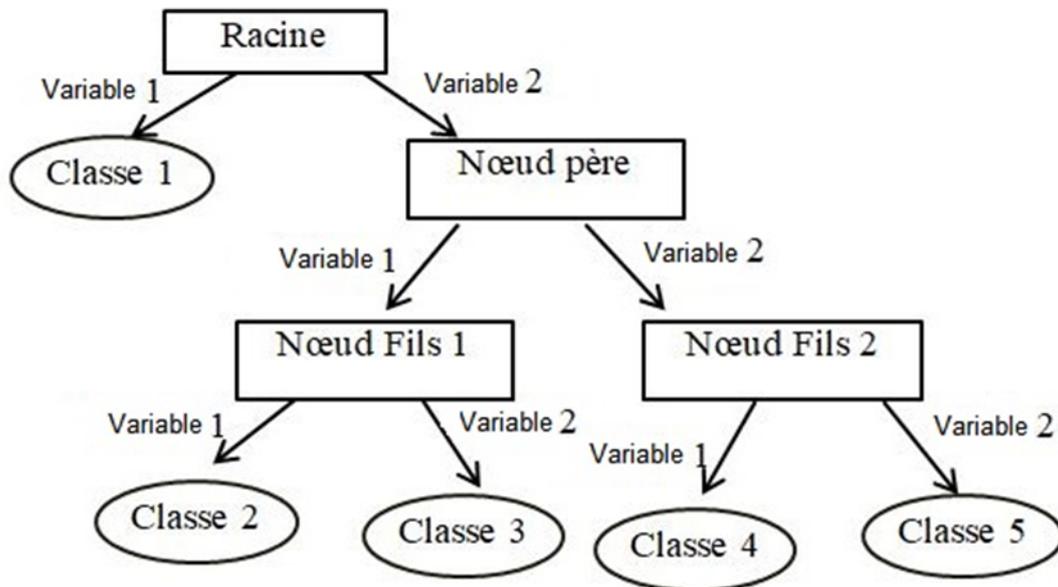


Fig 3.3: Architecture d'arbre de décision binaire

linéaire et des modèles statistiques comme la machine à vecteurs de support. Tous ces algorithmes seront utilisés dans notre méthode de classification des comportements des vaches dans les chapitres 4 et 5.

### 3.2.3.1 Les arbres de décision

L'arbre de décision est l'un des algorithmes les plus utilisés pour la classification de données volumineuses [146, 147]. Son avantage est qu'il est explicite et facile à suivre et permet d'avoir une performance de classification élevée pour un coût de calcul relativement faible. En effet, l'arbre de décision propose une structure hiérarchique arborescente où chaque nœud interne désigne une classification, chaque branche représente un résultat de la classification et chaque nœud feuille porte un label de classe. Le nœud le plus élevé de l'arbre est le nœud racine. La Fig 3.3 montre un exemple d'un arbre de décision binaire simple. Les nœuds internes sont désignés par des rectangles et les nœuds feuilles sont désignés par des ovales.

### 3.2.3.2 Les réseaux de neurones

Un réseau de neurones est une méthode de classification qui se base sur un algorithme biologiquement inspiré de la structure des neurones présents dans le cerveau humain [148]. Un réseau de neurones naturels est composé d'un ensemble d'unités de traitement appelées « neurones » interconnectées entre elles par un réseau de communication. Le fonctionnement

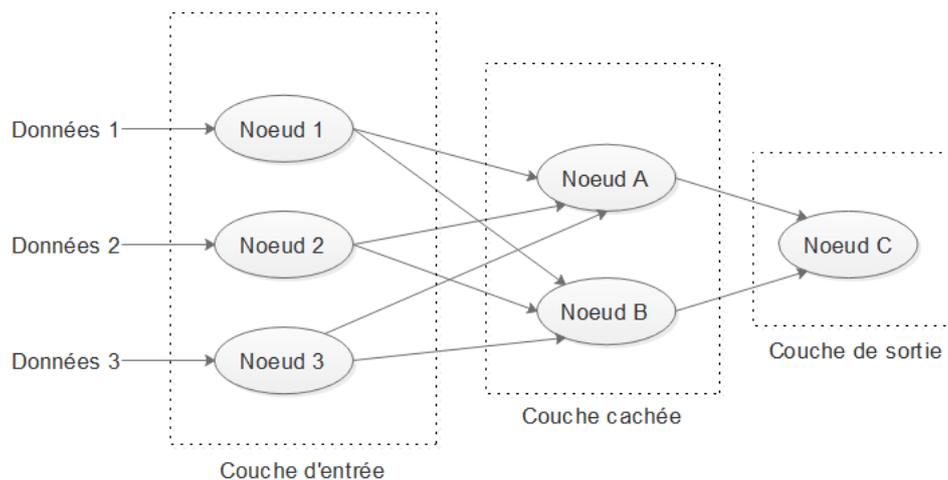


Fig 3.4: Architecture basique d'un réseau de neurones artificiels

collectif fournit une structure permettant de réaliser des calculs complexes. Les réseaux de neurones artificiels ont repris le même principe de fonctionnement que les réseaux de neurones réels. Ils sont utilisés lors de développement de modèles d'apprentissage automatique ou profond [149].

La topologie des unités de traitement dans un réseau de neurones artificiels est représentée comme un graphe dirigé. Les nœuds du graphe représentent les neurones et les arcs représentent les liens de communication. Les nœuds sont organisés en trois types de couches. Les neurones de la couche d'entrée permettent de recevoir les données, ceux des couches cachées permettent le déroulement de l'apprentissage et les neurones de la couche de sortie donnent la décision de classification. Cette architecture est illustrée par la Fig 3.4. L'augmentation du volume de données d'apprentissage a conduit à l'élaboration de nouvelles architectures de réseaux de neurones appelés « réseaux profonds » tels que : les réseaux de neurones à convolution (CNN) [150]. Ces types de réseaux assurent un entraînement étendu et accéléré.

### 3.2.3.3 Le modèle de mélange gaussien

Un modèle de mélange fini est une méthode probabiliste assurant une classification non supervisée. Le principe est qu'à partir d'un ensemble de données initial et un ensemble de groupe prédéfini, le modèle de mélange calcule la moyenne et la variance de chaque groupe, puis réalise une répartition de l'ensemble de données entre les groupes selon leurs distributions [151].

Un modèle de mélange gaussien suppose que toutes les populations de données sont issues à partir d'un mélange d'un ensemble de distributions normales appelées « gaussiennes » [152]. Les modèles de mélange peuvent être univariés ou multivariés. Les modèles de mélange

univariés étudient la distribution des données d'une seule variable. Cependant, les modèles multivariés étudient la distribution de données en combinant plusieurs variables.

### 3.2.3.4 La régression linéaire

La régression linéaire est une méthode d'analyse utilisée dans plusieurs domaines tels que : les sciences sociales, comportementales et la médecine [153]. La régression linéaire est utilisée pour trouver une relation linéaire entre la cible et un ou plusieurs prédicteurs. Il existe deux types de régression linéaire (simple et multiple) [154]. La régression linéaire simple est utile pour trouver une relation entre deux variables continues. L'une est un prédicteur (variable indépendante) et l'autre une réponse (variable dépendante). La régression linéaire multiple utilise plusieurs variables indépendantes pour prédire le résultat d'une variable de réponse. Elle a pour objectif de modéliser la relation linéaire entre les variables indépendantes X et la variable de réponse Y (dépendante) selon l'équation (3.1) ci-dessous.

$$Y = b + b_0X_0 + b_1X_1 + \dots + b_kX_k \quad (3.1)$$

Où :  $b, b_0, b_1, \dots, b_k$  sont les coefficients de la régression.

### 3.2.3.5 Les machines à vecteurs de support

Les machines à vecteurs de support, dites Support Vector Machines en anglais (SVM), sont des outils puissants de classification de données linéaires et non linéaires et d'analyse de régression. Ils sont utilisés dans l'apprentissage automatique, la reconnaissance de formes, l'optimisation [153], etc. Leur objectif est de séparer un ensemble de données en classes [154]. Les machines à vecteurs de support utilisent un hyperplan de séparation pour classer l'ensemble de données en classes. Le noyau des machines à vecteurs de support utilise des fonctions de différents types (linéaire, non linéaire, sigmoïde, etc) [155]. Son rôle est de transformer les données d'entrée vers une forme appropriée. La Fig 3.5 montre un exemple d'une fonction discriminante linéaire  $h(x) = w \cdot x + b$ , utilisant un hyperplan linéaire pour séparer des cercles rouges et des cercles bleus. La fonction  $h(x)$  est obtenue en combinant un vecteur d'entrée  $x$  avec un vecteur de poids  $w$ . Le vecteur de poids est déterminé en utilisant les données d'apprentissage.

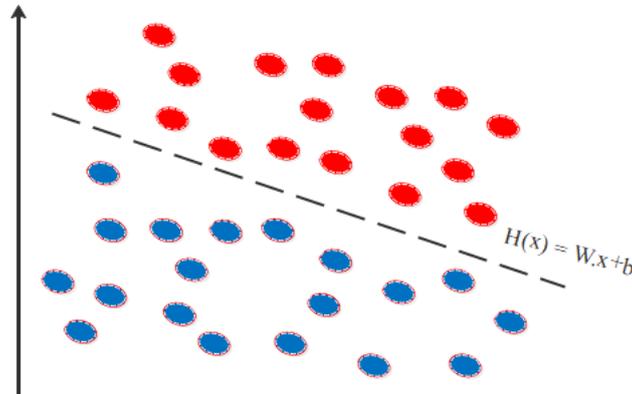


Fig 3.5: Hyperplan de base d'une machine à vecteurs de support linéaire

### 3.3 Partie 2 : Fouille de données pour la classification comportementale de vaches laitières

Le suivi des comportements (activité et alimentation) des vaches laitières est assuré en utilisant deux types de systèmes : les capteurs attachés (accéléromètre, GPS, etc.) et les caméras. Les données des capteurs attachés sont utilisées par des algorithmes d'apprentissage automatique (machine learning) pour localiser les vaches ou bien classer les différentes activités quotidiennes de chaque vache. Les données de caméra (images) sont utilisées par des algorithmes d'apprentissage profond (deep learning) pour assurer l'identification individuelle, la localisation et le suivi des comportements alimentaires des vaches. Dans ce qui suit, nous allons présenter la procédure utilisée par des études scientifiques pour détecter les comportements des vaches en utilisant ces deux types d'algorithmes.

#### 3.3.1 Classification comportementale par apprentissage automatique

Le suivi du comportement des vaches laitières prend de l'ampleur car il reflète l'état sanitaire des vaches comme il a été discuté dans le chapitre 2. Dans cette présente section, nous allons décrire la procédure de découverte de connaissances la plus communément utilisée dans les travaux concernant la classification des comportements. Cette procédure est illustrée par la Fig 3.6. Les travaux sont divisés en deux groupes. Les travaux qui se focalisent sur la construction de méthodes permettant un suivi des activités des vaches telles que « la position debout, couchée, la marche, etc. ». Ces activités peuvent être majeures ou secondaires. Les activités majeures sont la position debout, la position couchée, la marche et le stationnaire. Les activités secondaires sont incluses dans la position debout, on peut trouver le toilettage,

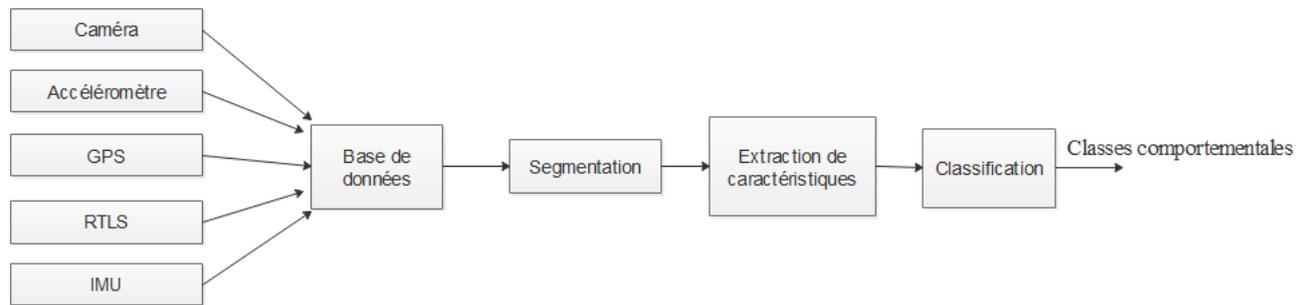


Fig 3.6: Processus d'apprentissage automatique appliqué dans la classification comportementale des vaches

les transitions entre la position debout et couchée, les coups de pied, etc. Dans le deuxième groupe des travaux, on trouve ceux qui se focalisent sur le suivi du comportement alimentaire. En effet, des systèmes d'exploration de données sont conçus pour calculer les variables liées à l'alimentation telles que le temps d'alimentation, le nombre de visites aux mangeoires, la quantité consommée, etc. La méthode de détection de tous ces comportements (activité et alimentation) se constitue de trois étapes qui sont : la construction de la base de données, la préparation de données et le déroulement de l'apprentissage. Le tableau 3.1 présente les étapes utilisées dans les articles scientifiques dans le cadre du monitoring comportemental. Dans ce qui suit, ces différentes étapes seront détaillées.

### 3.3.1.1 La construction de la base de données

Les travaux scientifiques classant les comportements de vaches laitières utilisent un ensemble de données préalablement acquises. Les bases de données contiennent souvent des données d'accéléromètres attachés au cou, au patte ou à la tête des animaux et des données de GPS. Par exemples, des données d'accéléromètres attachés au cou ont été acquises pour classer les activités debout, couchée et manger [29]. Dans [158], des données d'accéléromètres ont été utilisées pour classer les activités marche, couchée et le pâturage. La fusion de données d'accéléromètres et de GPS a été utilisée dans [159]. Récemment, les données d'unités de mesure inertielle ont été exploitées pour améliorer les résultats des accéléromètres [38, 160].

Pour concevoir la base de données, les capteurs d'acquisition sont déployés et ils se caractérisent par différents temps d'acquisition. Ces capteurs sont attachés à un nombre de vaches qui varie entre les études. De plus, ils fonctionnent selon plusieurs fréquences d'échantillonnages. En effet, le temps d'acquisition de données et le nombre de vaches utilisé différent grandement d'un travail à un autre. Dans [161] les auteurs ont récolté les données d'accéléromètre chez 5 vaches et durant 5 h par vache pour classer les activités debout et manger. Dans [156] les auteurs ont récolté les données d'accéléromètres chez 16 vaches pen-

Article	comportement	Capteurs	Échantillonnage	Segmentation	Caractéristiques	Algorithme D'apprentissage	Résultats
[38]	ruminer, voyager, couchée	GPS, accéléromètres	4 Hz et 10 Hz	10-s	la moyenne, l'écart type	Arbre de décision	Exactitude 90,5 %
[156]	Position debout, couchée et manger	Accéléromètres	1 Hz	1 min	Huit caractéristiques statistiques, énergie spectrale et entropie spectrale.	les k plus proches voisins, Naive Bayes, machine à vecteurs de support	Couchée précision (99 %) Manger précision 92 %
[157]	Manger, rumination	Accéléromètres	5 Hz	6,4 s, 12,8 s, 25,6 s, 51,2 s	minimum, premier quartile, médianes, troisième quartile, maximum, racine carrée moyenne, écart type et écart absolu moyen	les k plus proches voisins, machine à vecteurs de support et réseau de neurones	Manger précision 92,8 % rumination précision 93,7 %
[37]	En pâturage, position couchée, marcher	GPS	0.2 Hz	160 s	vitesse minimale, moyenne et maximale (m / s) et taux de changement de direction par segment	modèle de Markov caché	Exactitude 94 %
[29]	Position couchée, position debout, manger	Accéléromètres	50 Hz	1 min, 5 min et 10 min	accélération corporelle dynamique moyenne et vectorielle	Arbre de décision	couchée précision 98,63 % , debout précision 55,00 % , manger précision 93,10 %

Tableau 3.1: Synthèse des étapes d'acquisition, de prétraitement et de classification de données utilisées lors de la conception de systèmes de monitoring des comportements des vaches laitières par apprentissage automatique.

dant 6 h par vache pour classer les activités en position couchée, debout et manger. Dans [162] les auteurs ont conçu leur modèles à partir de données de 47 vaches avec un temps de 10 minutes par vache. De même, pendant le suivi du comportement d'autres espèces telles que les moutons, les chevaux et les agneaux, le temps d'acquisition de données et le nombre d'animaux utilisé varient considérablement. Dans [163] les auteurs ont récolté les données d'accéléromètre de 5 moutons durant 2,5 heures. Dans [164] les auteurs ont récolté les données d'accéléromètre chez 20 chevaux. Ils ont récolté les données durant 5 minutes pour chacune des activités (position debout et la marche). Dans [165] les auteurs ont récolté les données d'IMU de trois agneaux pendant 4 heures. Les travaux diffèrent aussi dans la taille de la fenêtre d'échantillonnage utilisée lors de l'acquisition. Les fréquences élevées permettent d'avoir une échelle de données complète. Cependant, elles alourdissent le processus d'apprentissage et posent des contraintes sur l'autonomie énergétique et la transmission des données des capteurs d'acquisition. Les basses fréquences fournissent une échelle de données d'une petite taille, mais garantissent une meilleure exécution en temps réel. Une grande fenêtre d'échantillonnage « 100 Hz » a été utilisée lors de l'acquisition des données d'accéléromètre dans plusieurs travaux, tels que [160, 166]. La taille moyenne de la fenêtre est généralement d'environ 10 Hz [27, 167, 168]. Cependant, d'autres travaux ont utilisé une petite fenêtre [156, 157, 161]. Il est a noté que les fréquences peuvent être sous échantillonné vers des fréquences plus basses. Par exemple, dans [156] les auteurs ont sous échantillonné les données de 1 Hz jusqu'à 0,25 Hz.

Une fois que les données brutes sont récoltées et insérées dans la base de données, des méthodes de préparation de données sont appliquées pour permettre de préparer les données pour la phase d'entraînement. Dans ce qui suit nous allons présenter la phase de préparation de données permettant de transformer la base de données brutes en une base de données d'entraînement.

#### **3.3.1.2 La préparation de données**

Le prétraitement de données est une phase importante dans la procédure d'apprentissage automatique. Les travaux qui classent le comportement des vaches laitières attribuent une grande importance à cette phase. Généralement, les auteurs segmentent les données brutes puis calculent ou extraient les variables pertinentes. Chaque segment est annoté selon le comportement visionné. Souvent, des caméras sont utilisées pour visualiser le comportement dans lequel les vaches se trouvent à chaque instant. Cette opération est faite pour créer la base de données d'entraînement.

Dans [169] une fenêtre de segmentation de 5 secondes a été utilisée. Puis les variables maximum, minimum, écart type et énergie des axes d'accéléromètre sont calculées. Dans

[156], les données des trois axes d'accéléromètre ont été fusionnées puis une fenêtre de segmentation de 1 minute a été appliquée. Pour chaque segment, les variables : minimum, premier quartile, médiane, troisième quartile, maximum, moyenne, racine carrée et écart type sont calculées. Dans [170] les auteurs ont réalisé la segmentation en temps réel dans le capteur de sorte qu'ils sauvegardent 10 secondes de données pendant chaque 15 minutes. Les segments ont été directement transformés en variables et cela en calculant la moyenne, la variance, le minimum et le maximum de chaque axe d'accéléromètre. Une fois que les données brutes ont été annotées, segmentées et transformées, la base de données sera utilisée lors de la phase d'apprentissage pour élaborer le modèle de classification.

### 3.3.1.3 Le déroulement de l'apprentissage

L'étape suivante est le déroulement de l'apprentissage où plusieurs algorithmes d'apprentissage automatique ont été utilisés. Pour assurer une meilleure exactitude, des algorithmes hybrides utilisant plusieurs classificateurs ont été aussi utilisés. Les méthodes d'apprentissage automatique qui sont souvent utilisées sont : les arbres de décision [29, 38], les machines à vecteurs de support [27, 171, 172], les réseaux de neurones [157], les k plus proches voisins [156, 157, 172] et les Modèles de Markov Cachés (MMC) [37, 173]. Les différents algorithmes peuvent être classés en deux catégories. La première est celle des algorithmes assurant un fonctionnement distribué. La structure hiérarchique simple proposée par les arbres de décision donne un grand atout car elle permet d'implémenter les modèles de classification directement dans les capteurs embarqués. Un arbre de décision ne nécessite pas une grande puissance de calcul et il est peu gourmand en énergie [174, 175]. Le capteur embarqué réalise la classification en temps réel et seulement une synthèse des résultats sera transmise aux utilisateurs. Cette synthèse est transmise dans un temps précis, généralement lorsque la vache est en traite. La deuxième catégorie est les algorithmes assurant un fonctionnement centralisé. Les algorithmes de SVM et MMC développés dans le suivi comportemental sont souvent gourmands en énergie et demandent un temps de calcul élevé. Pour cela, SVM et MMC sont souvent utilisés dans des systèmes centralisés de sorte que le capteur embarqué calcule seulement la caractéristique que suit l'algorithme puis la transmet par un réseau sans fil vers une borne centralisée. L'algorithme d'apprentissage est implémenté dans cette borne qui réalise tout le calcul. Les performances de SVM et MMC sont impressionnantes. Cependant, leurs implémentations dans un capteur embarqué n'est pas pratique [171, 174–176]. Les dispositifs embarqués doivent minimiser la consommation d'énergie alors que la mise en œuvre directe de ces techniques de classification peut nécessiter beaucoup de calcul, une puissance de traitement élevée et des mémoires de grande taille [174]. Pour cela, nous avons opté pour l'utilisation des arbres de décision comme modèle d'apprentissage lors de la conception

du système de détection des activités (chapitre 4). Des modèles d'apprentissage sophistiqués qui combinent plusieurs algorithmes d'apprentissage ont été développés. Ces modèles permettent d'optimiser les résultats de classification. Par exemple, dans [36] les algorithmes de classification et de regroupement ont été utilisés pour développer un modèle sophistiqué. Ces algorithmes sont l'analyse en composantes principales, k-moyennes floues et cartes auto-organisatrices pour le regroupement et les arbres binaires, l'analyse linéaire discriminante et réseaux bayésiens naïfs pour la classification.

Une fois que les algorithmes d'apprentissage ont été entraînés, les performances des modèles de classification élaborés doivent être calculées pour valider ces modèles.

#### 3.3.2 Classification comportementale par deep learning

L'apprentissage profond est une branche de l'apprentissage automatique qui améliore significativement les performances de l'analyse d'images et de la reconnaissance vocale. Il est récemment utilisé dans la classification des comportements de vaches laitières. Les algorithmes d'apprentissage profond sont utilisés pour classer les images issues de caméras installées dans des endroits fixes du bâtiment d'élevage qui sont quotidiennement fréquentés par les vaches. L'aspect non invasif offert par l'utilisation des caméras donne plus de confort aux animaux et permet d'éviter les contraintes des capteurs attachés (déchargement des batteries, endommagement des capteurs, etc). De plus, l'utilisation de caméras offre plusieurs avantages tels qu'un coût d'investissement faible et une aptitude à surveiller simultanément d'un nombre élevé de vaches avec seulement quelques caméras. Le traitement d'images apporte un autre avantage par sa capacité à comprendre la scène d'élevage. Les perspectives de son application sont illimitées.

Les différentes phases utilisées pour élaborer les modèles de monitoring des vaches en utilisant des méthodes de deep learning sont présentées dans la Fig 3.7. La première phase est la construction de la base de données. Tout d'abord, les caméras sont installées dans le bâtiment d'élevage afin de récolter les images et construire la base de données. En effet, les bases de données diffèrent d'un travail à un autre dans le nombre total d'images, la fréquence d'acquisition, la répartition des images d'entraînement et de test, la résolution des images acquises et le nombre de vaches. Dans [177] une base de données d'images de 10 vaches a été construite, elle contenait un ensemble de 1965 images. Ces images ont été divisées en 1667 images d'entraînement et 298 d'images de test. Le taux d'images acquises par seconde est de 25 images par seconde et la résolution des images est de  $1920 \times 1080$  pixels. Dans [178] une base de données de 46 430 images de 23 vaches a été construite où 90 % des images a été utilisé pour l'entraînement et 10 % pour le test.

La deuxième phase consiste à faire un prétraitement des images disponibles dans la base

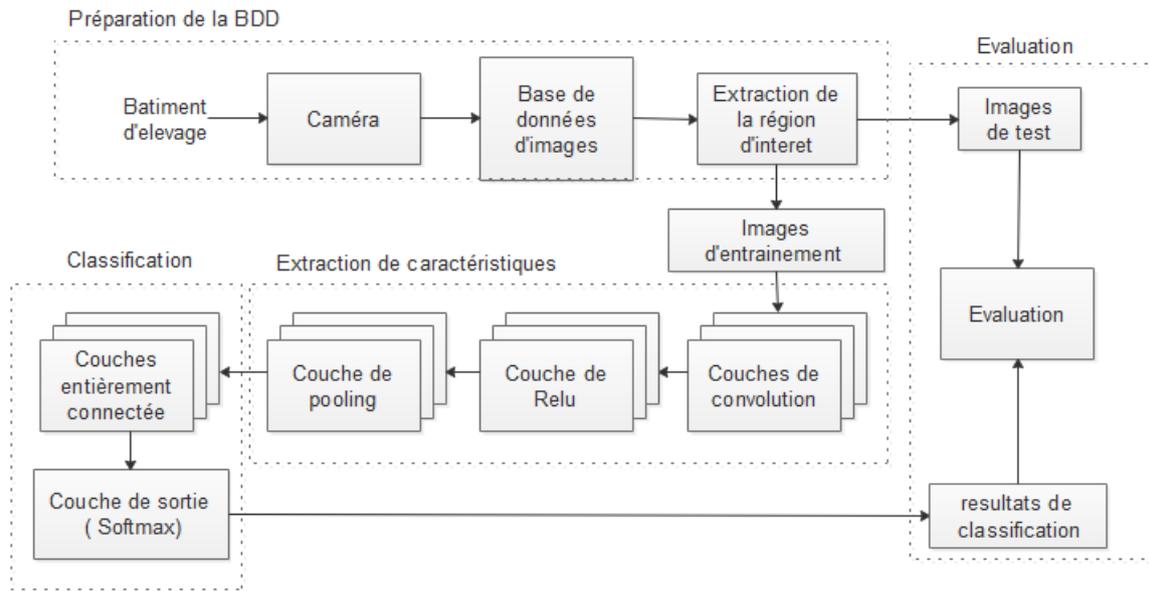


Fig 3.7: Architecture basée sur les réseaux de neurones à convolution utilisée lors de la classification comportementale des vaches

de données. Dans cette étape, des méthodes de réduction de données sont appliquées pour réduire le nombre d'images. Puis les images sont traitées pour segmenter la zone d'intérêt dite en anglais Region Of Interest (ROI). Cette zone est une partie physique de l'animal (le dos des vaches [177], le museau [179], la tête [180], etc) qui est suivie par l'algorithme d'apprentissage.

Pendant la troisième phase les algorithmes l'apprentissage profond sont utilisés pour l'extraction des caractéristiques et la classification des images. En effet, l'apprentissage profond permet une extraction automatique des meilleures caractéristiques lors du déroulement de l'apprentissage. Les réseaux de neurones à convolution sont le type d'algorithme le plus utilisé. Ils se composent de deux ensembles de couches, le premier est l'ensemble de couches d'extraction de caractéristiques dans lequel se trouve des couches de convolution et de pooling et le deuxième est l'ensemble de couches de classification. L'avantage des réseaux de neurones à convolution est que les caractéristiques sont extraites automatiquement pendant la phase d'apprentissage et que les erreurs de classification (pertes) sont minimisées afin d'optimiser les performances de classification. Cette optimisation est obtenue lors de l'ajustement des poids du réseau. Les réseaux de neurones à convolution donnent d'excellents résultats lors du monitoring du comportement alimentaire [178, 181–183] et ils peuvent être ré-entraînés dans le cadre de l'apprentissage par transfert (transfert learning). Pour cela, nous allons utiliser les réseaux de neurones à convolution de la conception de notre système d'analyse d'images permettant de suivre le comportement alimentaire (chapitre 5).

## 3.4 Conclusion

L'activité et l'alimentation des vaches laitières reflètent leurs états sanitaires. Le corps des vaches malades réduit significativement l'énergie liée au comportement afin d'augmenter sa résistance à la maladie ce qui conduit à une altération du comportement.

Dans ce troisième chapitre, nous avons présenté les différentes méthodes de fouille de données utilisées pour assurer le suivi automatisé du comportement des vaches. En effet, les méthodes de préparation des bases de données, le prétraitement de données et les algorithmes d'apprentissage utilisés par les études scientifiques ont été présentés. D'après ce chapitre, nous avons conclu que l'utilisation des arbres de décision est la moins coûteuse en énergie ce qui motive leur utilisation dans les systèmes de suivi d'activités de vaches. Pour cela, les arbres de décision seront utilisés comme méthode de classification par notre système de détection des activités. Les réseaux de neurones à convolution offrent, également, plusieurs avantages et ils sont très utilisés pour la classification d'images permettant la détection du comportement des vaches. Pour cela, les réseaux de neurones à convolution seront utilisés par notre système d'identification et de suivi du comportement alimentaire des vaches laitières.

Dans les prochains chapitres, nous allons présenter deux systèmes de monitoring comportemental. Le premier est développé pour assurer la classification des activités des vaches laitières à l'aide d'une méthode de classification non supervisée. Le second est un système qui se base sur l'utilisation de techniques d'analyse d'images pour assurer le monitoring du comportement alimentaire de vaches laitières.



# Chapitre 4

## Monitoring comportemental de vaches laitières à l'aide d'unités de mesure inertielle fixées sur leurs dos

### 4.1 Introduction

Les progrès technologiques ont permis la surveillance et la classification automatique des activités comportementales de nombreux animaux, y compris les espèces domestiquées telles que les vaches laitières [7]. La classification automatisée des comportements a le potentiel d'améliorer les processus de surveillance de la santé et du bien-être de l'animal. En effet, les enregistrements automatisés et continus du comportement peuvent être utilisés pour identifier les vaches malades au début du processus pathologique. Cela permet d'augmenter le succès du traitement [98, 184] et ainsi diminuer le taux de mortalité. Le suivi comportemental des vaches laitières prend de plus en plus d'importance dans la compréhension de la nutrition, la production, la gestion du bien-être et de la santé.

La surveillance des changements du comportement des animaux nécessite une observation de longue durée et une main-d'œuvre accrue. En pratique, cela est impossible dans les exploitations avec un grand nombre d'animaux ou pendant la nuit. Les systèmes de surveillance automatisés peuvent détecter tous les changements comportementaux d'un animal pendant 24 heures par jour. La détection de ces changements aide les éleveurs dans tout processus de prise de décision. Plusieurs animaux tels que les chèvres [185–187], les moutons [163, 185, 186], les truies [188], les porcs [189, 190], les iguanes roses [191], les chevaux [164, 192], les vaches laitières [27, 97] sont considérés par un suivi automatique à base de capteurs pour mesurer les données comportementales nécessaires au suivi de leurs santé, assurer leurs bien-être, les

protéger des attaques des animaux sauvages, détecter les périodes de chaleur qui indiquent le moment approprié pour l'insémination artificielle, etc.

Dans cette étude nous nous intéressons au suivi comportemental des vaches laitières. Par rapport aux autres animaux d'élevage (chèvres, mouton, etc.), l'élevage des vaches laitières est caractérisé par l'augmentation de la taille des troupeaux et donc l'observation visuelle du comportement devient de plus en plus complexe. Cette limitation de l'observation visuelle directe du comportement implique un besoin croissant d'une détection automatisée des changements de comportement.

L'objectif de cette présente étude est de développer, d'implémenter et d'évaluer une méthode de classification non supervisée des activités comportementales des vaches laitières logée en stabulation libre. Le but est de discriminer les comportements à partir de données brutes issues des unités de mesure inertielle attachées aux dos des vaches, et ce à l'aide d'une méthode de classification qui se base sur une analyse statistique et de probabilités dite méthode « des modèles de mélanges finis » permettant de créer un arbre de décision performant et peu exigeant en énergie. Cet arbre de décision permet le calcul du temps passé en position couchée et debout, le nombre et le moment de transition entre les positions debout et couchée et l'activité (marche et stationnaire). Les performances de cette classification sont évaluées en termes de sensibilité, spécificité, exactitude et précision. Les résultats de ce travail ont été publiés dans Achour et al, 2019 [35].

## 4.2 Matériel et méthodes

La procédure de développement du système de classification des activités comportementales d'une vache en temps réel se compose de six (6) étapes comme l'illustre la Fig 4.1.

1- Développement, déploiement et collecte de données: cette étape vise à concevoir des capteurs pour l'acquisition des données brutes (données d'IMU et de temps) et de les sauvegarder localement sur une carte mémoire.

2- Prétraitement et traitement des données brutes : les données brutes collectées sont divisées en deux sous-ensembles à savoir les données de test et les données de validation. L'ensemble des données sera transformé et agrégé selon différentes fréquences d'échantillonnage et différentes tailles de fenêtres de segmentation, permettant de construire une base de données pour l'analyse statistique.

3- Analyse statistique et construction de l'arbre de décision : dans cette étape, l'analyse des données prétraitées est réalisée dans le but de calculer et de sélectionner les variables les plus pertinentes. Ces variables seront ensuite utilisées pour créer l'architecture de l'arbre de décision.



Fig 4.1: Enchaînement de la procédure de développement du système de classification comportementale.

4- Ajustement de seuils : les seuils de l'arbre seront ajustés en utilisant des modèles de mélanges finis.

5- Validation et optimisation des résultats : l'arbre de décision sera ensuite validé grâce à l'ensemble des données de validation. Les résultats seront optimisés à l'aide d'une méthode de fusion de données et d'un algorithme issu d'automates à états finis.

#### 4.2.1 Étape 01 : Développement, déploiement et collecte de données

Le monitoring comportemental des vaches laitières se base sur des capteurs de mouvement ou de localisation. Les accéléromètres sont les capteurs de mouvement les plus utilisés pour surveiller les activités des vaches laitières. La surveillance du comportement à l'aide d'un système de localisation permet seulement de fournir des informations sur l'emplacement du bétail (zone de repos, mangeoire, traite, abreuvoir, etc.). Actuellement, on utilise beaucoup les unités de mesure inertielle (IMU) pour collecter les paramètres qui permettent la classification comportementale. Ces IMU sont composées d'un accéléromètre 3 axes, d'un gyroscope 3 axes et d'un magnétomètre 3 axes. L'avantage des IMU est qu'en plus de la mesure de l'accélération linéaire (offerte déjà par les accéléromètres), les IMU mesurent aussi l'accélération angulaire et estiment la trajectoire, ce qui élargit l'ensemble des données utilisé. Pour cela, dans notre étude nous allons utiliser les données d'IMU pour classer les activités des vaches laitières.

L'équipement d'acquisition que nous avons conçu est illustré par la Fig 4.2-b. Il se compose d'une unité de mesure inertielle (SparkFun MPU-9250 9 Dof IMU Breakout), d'un microcontrôleur (ATmega2560, Arduino Mega ADK Rev3), d'un module de stockage de données (SD Card 8 Go) et de deux Leds. Tous ces composants ont été raccordés et scellés sur une ceinture attachée au dos de l'animal. L'IMU dispose d'un module MPU-9250 (9 degrés de liberté) avec un système intégré qui combine deux puces : le MPU-6500 qui contient un gyroscope à 3 axes ainsi qu'un accéléromètre à 3 axes et l'AK8963 qui comprend un magnétomètre à 3 axes.

Dans notre étude, contrairement aux autres, nous avons choisi d'attacher le capteur aux dos des vaches comme le montre la Fig 4.2-a. Ce choix est motivé par le fait que le dos est un endroit stable et sûr (moins de perturbations et de risque d'éraflure ou de briser le capteur) qui est directement impacté par les activités des vaches. En plus, cet emplacement permet d'intégrer plusieurs capteurs et offre la possibilité de suivre un large éventail de phénomènes sanitaires, reproductifs et morphologiques tels que la contraction musculaire, l'acceptation de chevauchement, la température des mamelles, la mesure de la circonférence des vaches, etc.

Le système a été programmé pour collecter des données IMU à 80 Hz (soit 288 000 lignes de données/heure). A la fin de l'acquisition, les cartes de stockage sont retirées et les données sont téléchargées. Le système d'annotation étant un système d'enregistrement vidéo, les données acquises ont été annotées à l'aide d'analyses visuelles du comportement de la vache.

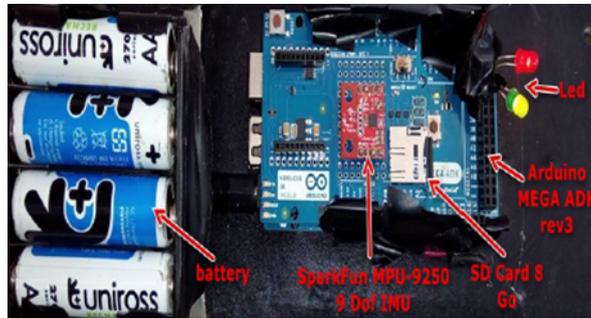
Les expériences sur terrain se sont déroulées en décembre 2017 dans une ferme à stabulation libre pour vaches laitières située à Draâ Ben Khedda wilaya de Tizi Ouzou. L'acquisition de données a duré huit jours avec une moyenne de 6 h par jour (entre 9 h et 15 h). Le capteur a été protégé avant d'être installé sur les animaux (imperméabilité à l'eau).

Huit vaches de races différentes (deux PRIM HOLSTEIN, trois MONTBÉLIARDS et trois FLECKVIEH) ont été sélectionnées pour l'expérience. La capture de données fournit un total de 16 signaux : 9 signaux sont collectés par l'IMU (accéléromètre 3D (Acc\_x, Acc\_y, Acc\_z), le gyroscope 3D (Gyr\_x, Gyr\_y, Gyr\_z) et le magnétomètre 3D (Mag\_x, Mag\_y, Mag\_z). Ensuite, le microcontrôleur ATmega2560 reçoit ces signaux en entrée et utilise le filtre de Madgwick [193] pour calculer les quaternions 4D (q0, qx, qy, qz). Après cela, la conversion des quaternions aux angles d'Euler 3D (Yaw, pitch, Roll) est effectuée en temps réel. En plus de ces 16 signaux, 3 signaux temporels (heure, minute, seconde) sont enregistrés.

Le schéma fonctionnel donné dans la Fig 4.3 est un schéma montrant le flux de données échangé entre les composants matériels et logiciels. Les flèches se pointent de la source vers la destination. Les rectangles représentent les composants matériels et les ovales représentent des modules logiciels. Nous utilisons ce graphe de flux de données dans la conception de haut niveau car il décrit le fonctionnement global du système, tout en masquant les détails de la



(a) Emplacement de l'équipement d'acquisition sur le dos de vache laitière



(b) Composants du système

Fig 4.2: Equipement d'acquisition de données

façon dont il fonctionne.

Les données collectées de trois vaches laitières (vache 01, vache 02, vache 03) sont utilisées lors de l'étape de développement du modèle de classification et les données acquises des huit vaches (vache 01, vache 02..., vache 08) sont utilisées lors de l'étape de validation du modèle comme indiqué dans le tableau 4.1.

#### 4.2.2 Étape 2 : Prétraitement et traitement des données brutes

**Choix de comportement suivi : Modélisation de comportement à l'aide d'automates à états finis déterministes** Un automate à états finis déterministes (AFD) permet de modéliser de nombreux systèmes. Les AFD utilisent une description formelle et montrent la manière dont les systèmes se comportent et évoluent. Un AFD est composé d'un ensemble fini d'états (représentés graphiquement par des cercles), d'une fonction de transition décrivant l'action qui permet de passer d'un état à un autre (les flèches), d'un ou plusieurs états initiaux (marqués par un arc sans nœud d'origine), d'un ou plusieurs états finaux (désignés par un double cercle).

Dans ce travail, nous avons modélisé le comportement à l'aide d'automates d'états finis

Vache Id	Durée de la vidéo		durée d'acquisition		Étapes utilisant les données
	[h]	[min]	[h]	[min]	
Vache 01	6	27	6	17	development/validation
Vache 02	6	30	6	21	development/validation
Vache 03	5	8	5	0	development/validation
Vache04	6	44	5	10	validation
Vache 05	5	41	5	34	validation
Vache 06	6	42	6	35	validation
Vache 07	8	37	8	30	validation
Vache 08	7	14	7	35	validation
<b>durée totale</b>	53	3	50	3	--

Tableau 4.1: Description de la base de données

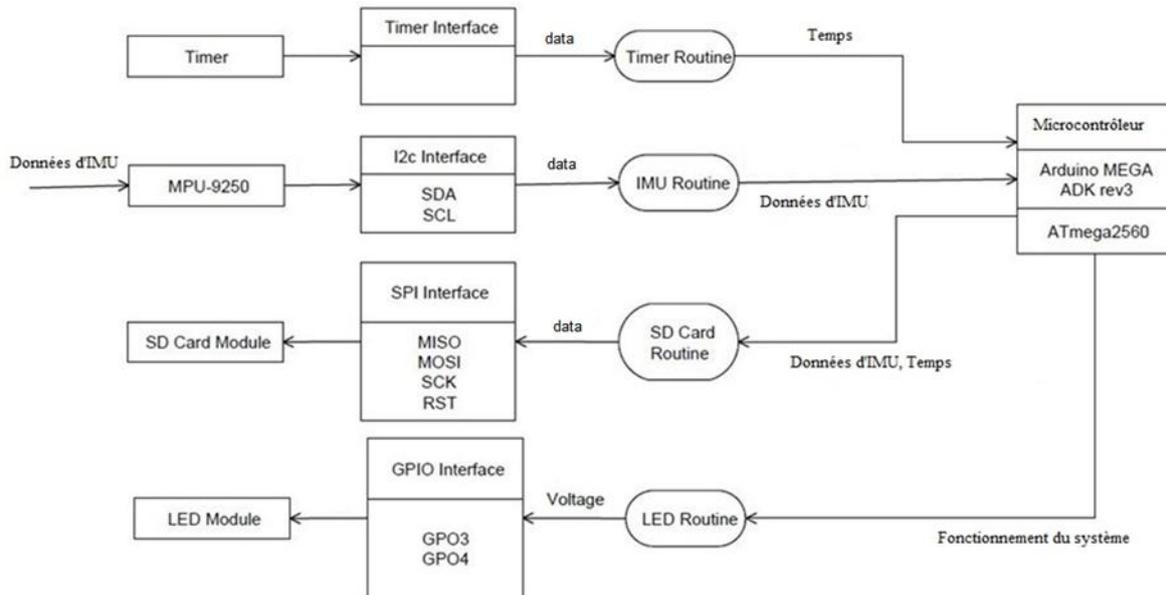
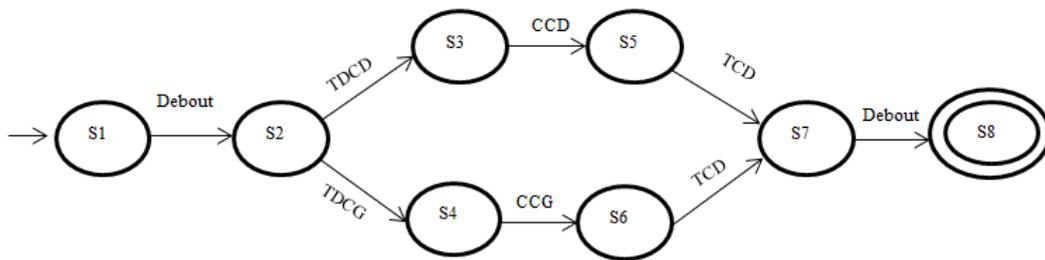


Fig 4.3: Le schéma fonctionnel matériel du système d'acquisition de données.

déterministes pour visualiser les états comportementaux dans lesquels une vache pourra se trouver et pour décrire les transitions possibles entre ces classes comportementales selon un ordre valide. Le comportement global de l'animal peut être modélisé à l'aide de deux automates. Le premier est illustré par la Fig 4.4. Il décrit les comportements liés à la posture de l'animal (position debout, couchée et transition entre ces positions) et permet également de poser des contraintes sur l'enchaînement de ces comportements. Par exemple la série comportementale : vache debout – vache couchée – vache marche est fautive (une vache ne peut pas marcher en étant en position couchée), par contre la série comportementale : vache debout - vache en transition debout/couchée – vache couchée – vache en transition couchée/debout - vache debout est valide. Le second automate est illustré par la Fig 4.5. Il montre les différents comportements dans lesquels une vache pourra se trouver tout en étant debout. L'arborescence donnée par la Fig 4.6 montre les différents comportements suivis dans la présente étude.

**Transformation et traitement des données** Les mouvements des vaches ont été observés sur les vidéos et classifiés manuellement en trois classes principales et six sous-classes comme l'illustre le tableau 4.2.

Les données de L'IMU ont été visualisées sur une interface Matlab que nous avons développée. Cette interface est montrée par la Fig 4.7. Elle permet d'afficher les signaux IMU en fonction du temps. La comparaison de ces signaux avec les activités visualisées dans les enregistrements vidéo permet de déterminer le temps minimal requis pour effectuer chacune des



Où : TDCD : Transition Debout Couchée côté Droit. TDCG : Transition Debout Couchée côté Gauche. CCD : Coucher vers le Côté Droit. CCG : Coucher vers le Côté Gauche. TCD : Transition Couchée/Debout.

Fig 4.4: Automate de modélisation des comportements de base d'une vache laitière.

activités comportementales suivies dans cette étude. Les résultats de cette comparaison ont permis de définir les temps suivants : 3 secondes pour les activités marche et stationnaire, 10 secondes pour les activités se coucher et se lever et 5 secondes pour la position debout ou couchée. Ces temps seront utilisés comme fenêtres de segmentation lors du prétraitement des données.

Les données brutes sont tout d'abord segmentées en utilisant les trois fenêtres de segmentation précédemment définies. Après l'étape de segmentation, plusieurs variables statistiques ont été calculées en appliquant différentes fréquences d'échantillonnage à chaque segment. Il est à noter que la fréquence de 80 Hz a été utilisée lors de l'acquisition des données. Après avoir testé plusieurs fréquences sous-échantillonnées, nous avons constaté que les fréquences 1 Hz et 4 Hz donnaient de meilleurs résultats de classification. Pour cela, pendant toutes les autres étapes, ces fréquences sont utilisées. En effet, à partir de chaque segment, un vecteur de caractéristiques a été construit. Ces vecteurs contiennent 14 caractéristiques. Pour chaque caractéristique, huit variables statistiques ont été calculées. Les mesures extraites pour obtenir le vecteur de caractéristiques sont données par le tableau 4.3. Les différentes étapes de segmentation, du sous échantillonnage et du calcul des variables statistiques sont représentées dans la Fig 4.8.

### 4.2.3 Étape 03 et 04 : Analyse statistique et construction de l'arbre de décision

L'étape précédente permet de calculer 112 variables statistiques. L'objectif de l'analyse statistique est de sélectionner les variables pertinentes pour discriminer les activités comportementales. Ensuite, ces variables seront utilisées pour établir l'architecture de l'arbre de décision. Comparés aux autres algorithmes de classification (régression logistique, SVM, etc.), les ar-

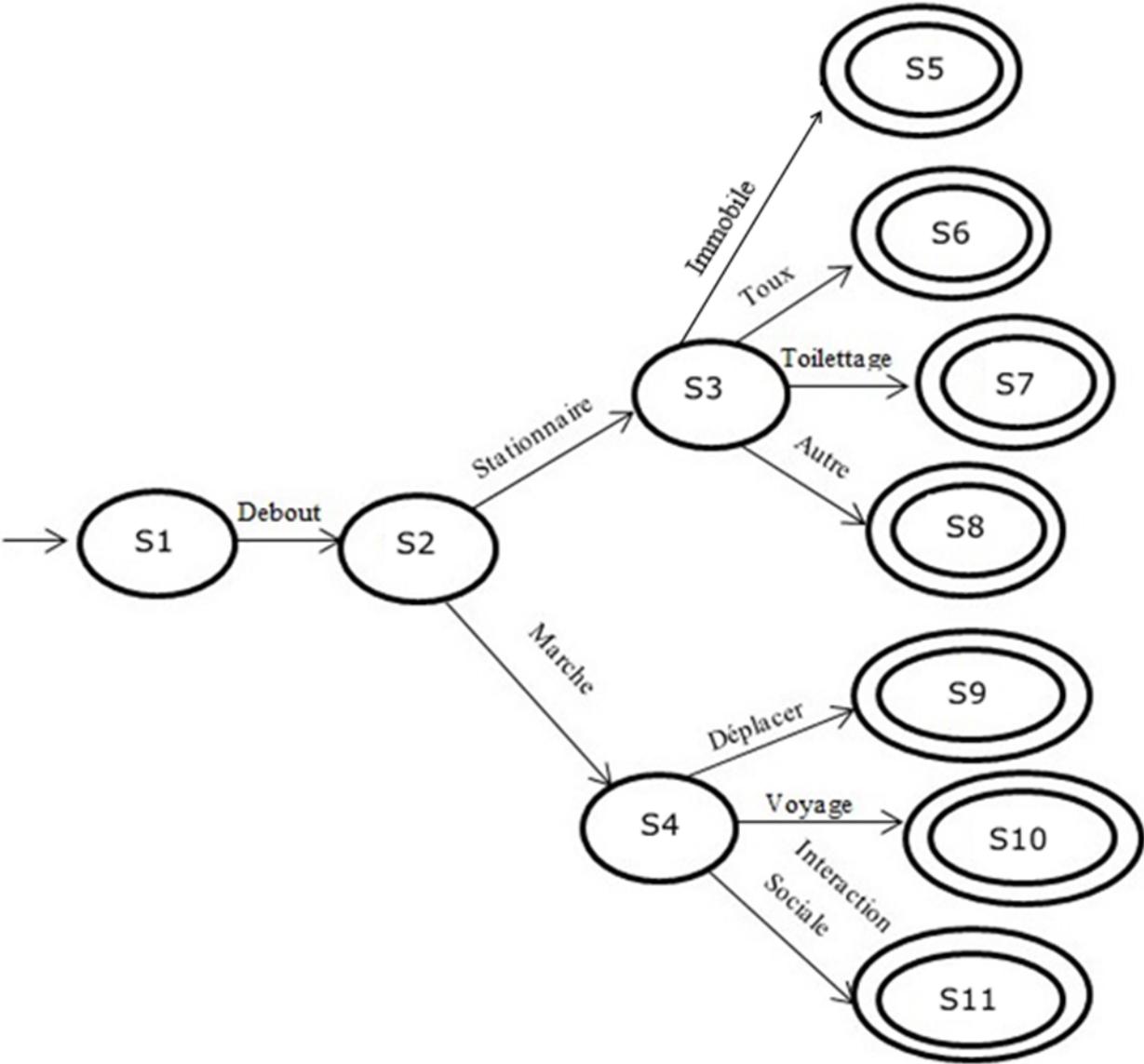


Fig 4.5: Automate de modélisation des différents comportements d'une vache se trouvant dans l'état debout.

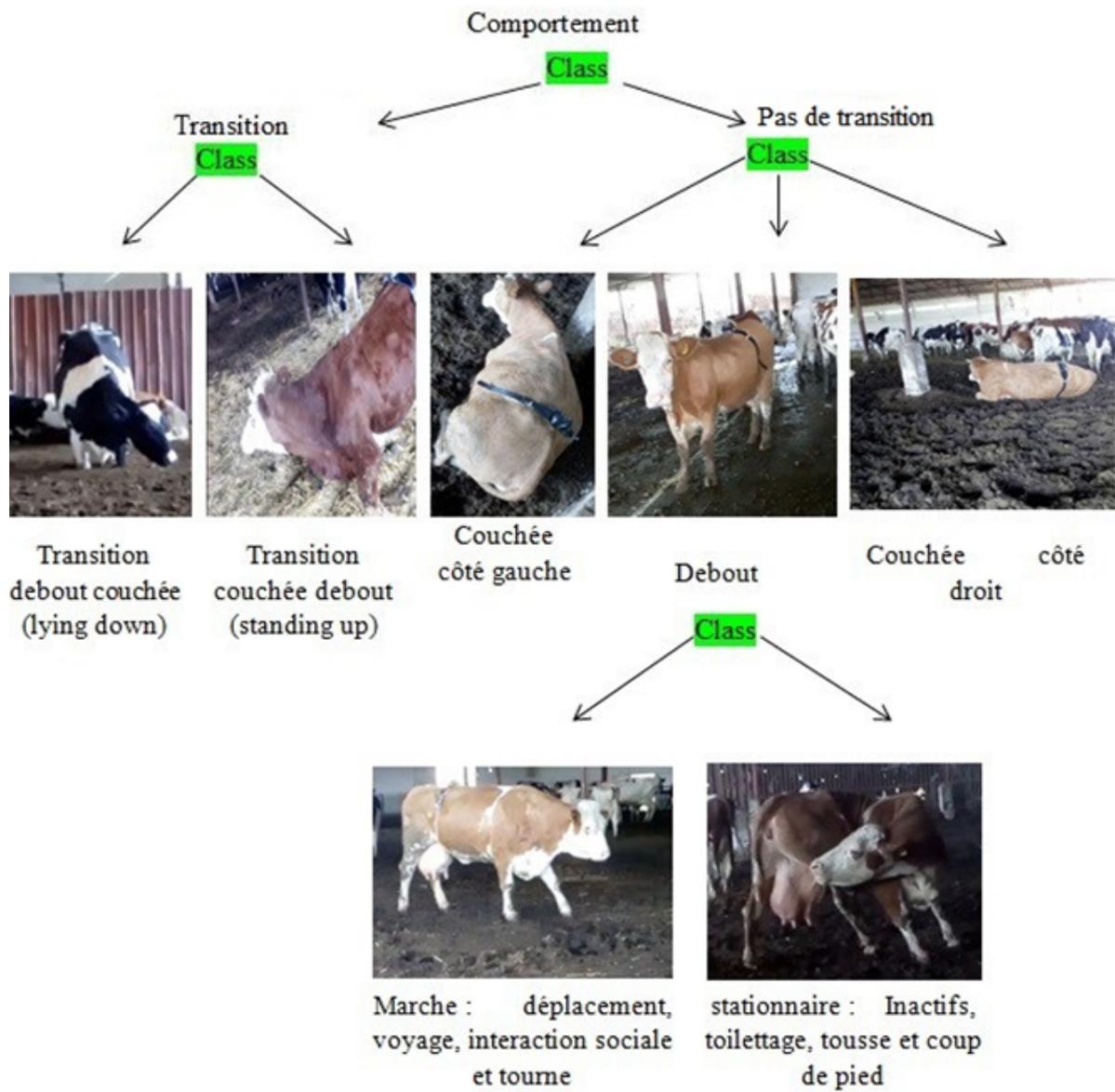


Fig 4.6: Arborescence de modélisation des comportements suivis.

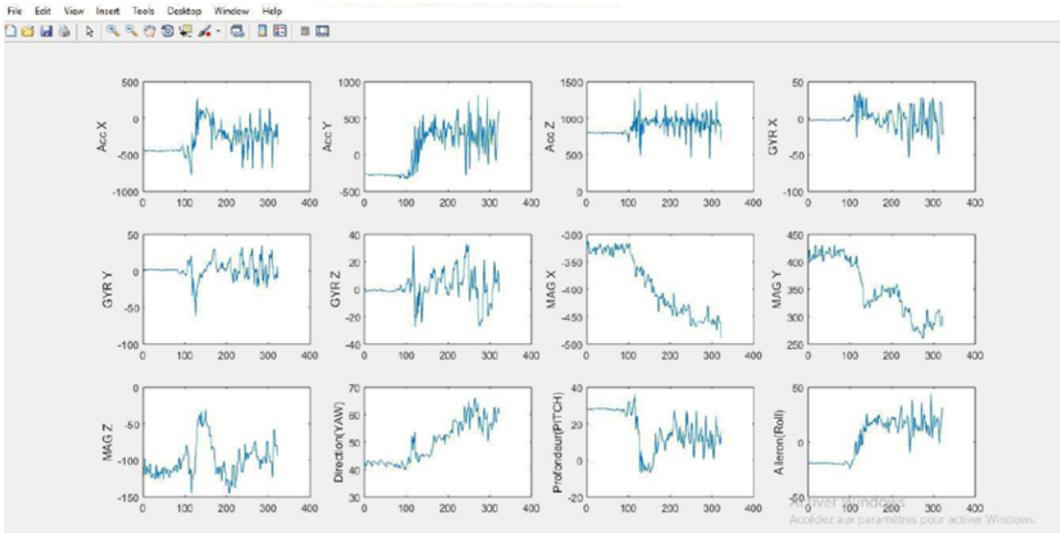


Fig 4.7: Interface Matlab permettant de visualiser les signaux acquis.

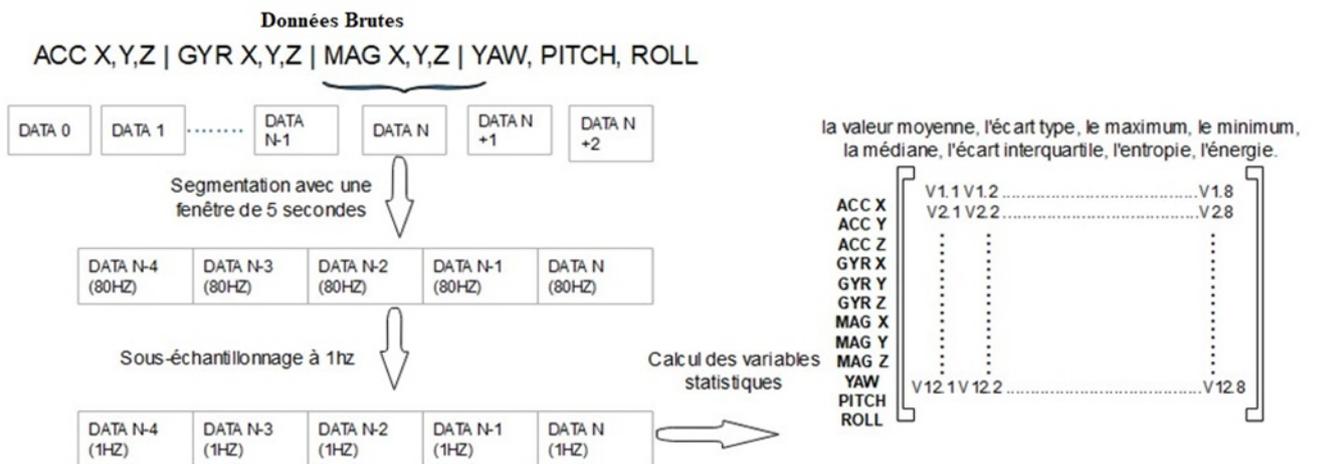


Fig 4.8: Exemple montrant les étapes de la procédure de traitement des données en utilisant une fenêtre de segmentation de 5 secondes et une fréquence de 1 Hz.

Comportement	prévalence	Définition	Sous comportement
<b>Debout</b>	74,05 %	Lorsque la vache est en position debout et aucune de ses pattes ne frottent la terre	<p><b>Stationnaire</b> : lorsque la vache est inactive (immobile) ou bien elle effectue un mouvement actif dans lequel elle utilise une patte (toiletage, coup de pied, etc.) ou moins (toiletage, toux, etc.).</p> <p><b>Marche</b> : lorsque la vache effectue un mouvement dans lequel elle utilise ses 4 pattes (marche, voyage, en déplacement) ou elle utilise deux pattes (rotation ou interaction sociale).</p>
<b>Couchée</b>	25,67 %	Lorsque la vache est en position couchée, deux de ses pattes frottent la terre	<p><b>Couchée côté droit</b> : lorsque la vache est strictement couchée vers le côté droit de son corps, elle frotte le sol avec ses pattes droites.</p> <p><b>Couchée côté gauche</b> : lorsque la vache est strictement couchée vers le côté gauche de son corps, elle frotte le sol avec ses pattes gauches.</p>
<b>Transition</b>	0,28 %	Lorsque la vache est en période de transition entre les positions debout et couchée	<p><b>Se coucher</b> : lorsque la vache est en transition de la position debout vers la position couchée. Ce mouvement commence lorsque ses deux pattes avant frottent le sol et se termine lorsque la vache se trouve en position couchée.</p> <p><b>Se lever</b> : lorsque la vache est en transition de la position couchée vers la position debout. Ce mouvement commence lorsqu'elle soulève ses pattes arrière pour se lever et se termine lorsque la vache est en position debout.</p>

Tableau 4.2: Classes comportementales suivies

Caractéristique extraite	Variable calculée
Accélération Total Ax, Ay, Az	la valeur moyenne, l'écart type,
Valeur de gyroscope Gx, Gy, Gz	l'entropie, la médiane,
Valeur de magnétomètre Mx, My, Mz	l'intervalle interquartile, l'énergie, le maximum et le minimum.
Ongle d'Euler Yaw, Pitch, Roll	
Combinaison de signaux d'accéléromètre [Acc_x , Acc_y , Acc_z]	
Combinaison de signaux de gyroscope [Gyr_x, Gyr_y, Gyr_z]	

Tableau 4.3: Caractéristiques extraites de l'IMU et variables calculées pour chaque caractéristique

bres de décision sont extrêmement intuitifs et fournissent une représentation graphique explicite et facile à interpréter.

Pour choisir les variables statistiques qui ont la plus grande aptitude pour la classification des différentes activités, nous utilisons la régression linéaire à effets mixtes. En effet, un modèle de régression linéaire est un modèle de régression qui cherche à établir une relation linéaire entre une variable dite expliquée et une ou plusieurs variables dites explicatives. Le modèle linéaire à effets mixtes est une extension du modèle linéaire qui prend en compte la variabilité liée aux individus. Ce modèle est composé d'une partie fixe et d'une partie aléatoire. La partie fixe est identique pour chaque individu et représente l'effet de population. La partie aléatoire est propre à chacun des individus et traduit la variabilité liée à chaque sujet. De ce fait, nous avons utilisé le modèle linéaire à effets mixtes pour décrire la variation de chaque caractéristique en fonction de la classe comportementale dans laquelle se trouve une vache. Cela, permet de sélectionner les meilleures caractéristiques qui montrent la plus grande variation entre les classes comportementales.

#### 4.2.4 Étape 05 : Ajustement des seuils

L'étape d'analyse statistique des variables a permis de créer l'architecture globale de l'arbre de décision et de choisir les meilleures caractéristiques qui reflètent au mieux la variabilité entre les classes comportementales. Le but de cette étape est d'ajuster les seuils de cet arbre de décision en utilisant le modèle des mélanges finis. Ce dernier fait partie des méthodes de

modélisation statistique ayant des typologies qui ne reposent pas sur des modèles statistiques formels, mais plutôt sur une démarche itérative et intuitive qui consiste à regrouper les individus les plus semblables et éloigner les individus les plus différents. L'intérêt majeur du modèle des mélanges finis est de fournir un cadre rigoureux pour évaluer le nombre de groupes (clusters) dans un contexte de classification non supervisée [194].

Le modèle des mélanges finis consiste à supposer que les données proviennent d'une source contenant plusieurs sous-populations homogènes appelées composants. La population totale étant un mélange de ces sous-populations. Une loi de mélanges finis « p » sur un espace « X » est une loi de probabilité s'exprimant comme une combinaison linéaire de plusieurs lois de probabilité  $P_1, P_2, \dots, P_g$  sur X. autrement dit, il existe g coefficients

$\pi_1, \pi_2, \dots, \pi_g$  pour tout  $x_1 \in X$ , tel que :

$$P(x_1) = \sum_{K=1}^g \pi_K P_K(x_1) \quad (4.1)$$

les  $\pi_K$  et  $P_K$  sont respectivement appelées proportions et composantes du mélange.

La classification automatique par loi des mélanges finis est une procédure mathématique qui propose une ou plusieurs partitions parmi une multitude de possibilités offertes. Les lois de probabilité les plus connues sont : la loi normale ou loi de Laplace Gauss, la loi exponentielle, la loi du khi-deux, la loi de Student, la loi de Fisher Snedecor (loi F), Loi de Weibull, etc. et chaque composante du mélange montre le meilleur ajustement à une distribution particulière. Dans cette présente étude, la nature et le nombre de lois composant chaque mélange sont déterminés visuellement en observant la distribution des fonctions de densité de chaque population de points de données. La combinaison de distributions avec le meilleur ajustement a été sélectionnée et cela à l'aide du Package mixtools disponible sur R [195].

## 4.2.5 Étape 06 : Validation et optimisation des résultats

### 4.2.5.1 Validation

L'objectif de l'arbre est de détecter précisément les classes comportementales dans lesquelles se trouve l'animal. Cependant, par erreur l'arbre peut classifier une classe comme étant une autre classe. De ce fait, pour estimer la performance de la classification de l'arbre de décision de nombreuses mesures peuvent être utilisées. Nous considérons quatre critères d'évaluation de performance : la sensibilité, la spécificité, la précision (valeur prédictive positive) et l'exactitude de la classification. La sensibilité d'un test mesure sa capacité à donner un résultat positif lorsqu'une hypothèse est vérifiée. À l'opposé, la spécificité mesure la capacité d'un test à donner un résultat négatif lorsque l'hypothèse n'est pas vérifiée. Les fonctions des

Mesure de validation	Equation	Signification
<b>Sensibilité (Se)</b>	$VP/(VP+FN)$	Aptitude du système à identifier toutes les classes
<b>Spécificité (Sp)</b>	$(VN)/(VN+FP)$	Aptitude du système à n'identifier que les vrais changements de classes
<b>Exactitude (Exct)</b>	$(TP+TN)/(TP+TN+FP+FN)$	L'exactitude globale de la classification
<b>Précision (Pr)</b>	$VP/(VP+FP)$	Probabilité qu'une vache classée est réellement dans cette classe

Où : VP (vrai positif) : le nombre d'instances où l'état comportemental a été correctement classé par l'algorithme après validation par l'observateur visuel. FN (faux négatif) : le nombre d'instances où l'état comportemental a été observé visuellement dans la réalité, mais a été incorrectement classé comme un autre comportement par l'algorithme. FP (faux positif) : le nombre de fois que l'état comportemental était incorrectement classé par l'algorithme, mais non observé dans la réalité. VN (vrai négatif) : nombre d'instances où l'état comportemental a été correctement classé comme n'étant pas observé.

Tableau 4.4: critères d'évaluation de la validation de l'arbre de décision.

quatre critères de validation sont présentés dans le tableau 4.4.

#### 4.2.5.2 Optimisation des résultats et validation finale

Une analyse plus poussée a été effectuée afin d'atteindre des taux de classification plus élevés. Cette analyse se base sur la fusion de données qui permet de réduire considérablement les taux de faux positifs [196]. A titre d'exemples, la fusion des variables d'accéléromètre et de gyroscope lors de la détection des chutes chez les personnes âgées a optimisé considérablement les performances de détection [197, 198] et la fusion des variables issues de l'accéléromètre tout seul montre aussi une bonne capacité d'optimisation [199].

Dans notre cas, la méthode de fusion de données est composée de deux phases, comme le montre la Fig 4.9.

- Phase 1 : Sélection de variables pertinentes ;
- Phase 2 : Classification par modèles de mélanges multivariés, construction de l'arbre de décision finale et la validation.

Dans la première phase, les variables qui montrent une excellente sensibilité pour la classification d'une activité quelconque ont été sélectionnées et ajoutées à la liste de variables pertinentes. Pour construire cette liste, une classification par modèle de mélange univarié a été réalisée pour chacune des 112 variables calculées pendant l'étape d'analyse statistique. Le nombre de lois utilisées pour cette classification est identique à celui choisi pour la classification des variables sélectionnées par l'analyse statistique. Dans la seconde phase, une classification multivariée a été effectuée à l'aide de ces variables pertinentes en utilisant l'outil Mclust sur R [200]. Pour avoir un compromis entre la sensibilité et la précision de classifica-

tion, seulement des mélanges de deux ou de trois variables ont été réalisés. Puis un arbre de décision finale a été construit. Cet arbre permet de mélanger et de fusionner toutes les variables sensibles à la détection de la classe souhaitée. Cela permet de garder une bonne sensibilité de classification tout en optimisant l'exactitude, la spécificité et la précision. Les nouveaux taux de classification ont été calculés et les résultats de classification ont été comparés à ceux issus des arbres de décision à variable unique (voir section 4.3.3.2).

## 4.3 Résultats

### 4.3.1 Analyse statistique

Les résultats de cette phase montrent que chaque comportement impacte une variable particulière et que chaque variable donne des résultats optimaux selon différentes fenêtres de segmentation.

Les variables médiane et intervalle interquartile de l'axe-x de l'accéléromètre ont une grande capacité de classification, qu'elles gardent même aux plus faibles fréquences (1Hz) comme le montre le tableau 4.5. La médiane de l'axe-x de l'accéléromètre distingue les classes d'activités liées à la position du corps de la vache (Fig 4.10-a). Cette variable est dotée d'une grande précision dans la séparation des classes debout, couchée vers le côté droit et couchée vers le côté gauche. En effet, la Fig 4.10-a montre les résultats de la régression linéaire mixte basée sur les données de trois vaches laitières pour la classification en position debout et couchée (à gauche ou à droite) en utilisant la variable médiane de l'accéléromètre de l'axe-x avec une segmentation de 5 secondes et une fréquence de 1 Hz. Pendant l'acquisition des données de la vache 3, aucune période de couchée vers le côté droit n'a été observée, ce qui explique l'absence de cette population de données dans la Fig 4.10-a. L'intervalle interquartile de l'axe-x de l'accéléromètre indique toutes les transitions entre les positions debout et couchée (Fig 4.10-b).

L'énergie de l'axe-y du gyroscope (Fig 4.10-c) a une bonne capacité pour classifier les mouvements de l'animal en marche ou stationnaire comme l'indique le tableau 4.5-b. Mais classification n'est pas optimale car elle dépend de la manière dont l'animal effectue son activité. Généralement, cette variable reflète trois classes :

- Classe 1 : Les activités à haute énergie correspondant à l'un des mouvements suivant : marche, tourne ou opère une interaction sociale.
- Classe 2 : Les activités à moyenne énergie correspondant au comportement de toilettage, coup de pied ou toux.

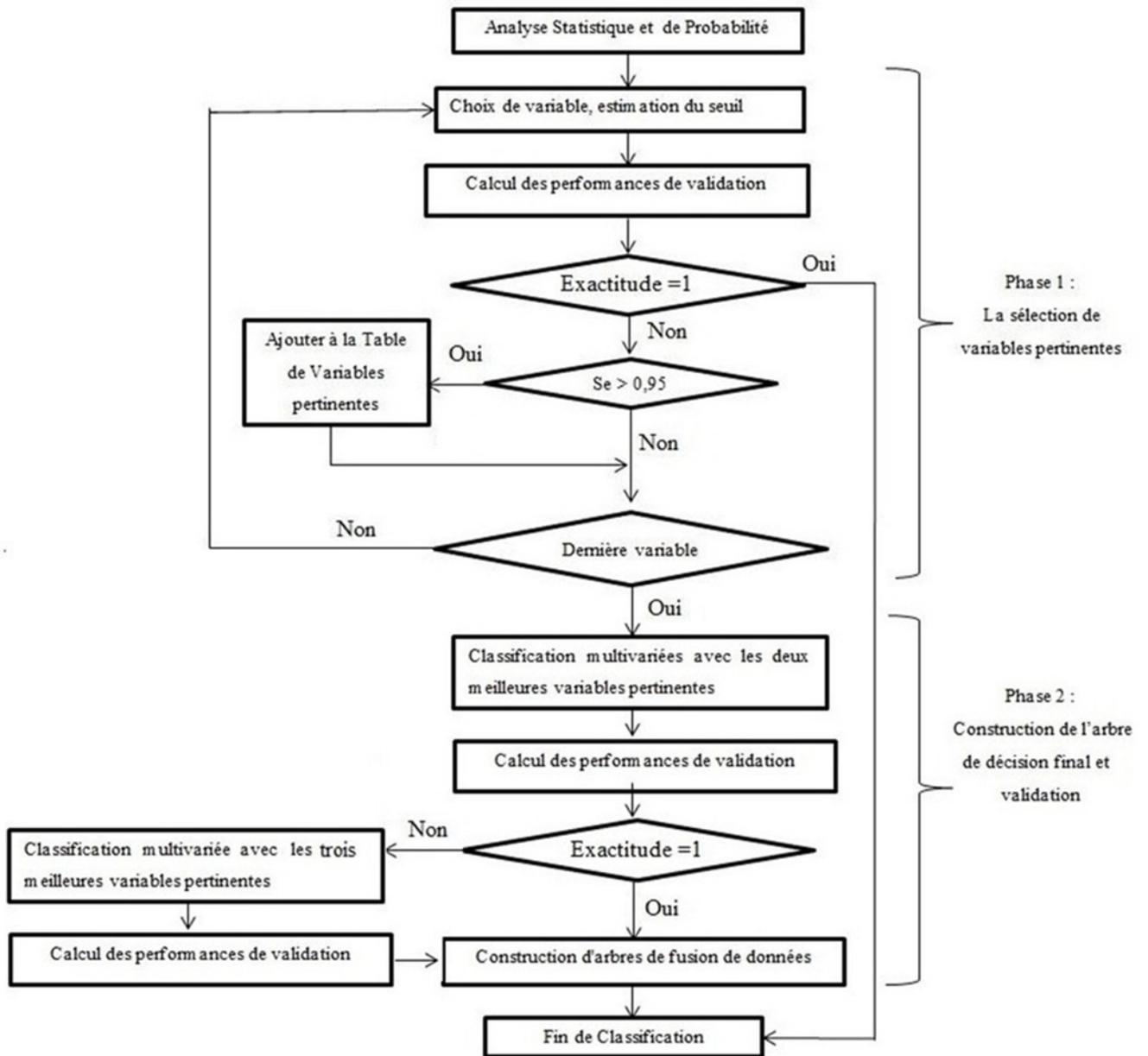


Fig 4.9: Méthode de fusion de données.

- Classe 3 : Les activités à basse énergie où l'animal est inactif.

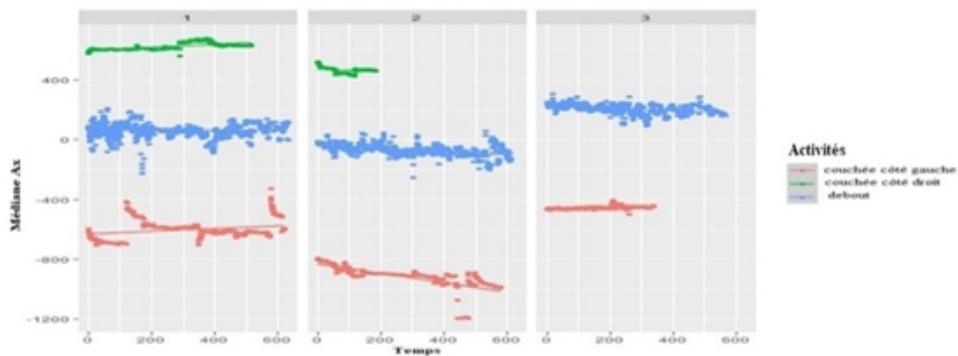
Occasionnellement, des activités de la deuxième classe peuvent se trouver dans la classe une. Par exemple, l'énergie de l'activité de toilettage faite de façon agressive en levant sa patte sera élevée et elle correspondra, de ce fait, à l'énergie des activités la classe une. Inversement, lorsque la vache se déplace très lentement sur une courte distance l'énergie de cette activité sera faible et correspondra à celle fournie par la classe deux.

L'analyse par régression linéaire mixte à variable aléatoire a indiqué deux types de variables. Les premières sont celles qui ne sont pas affectées par la variabilité inter animal : se sont les variables qui séparent les classes debout, couchée et transition. L'utilisation de ces variables offre la possibilité de déterminer de manière optimale les seuils directement à partir du capteur sans analyse centralisée (Fig 4.10-a, Fig 4.10-b). Les secondes sont celles qui sont affectées par la variabilité inter animal : se sont les variables qui séparent les classes marche et stationnaire. Elles ont une bonne précision pour séparer ces deux classes comportementales (marche et stationnaire). Mais les seuils changent entre les vaches. Les seuils de ces variables doivent être uniques pour chaque vache et sont déterminés par un système centralisé (Fig 4.10-c).

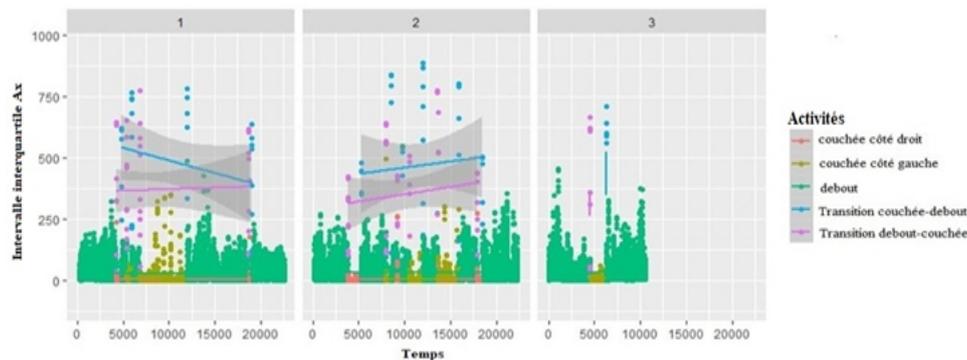
Lorsqu'une vache est en position debout ou couchée, l'orientation de l'IMU change considérablement. Ce changement peut être exploité directement par le capteur pour en déduire la position de la vache. Cependant, l'énergie moyenne de l'animal change considérablement selon sa race, sa parité et son état physiologique. Lorsque l'animal marche, tourne ou opère une interaction avec un autre animal, la valeur énergétique de l'axe-y du gyroscope est élevée et donne une bonne aptitude de classification pour extraire ces classes comportementales. De plus, plusieurs autres variables telles que : l'énergie de l'axe-x de l'accéléromètre et l'énergie de l'axe-y de l'accéléromètre sont affectées par ces activités comme le montre le tableau 4.5.b.

### 4.3.2 Résultats d'ajustement de seuil

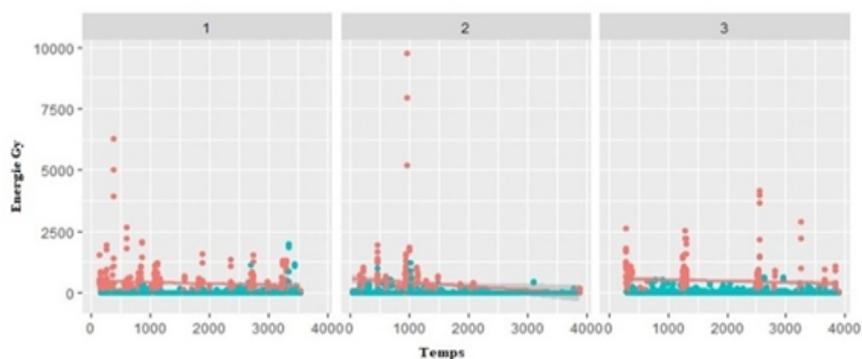
Sur la base de son habilité à séparer les classes debout, couchée côté droit et couché côté gauche, la variable médiane de l'axe-x de l'accéléromètre a été sélectionnée pour vérifier que cette variable contient les différentes populations de données qui peuvent représenter les comportements. Tandis que l'intervalle interquartile de l'axe-x de l'accéléromètre a été choisi pour séparer les classes comportementales debout, couchée et transition. La variable énergie de l'axe-y du gyroscope a été sélectionnée pour différencier l'état stationnaire de l'état marche. Les distributions de fonction de densité de ces trois variables sélectionnées pour chaque comportement sont présentées dans la Fig 4.11.



(a) Résultat de la régression linéaire mixte à variable aléatoire sur les données de trois vaches de races différentes et de parités différentes pour la classification des activités debout, couchée vers le côté droit et couchée vers le côté gauche à l'aide de la variable médiane d'accélération de l'axe-x (segmentation 5 s ; 1 Hz).



(b) Régression linéaire mixte à variable aléatoire sur les données de trois vaches de races différentes et de parités différentes pour la classification des activités se coucher et se lever à l'aide de la variable intervalle interquartile de l'axe-x de l'accéléromètre (segmentation de 10 s ; 1 Hz).



(c) Régression linéaire mixte à variable aléatoire sur les données de trois vaches de races différentes et de parités différentes pour la classification des activités stationnaire et marche à l'aide de la variable énergie de l'axe-y du gyroscope (segmentation de 3 s ; 1 Hz).

Fig 4.10: Régression linéaire mixte sur les données de trois vaches de races différentes et de parités différentes pour la classification de diverses activités à l'aide de plusieurs variables.

variable	Axe	Couchée Droit	Debout	Couchée-Gauche	Se-coucher	Se-lever	RSE *	R <sup>2</sup> *	p-valeur
médiane Acc 1 Hz 5 sec	X	462,675	-137,359	-934,347	-132,282	-234,85	68,62	0,970 6	< 2,2 e-16
	Y	143,856	-128,932	122,732	-80,59	116	67,51	0,757 9	< 2,2 e-16
Intervalle interquartile Acc 1Hz 10 sec	Z	469,560	808,331	518,835	773,307	564,535	59,83	0,844 1	< 2,2 e-16
	X	5,217	28,585	7,481	456,934	532,308	33,22	0,394 1	< 2,2 e-16
	Y	6,0265	28,0762	7,7714	219,073	259,268	26,72	0,241 3	< 2,2 e-16
	Z	4,801 1	12,1277	7,1547	197,433	198,923	17,65	0,250 3	< 2,2 e-16

(a) Régression linéaire entre les comportements et les caractéristiques (médiane et intervalle interquartile des axes de l'accéléromètre)

variable	Axe	Stationnaire	Marche	RSE	p-valeur	R <sup>2</sup>
Gyr Énergie 1Hz 3 Sec	X	5,525	137,540	132,8	< 2,2 e-16	0,089 85
	Y	11,13	382,17	207,1	< 2,2 e-16	0,242 6
Acc Énergie 1Hz 3 Sec	Z	31,525	363,016	185,9	< 2,2 e-16	0,240 7
	X	14633,6	1308	5601	< 2,2 e-16	0,361 1
	Y	10307,2	831	4396	< 2,2 e-16	0,316 8
	Z	5322,5	319,4	2829	< 2,2 e-16	0,238

(b) Régression linéaire entre les comportements (marche et stationnaire) et les caractéristiques (énergie des axes du gyroscope et énergie des axes de l'accéléromètre)

Tableau 4.5: Régression linéaire qui montre l'influence des activités comportementales sur les variables suivies.

Où : RSE est l'erreur standard résiduelle dite en anglais Residual Standard Error, R<sup>2</sup> le coefficient de détermination et P-valeur est la valeur de probabilité que le résultat de la régression se produit essentiellement par hasard.

### 4.3.2.1 Détermination du nombre de lois de mélange fini

Pour la médiane de l'axe-x de l'accéléromètre, en opérant une segmentation de 5 secondes et une fréquence de 1 Hz, les populations de points de données correspondant à la position debout, la position couchée côté droit et la position couchée côté gauche sont bien séparées (Fig 4.11-a). L'analyse de la Fig 4.11-a indique la présence de trois populations clairement distinctes. La première population inférieure à - 400 unités de gravité représente la position couchée vers le côté droit, la deuxième population comprise entre -400 et 200 unités de gravité représente la position debout et la troisième population supérieure à 200 unités gravité représente la position couchée vers le côté gauche. Suite à cette analyse la suggestion suivante est émise : les valeurs des seuils pour la séparation de ces activités pourront être ajustées à l'aide d'un mélange fini de trois lois normales. Grâce aux grands espacements qui existent entre les différentes populations de données cette séparation se fera d'une façon stricte avec des taux de classification optimale.

La distribution des fonctions de densité de la variable intervalle interquartile de l'axe-x de l'accéléromètre avec une segmentation de 10 secondes et une fréquence de 1 Hz est caractérisée par deux populations de points de données. La première population représente les états dans lesquels la posture de la vache a été soit debout ou couchée et la seconde population de points représente les périodes de transition entre la position debout et la position couchée.

L'analyse de la Fig 4.11-b révèle deux populations clairement séparées et distinctes. La première population inférieure à 200 unités de gravité représente la position debout ou couchée vers l'un des deux côtés et la deuxième population supérieure à 200 unités de gravité représente l'état de transition entre la position debout et la position couchée. Suite à cette analyse la suggestion suivante est émise : la valeur seuil pour la séparation de ces activités pourra être ajustée à l'aide d'un mélange fini de deux lois normales. Le grand espacement qui existe entre les différentes populations de données montre que des taux de classification élevés peuvent être atteints. Cependant, sur l'échelle de données les états debout et couché regroupés représentent 99,72 %. Tandis que, les transitions représentent 0,28 % (tableau 4.2). Cette grande différence oblige de faire des séparations préliminaires d'une façon récursive avant d'atteindre la séparation finale et cela en appliquant plusieurs modèles de mélange afin d'arriver à un modèle de mélange final qui permettra de trouver le bon seuil. Le nombre et la nature de la loi de ces mélanges préliminaires et du mélange final doivent être identifiés.

La distribution de fonctions de densité de la variable énergie de l'axe-y du gyroscope avec une segmentation de 3 secondes et une fréquence de 1 Hz est caractérisée par trois populations de points de données. La première population représente les états dont lesquels la vache était inactive (basse énergie, vache immobile) et n'effectuait aucun mouvement significatif. La seconde population de points représente la vache en mouvement de moyenne énergie (vache

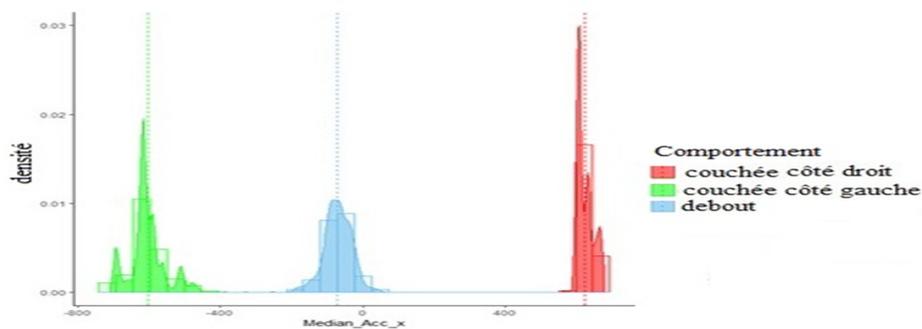
bouge mais n'utilisant qu'une seule ou aucune patte) tel que le toilettage, coup de pied, etc. La dernière population correspond au mouvement de haute énergie tel que la marche (déplacement, voyage) et l'interaction sociale (vache utilisant deux pattes ou plus pour effectuer ces activités). Les deux premières populations correspondent au comportement de stationnaire et la troisième population correspond au comportement de marche.

L'analyse de la Fig 4.11-c montre qu'au moins deux populations de données peuvent être séparées. La première population inférieure à 85 unités de gravité représente l'activité stationnaire et la deuxième population supérieure à 85 unités de gravité représente l'activité marche. Suite à cette analyse la suggestion suivante est émise : la valeur de seuil pour la séparation de ces activités pourra être trouvée à l'aide d'un mélange fini de deux ou trois lois normales. Les résultats varieront selon le choix du nombre de lois. Le choix de deux lois normales (seuil = 40 unités gravité) favorise la sensibilité de détection de sorte que tous les états de marche vont être détectés et le choix de trois lois normales (seuil = 85 unités gravité) favorise la précision de classification de sorte que tous les états détectés comme étant une marche seront exacts. Donc la séparation globale se fera d'une façon moins optimale avec des niveaux de classification acceptables et cette séparation nécessite une analyse plus poussée en combinant d'autres signaux tels que : l'énergie de l'axe-x et de l'axe-y de l'accéléromètre.

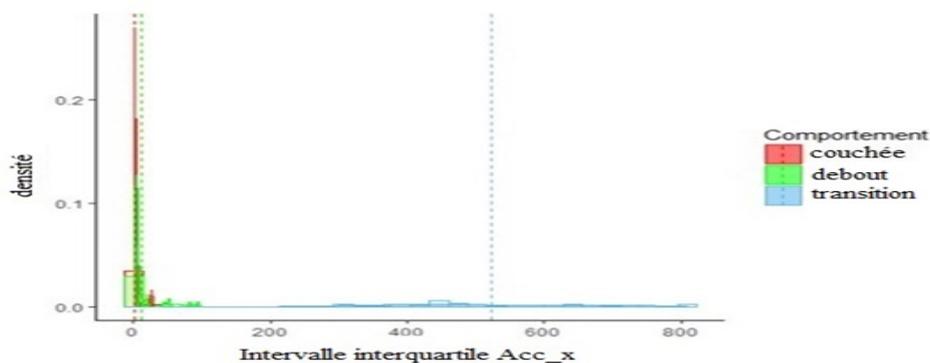
#### **4.3.2.2 Détermination des seuils à l'aide des modèles de mélange**

Pour la médiane des données de l'axe-x de l'accéléromètre, le meilleur ajustement est un mélange de trois distributions de lois normales (Fig 4.12) pour séparer les activités debout, couchée vers le côté droit et couchée vers le côté gauche. L'intervalle interquartile des données de l'axe-x de l'accéléromètre a donné le meilleur ajustement par l'utilisation de trois mélanges consécutifs. Les deux premiers mélanges de deux lois normales sont préliminaires (Fig 4.13-a, Fig 4.13-b) et servent à réduire l'échelle de données et ils sont suivis d'un troisième mélange de deux lois normales (Fig 4.13-c). Ce dernier permet de trouver le seuil optimal qui sépare la classe transition entre les positions debout et couchée des autres classes. Un autre modèle avec trois mélanges consécutifs sépare les classes stationnaire et marche à l'intérieur de la classe debout préalablement séparée. Le premier mélange de deux lois normales est préliminaire (Fig 4.14-a) et il permet de séparer l'activité immobile des autres activités. Ce modèle est suivi d'un deuxième modèle qui se compose d'un mélange de deux lois normales (Fig 4.14-b) et sert à séparer les classes marche et stationnaire en optimisant la sensibilité de la méthode. De même le troisième mélange sert à séparer les classes marche et stationnaire, mais en optimisant la précision de classification (Fig 4.14-c). Ce troisième mélange se compose d'un mélange de deux lois normales.

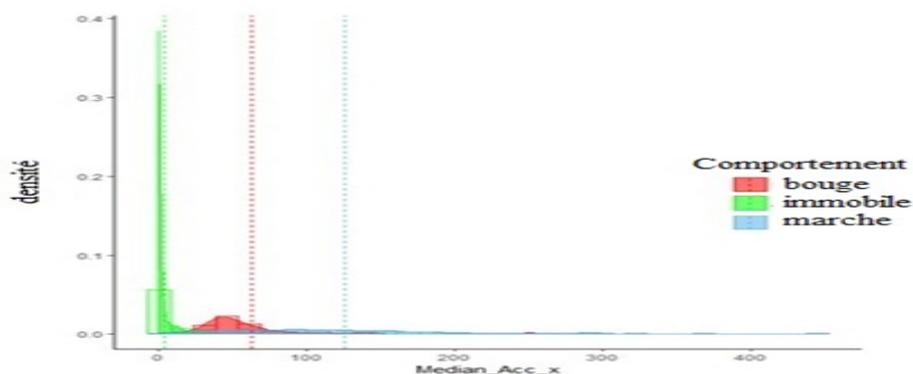
La valeur du seuil A (Fig 4.13-c) sépare la population de données de la transition de-



(a) Les trois fonctions de densité des classes comportementales debout, couché côté droit, et couché côté gauche utilisant la variable médiane de l'axe-x d'accéléromètre avec une segmentation de 5 secondes et une fréquence de 1 Hz.



(b) Les trois fonctions de densité des classes comportementales debout, couchée et transition entre position debout et couchée en utilisant la variable intervalle interquartile d'axe-x de l'accéléromètre avec une segmentation de 10 secondes et une fréquence de 1 Hz.



(c) Les trois fonctions de densité des classes comportementales immobile, en mouvement (toux, coup de pied, toiletage) et marche (déplacement, voyage, interaction sociale, tournée) en utilisant la variable énergie de l'axe-y du gyroscope avec une segmentation de 3 secondes et une fréquence de 1 Hz.

Fig 4.11: Fonctions de densité spécifiant les classes comportementales suivies selon les variables choisies.

bout/couchée des autres classes. L'intersection entre la première et la deuxième population de données du troisième mélange (Fig 4.13) est utilisée comme première étape de la classification des points de données et permet de trouver les moments exacts de transition et de les séparer de l'échelle globale de données.

Les valeurs des seuils B et C (Fig 4.12) séparent les populations de données des positions debout, couchée côté droit et couchée côté gauche. L'intersection entre ces populations de données est utilisée comme deuxième étape dans le procéder de classification des points de données.

Les transitions entre les états décrites dans la section précédente par le premier automate d'état fini déterministe et montré dans la Fig 4.4 est utilisé pour identifier le type de transition entre les positions debout et couchée sans utiliser une nouvelle variable. Cet automate permet de faire une annotation automatique de l'ensemble des transitions en quatre types : 1. transition debout/couchée côté droit, 2. transition debout/couchée côté gauche, 3. transition couchée côté droit/debout et 4. transition couchée côté gauche/debout et cela en combinant le résultat des classifications précédentes (debout, couché côté droit, couché côté gauche et transition) avec leur ordre chronologique d'apparition.

Le seuil D est déterminé à partir de deux mélanges consécutifs. Le premier mélange montré par la Fig 4.14-a est préliminaire et sert à extraire une partie de l'activité stationnaire (vache immobile) de l'échelle de données. Puis le deuxième mélange sert à déterminer la valeur du seuil D (Fig 4.14-b) qui sépare la population de données des activités stationnaire et marche incluses dans la position debout. La valeur du seuil D correspond à la valeur de l'énergie de l'axe-y du gyroscope où la première et la deuxième fonction de densité de probabilité du mélange B se croisent. La première population représente la classe stationnaire et la deuxième population représente la classe marche. Cette dernière classe est elle-même divisée en utilisant un troisième mélange de deux nouvelles lois normales afin de trouver le seuil E (Fig 4.15-c). Le seuil E est la valeur de l'énergie de l'axe-y du gyroscope où la première et la deuxième fonction de densité de probabilité du mélange C se croisent. Les seuils D et E ont le même rôle dans la classification. Mais, leurs résultats sont différents de sorte que le seuil D optimise la sensibilité et le seuil E optimise la précision de classification.

### 4.3.3 Optimisation des résultats

Les résultats de classification obtenus avec le modèle de mélange univarié étaient très satisfaisants pour les classes debout et couchée du côté droit ou côté gauche comme indiqué dans le tableau 4.6 et acceptables pour toutes les autres classes comme indiqué dans les tableaux 4.5 et 4.7. Pour optimiser les résultats de la classification, nous avons procédé comme suit :

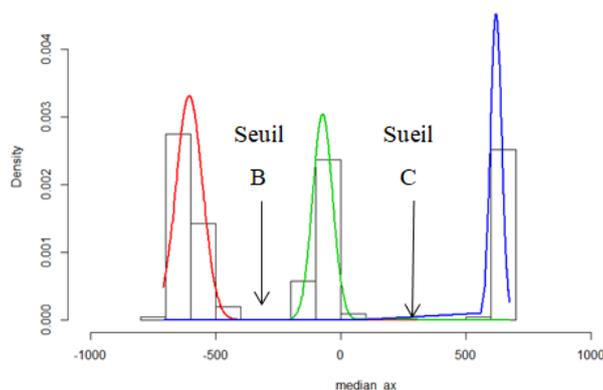


Fig 4.12: Modèle de mélange de trois lois normales de la variable médiane de l'axe-x de l'accéléromètre avec une segmentation de 1 seconde et une fréquence de 1 Hz (vache1).

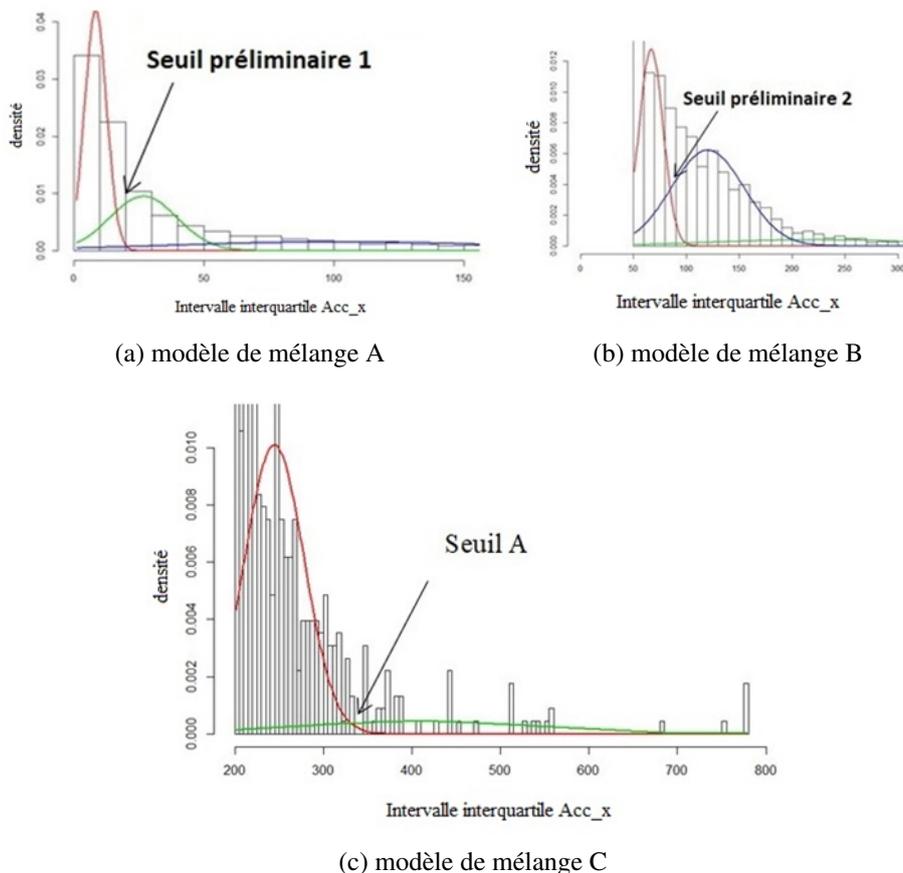


Fig 4.13: Modèles de mélange consécutifs (A, B, C) de 3 lois normales (a) suivies de 3 lois normales (b) et de 2 lois normales (c) de la variable intervalle interquartile de l'axe-x de l'accéléromètre utilisée pour séparer la classe transition des autres classes.

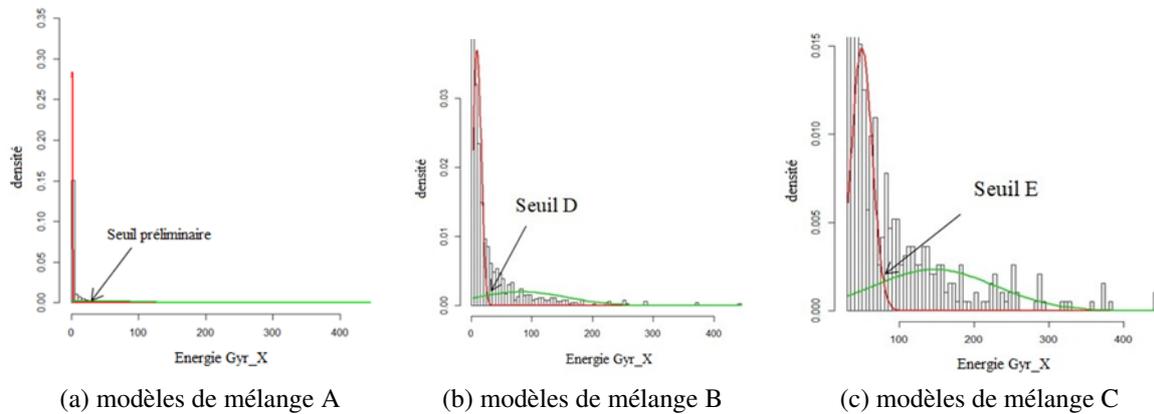


Fig 4.14: Modèles de mélange consécutifs (A, B, C) utilisés pour la séparation des classes stationnaire et marche. (a) Mélange préliminaire de deux lois normales, (b) Mélange de deux lois normales qui déterminent le seuil D et (c) Mélange de deux lois normales qui déterminent le seuil E.

#### 4.3.3.1 Par automates

La classification des transitions (couchée/debout) par un modèle de mélange univarié a une sensibilité de 100 %, mais une précision de 59 % seulement. Le modèle de mélange univarié détecte toutes les transitions sans faux négatifs. La seule erreur qui existe est liée aux faux positifs. Pour cela, chaque transition trouvée par ce modèle doit être vérifiée. Selon l'automate, une transition valide se situe certainement entre les états debout et couchée. Pour vérifier cela, une étude sur les données qui précèdent et succèdent le moment des transitions est réalisée.

Les contraintes définies par l'automate sont utilisées pour valider chaque transition trouvée de sorte que chaque transition valide doit être comprise entre deux états différents (position debout ou position couchée). Pour valider chaque transition une nouvelle variable qui reflète l'espacement (ESP) entre les données qui précèdent et qui succèdent le moment de transition est déterminée. Si la valeur de cette variable est significative (supérieure au seuil  $S = 200$  unités de gravité) la transition est valide sinon la transition est invalide (Fig 4.15). L'espacement est déterminé en calculant l'éloignement qui existe entre les données des quatre minutes qui précèdent et succèdent la transition lorsqu'un modèle de mélange fini de deux lois normales est appliqué. Cette validation automatique permet d'atteindre un taux de précision optimal (100 %) comme le montre le tableau 4.7.

L'équation (4.2) ci-dessous est utilisée pour estimer la variable d'espacement en temps réel. Pour ce faire, le capteur enregistre en continu la valeur médiane de l'axe-x de l'accéléromètre (avec une fréquence de 1 Hz et une segmentation de 5 secondes) lorsque la valeur de l'écart type du même segment de données est petite (inférieure à 10 unités de gravité, ce qui reflète

Modèle	couchée côté droit				debout				couchée côté gauche			
	Se	Sp	Pr	Ex	Se	Sp	Pr	Ex	Se	Sp	Pr	Ex
Uni varié	1,00	0,99	0,99	0,99	0,99	1,00	1,00	0,99	1,00	1,00	1,00	1,00
Multi varié	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

Tableau 4.6: Résultats de classification des classes debout, couchée côté droit et couchée côté gauche à l'aide des deux modèles de mélange univarié et multivarié.

une faible variance). L'espacement (ESP) est déterminé en soustrayant la première médiane après la transition de la dernière médiane avant la transition. Cette validation automatique atteint une précision de 100 %. Le tableau 4.7 présente les résultats obtenus.

$$Esp = (\text{Dernière médiane avant la transition}(\text{Axe}_x)[\text{Écart-type}(\text{Axe}_x) < 10] - \text{Première médiane après la transition}(\text{Axe}_x)[\text{Écart-type}(\text{Axe}_x) < 10]) \quad (4.2)$$

#### 4.3.3.2 Fusion de données (modèle de mélange multivarié)

La classification univariée a montré une excellente capacité pour la classification des classes debout, couchée et la transitions entre ces positions et cela en utilisant deux variables : la médiane et l'intervalle interquartile de l'axe-x de l'accéléromètre. Cette capacité de classification a été améliorée à l'aide de la classification multivariée.

Pour la classification des classes debout, couchée côté droit et couchée côté gauche deux variables ont été fusionnées : la médiane de l'axe-x et de l'axe-y de l'accéléromètre avec une segmentation de 5 s et une fréquence de 1 Hz (Fig 4.16). Cette combinaison permet d'améliorer légèrement les résultats fournis par le modèle univarié (tableau 4.6) avec des populations de points mieux séparées et un résultat de classification optimal (100 %).

Pour la classification de la classe transition debout/couchée et les autres classes, une similitude de résultat a été trouvée grâce à une combinaison de deux signaux qui sont l'intervalle interquartile de l'axe-x et de l'axe-y de l'accéléromètre (Fig 4.17), en utilisant quatre mélanges consécutifs (A, B, C, D). Tous les mélanges se composent de deux lois normales. Les mélanges A, B et C sont préliminaires et le dernier mélange sert à séparer la classe transition. Cette nouvelle classification a réduit le nombre de transitions invalides comme le montre le tableau 4.7.

Pour la classification des classes stationnaire et marche trois variables ont été fusionnées : l'énergie de l'axe-y du gyroscope à 1 Hz, l'énergie de l'axe-x de l'accéléromètre à 4 Hz et l'énergie de l'axe-y de l'accéléromètre à 4 Hz avec une même fenêtre de segmentation de trois secondes pour les trois variables et à l'aide d'un mélange de deux lois normales. Pour la variable énergie de l'axe-y du gyroscope les valeurs des seuils D et E déjà déterminées par l'analyse univariée ont été maintenues. La fusion a permis de déterminer les valeurs des seuils

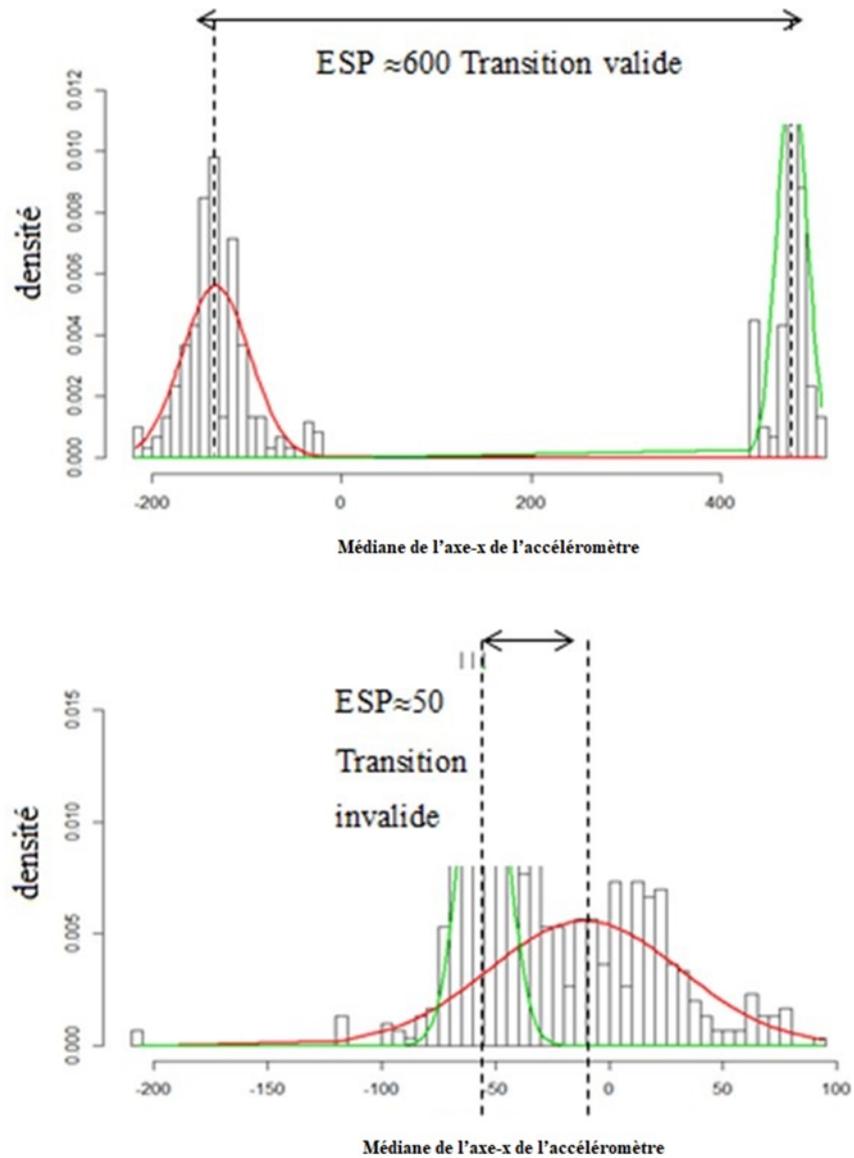


Fig 4.15: Validation de la transition debout/couchée par l'utilisation de l'espace entre un mélange de deux lois normales des données des 4 minutes qui précèdent et qui succèdent la transition.

Type de modèle	Seuil A	Nombre de transitions	Transition trouvée sans validation	Transition trouvée après validation	Précision avant validation	Précision après validation
Uni varié	332	36	56	36	0,70	1,00
Multi varié	376	36	44	36	0,85	1,00

Tableau 4.7: Résultats de la validation de la classification de la classe transition debout/couchée.

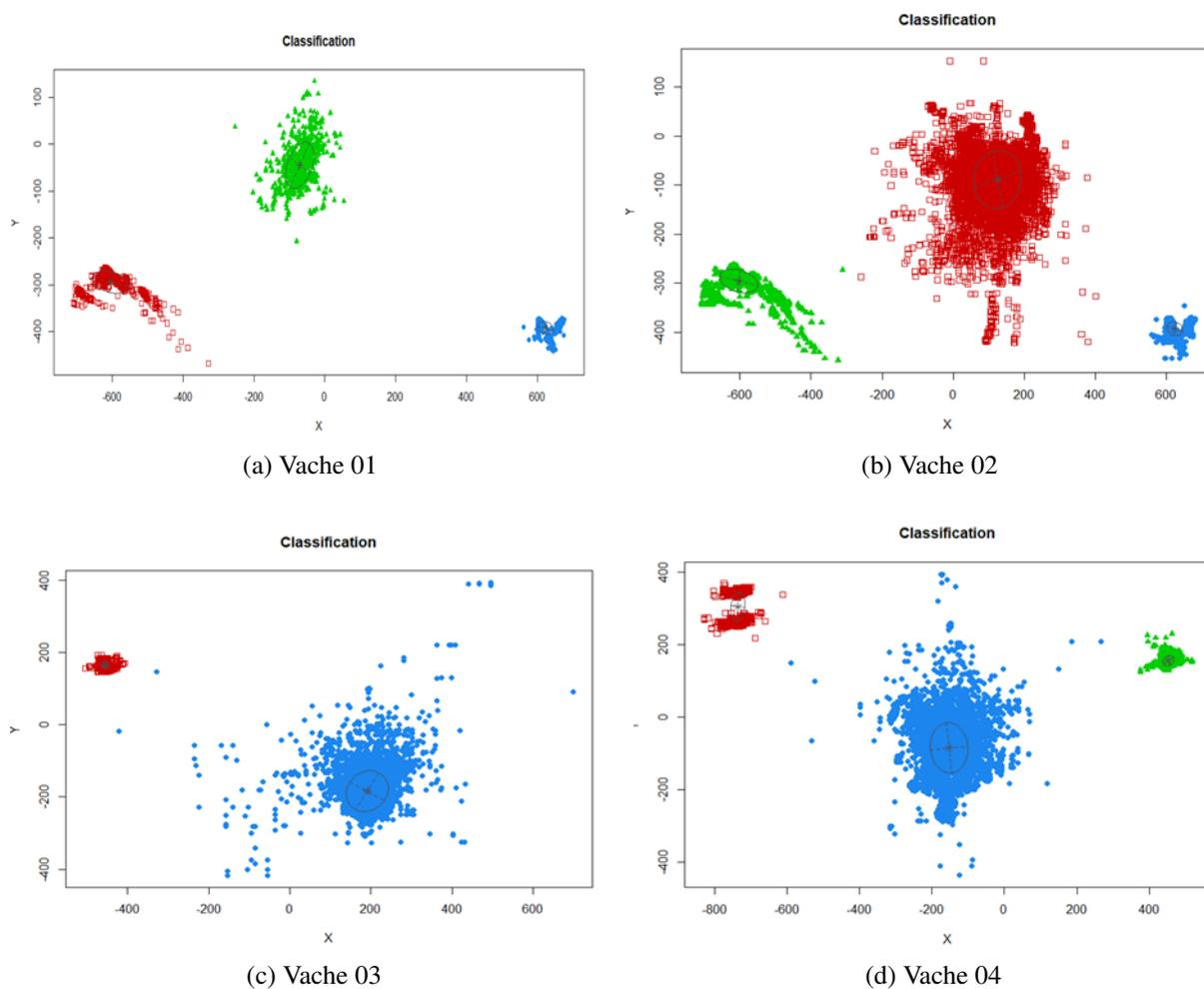


Fig 4.16: Classification multivariée des classes position debout, position couchée côté droit et position couchée côté gauche à l'aide des deux variables (médiane de l'axe-x et médiane de l'axe-y de l'accéléromètre).

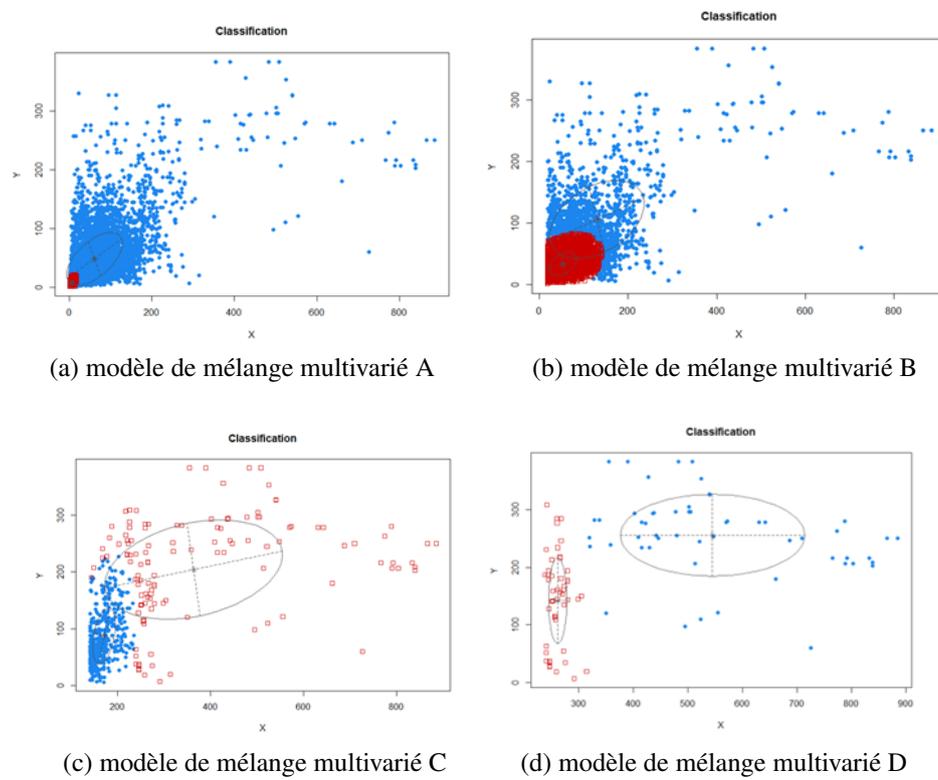


Fig 4.17: Les quatre mélanges multivariés consécutifs séparant la classe comportementale transition debout/couchée à l'aide de deux variables (intervalle interquartile de l'axe-x et de l'axe-y de l'accéléromètre).

Méthodes	Seuils moyens			Marche				Stationnaire			
	D / E	G	H	Se	Sp	Pr	Exct	Se	Sp	Pr	Exct
Méthode 1	D=47	-	-	0,98	0,98	0,84	0,98	0,97	0,97	0,99	0,98
Méthode 2	D=47	29,6	17,6	0,96	0,99	0,91	0,98	0,99	0,96	0,99	0,99
Méthode 3	E=142,4	-	-	0,93	0,99	0,93	0,99	0,99	0,95	0,99	0,99
Méthode 4	E=142,4	29,6	17,6	0,89	0,99	0,96	0,99	0,99	0,90	0,99	0,99

Tableau 4.8: Résultats de la classification de la classe marche et stationnaire avec 4 méthodes. Où : Méthode 1 : un mélange univarié de deux fonctions de densité de probabilité (seuil D) Méthode 2 : un mélange multi variée de trois variables optimisant la méthode 1 (seuil D, G, H) Méthode 3 : un mélange univarié consécutif de deux fonctions de densité de probabilité suivis de deux autres fonctions de densité de probabilité (seuil E) Méthode 4 : un mélange multi variée de trois variables optimisant la méthode 3 (seuil E, G, H).

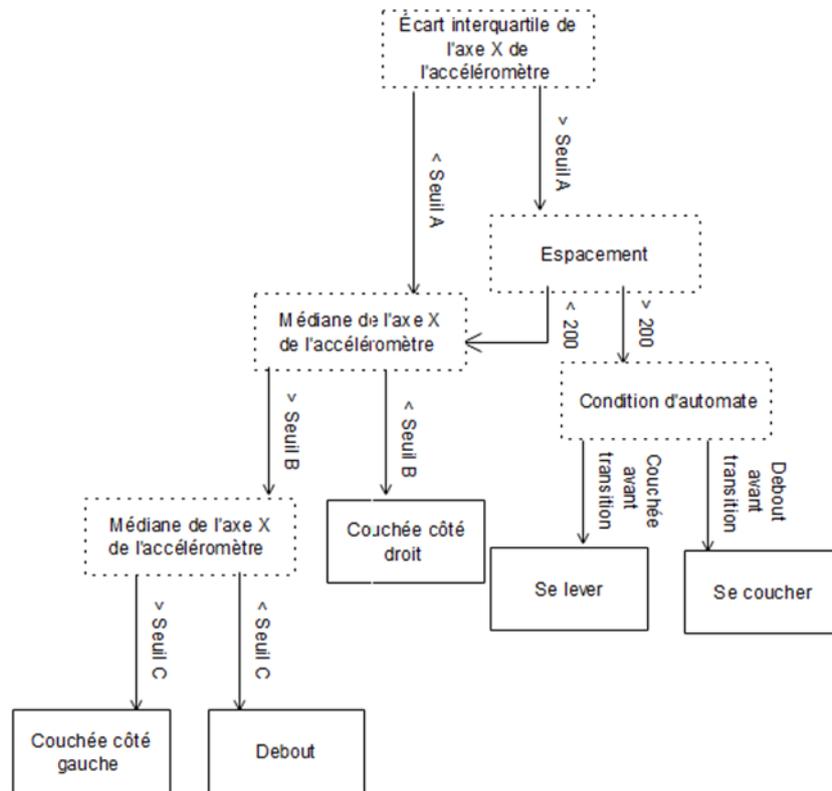
G et H. Le seuil G correspond au seuil de la variable énergie de l'axe-x de l'accéléromètre et le seuil H au seuil de la variable énergie de l'axe-y de l'accéléromètre. En utilisant ces différents seuils, quatre méthodes sont testées et les résultats varient considérablement entre ces méthodes comme le montre le tableau 4.8. La fusion a permis d'améliorer considérablement la précision de classification de la classe marche de 9 % (de 89 % jusqu'à 98 %), et la sensibilité de la classe stationnaire de 2% (de 97 % jusqu'à 99 %).

#### 4.3.4 Construction de l'arbre de décision

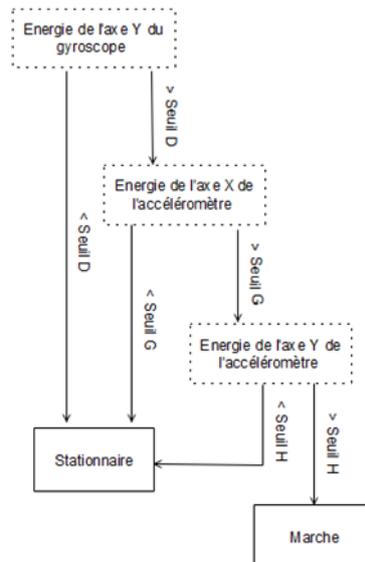
L'utilisation des variables et des seuils trouvés dans l'analyse par modèle de mélange univarié et multivarié et le respect des contraintes posées lors de la modélisation par automate ont permis de créer deux arbres de décision comme l'illustre la Fig 4.18. Le premier arbre de décision (Fig 4.18.a) sert à faire la classification des cinq classes (debout, couchée côté droit ou gauche, se coucher, se lever) et le deuxième (Fig 4.18.b) sert à séparer la classe debout en deux sous-classes qui sont le stationnaire et la marche.

## 4.4 Discussion

La classification comportementale des activités des vaches laitières à l'aide de capteur de mouvement a été considérée par plusieurs travaux [27, 28, 36, 38, 160, 161]. Ce qui distingue notre travail des travaux antérieurs se résume dans les points suivants : la multitude de races de vaches suivies, l'emplacement du capteur (dos de l'animal), la fréquence d'acquisition des données (80 Hz), le nombre de variables extraites et le nombre d'activités suivies par la méthode de classification. La variation de la fréquence des données collectées a été un point



(a) Architecture de l'arbre de décision permettant de faire la séparation des classes debout, couchée côté droit, couchée côté gauche et transition entre debout et couchée.



(b) Architecture de l'arbre de décision permettant de faire la séparation des classes marche et stationnaire incluses dans l'état debout.

Fig 4.18: Architecture des arbres de décision résultants.

culminant dans la méthodologie car elle permet de garder des taux de classification élevés tout en ayant la fréquence la plus minimale possible (jusqu'à 1 Hz). Le taux très élevé de collecte de données a permis d'avoir une grande échelle de données et le prétraitement a permis de faire des variations dans cette fréquence pour atteindre le niveau de classification le plus élevé possible avec la plus petite fréquence possible. L'avantage est que la diminution de ce taux de fréquence vers un niveau bas n'a engendré aucune perte de classification pour la plupart des comportements et que les petites fréquences permettent au capteur d'avoir une autonomie énergétique très élevée de sorte que sa durée de vie soit très grande et son déploiement soit plus bénéfique. Par exemple, avec une fréquence de 1 Hz le capteur s'allume juste une seule fois chaque seconde pour faire l'acquisition et le calcul de la variable souhaitée puis le capteur se met en mode économie d'énergie pendant le reste de la seconde ce qui augmente la durée de vie du composant électronique comparé à une fréquence de 80 Hz où le capteur est maintenu fonctionnel durant toute la seconde.

Contrairement à la plupart des méthodes de classification des comportements des vaches laitières, la classification par modèle de mélange univarié ou multivarié et l'utilisation des arbres de décision offre plusieurs avantages. Principalement, son fonctionnement automatique et non supervisé. Le modèle de mélange n'a besoin d'aucune phase d'apprentissage pour assurer son déroulement et son fonctionnement est moins complexe et plus proche des interprétations humaines. Le rôle du mélange des fonctions de densité est de trouver automatiquement le meilleur seuil qui permettra de faire la séparation entre les classes comportementales dans l'arbre de décision. Les arbres de décision sont facilement implementables dans les capteurs embarqués et sont peu gourmands en énergie par rapport aux autres classificateurs. Par conséquent, ils offrent l'avantage d'une grande autonomie de batterie et d'une grande durée de vie pour le capteur.

La position couchée est caractérisée par un changement considérable dans l'orientation du dos avec des taux très faibles d'énergie. En étant couchée, le dos de la vache s'incline vers deux orientations distinctes qui sont la droite ou la gauche, ce qui permet de différencier entre les états couchée vers le côté droit et couchée vers le côté gauche. La transition entre les positions debout et couchée est caractérisée par un changement remarquable dans l'orientation du dos également. L'activité de se coucher où de se lever implique une grande inclinaison du dos vers le bas ou vers le haut avec une activité d'énergie intense. Cependant, la position debout est caractérisée par une orientation inchangée du dos, mais la valeur énergétique diffère selon plusieurs sous-comportements inclus dans l'état debout : stationnaire et marche. L'état stationnaire se caractérise par plusieurs sous-mouvements dans lesquels la vache n'est pas en déplacement et elle utilise moins de deux pattes pour effectuer ses activités. La vache peut être immobile, en toilettage, tousse ou donne des coups de pied. Lors de l'immobilité, la vache

n'exerce aucune activité remarquable et le dos garde un état énergétique très faible. Pendant le toilettage la vache se lèche le corps en utilisant aucune ou une seule patte, l'état énergétique du dos est moyen pendant la plupart des évènements, mais peut atteindre des niveaux élevés. L'activité "toux" est rarement rencontré dans l'échelle de données et son état énergétique est élevé. L'état énergétique du dos de la vache faisant des coups de pied en utilisant une de ses pattes arrières est moyen chez la plupart des vaches. La plupart des activités incluses dans le stationnaire ont un impact énergétique moyen à faible sur le dos sauf dans des cas peu fréquents de toilettage et coup de pied qui montrent une énergie élevée. La marche se caractérise par d'autres sous comportements tels que le déplacement, le voyage, l'interaction sociale et la rotation. Tous ces comportements se caractérisent par un état énergétique élevé à très élevé, seuls quelques cas isolés de déplacement ont manifesté une énergie moyenne.

Le travail en cours présente des taux de classification optimaux pour quelques classes liées à la posture de l'animal (debout, couchée, transition) et très élevés pour d'autres classes liées au mouvement (stationnaire, marche). Ces taux de classification sont considérablement plus élevés que ceux rapportés par d'autres travaux en termes de sensibilité, spécificité, précision et exactitude comme le montre les tableaux 4.9, 4.10 et 4.11. Ce qui confirme l'hypothèse que le dos de l'animal peut être utilisé comme emplacement efficace pour faire une classification des activités de vaches laitières.

La classification des comportements debout, couchée et transition debout/couchée s'avère optimale (100 %) et peut être utilisée comme méthode de référence en remplaçant l'annotation manuelle faite après observation des enregistrements vidéo utilisée actuellement dans la recherche. En effet, les vaches peuvent se trouver hors de la portée des caméras utilisées pour l'annotation engendrant des pertes dans l'échelle de données à traiter.

La détermination de la direction vers laquelle la vache s'est couchée (côté droit ou gauche) a été effectuée avec succès. Notons qu'aucun autre travail scientifique n'a réalisé cette classification, et pourtant cette séparation ramène une nouvelle valeur informative et peut aider à mieux comprendre l'état comportemental de la vache.

L'utilisation de la fusion de données permet d'optimiser les taux de classification et cela en réduisant le nombre de faux positifs. La fusion des données de la médiane de l'axe-x et la médiane de l'axe-y de l'accéléromètre avec une segmentation de 5 secondes et une fréquence de 1 Hz permet d'atteindre un taux optimal de 100 % comparé à celui trouvé à l'aide d'une seule variable (qui reste élevé, 99 %). La fusion de données de l'intervalle interquartile de l'axe-x et de l'axe-y de l'accéléromètre à 1 Hz avec une segmentation de 10 secondes permet d'améliorer les résultats apportés par le modèle univarié (précision= 70 %) jusqu'à atteindre un niveau de précision de 85 %, les deux modèles (univarié et multivarié) ont une sensibilité de 100 %. La fusion de données de l'énergie de l'axe-y du gyroscope à 1 Hz avec l'énergie

de l'axe-x et l'énergie de l'axe-y de l'accéléromètre à 4 Hz a permis de réduire considérablement les taux de faux positifs et d'améliorer la précision et la sensibilité. Cependant, cette fusion a aussi causé une perte dans les taux de vrais positifs qui reflètent la sensibilité du système à détecter l'activité marche. En utilisant deux mélanges consécutifs afin de trouver le seuil de la variable énergie de l'axe-y du gyroscope et en ajoutant les nouveaux seuils liés à l'accéléromètre trouvé par la méthode de fusion de données, la précision a augmenté de 84 % à 91 %, la spécificité a augmenté de 98 % à 99 %, la sensibilité a diminué de 98 % à 96 % et l'exactitude garde un même taux de 98 %. Suite à l'utilisation de trois mélanges consécutifs pour trouver la valeur du seuil de la variable de l'énergie de l'axe-y du gyroscope et la fusion avec les deux variables provenant de l'accéléromètre, la précision a augmenté de 93 % jusqu'à atteindre 96 %, la sensibilité a diminué de 93 % jusqu'à 89 %, la spécificité et l'exactitude ont gardé un même taux de 98 %.

L'utilisation de l'automate d'états finis déterministes a permis de mieux comprendre les relations qui existent entre les comportements et de poser plusieurs contraintes sur l'ordre chronologique des comportements. Cet ordre temporel permet de trouver le type de transition et de différencier entre l'activité se lever et se coucher. L'autre utilisation de ces contraintes est lors de la validation des transitions de sorte qu'une transition valide ne peut se trouver entre deux états similaires. Cet ordre de validation a conduit au calcul d'une variable d'espacement entre la population de données qui précède et qui succède la période de transition et a permis d'optimiser les résultats de classification pour atteindre un taux de 100 %.

L'utilisation de deux mélanges consécutifs ou de 3 mélanges consécutifs de lois normales pour la classification des classes marche et stationnaire dans le modèle univarié à l'aide de la variable énergie de l'axe-y du gyroscope a apporté des résultats de classification différents. L'utilisation de deux mélanges consécutifs de deux lois normales a donné une sensibilité, une spécificité et une exactitude élevées de l'ordre de 98 %, quant à la précision, elle était moyenne (84 %). Cette classification permet de détecter la plupart des états marche qui existent dans l'échelle de données. Cependant, vu que la valeur du seuil donnée par ce mélange est petite, plusieurs activités incluses dans le stationnaire peuvent franchir ce seuil causant un nombre remarquable de faux positifs ce qui explique cette perte de précision. En outre, l'utilisation de trois mélanges consécutifs de deux lois normales a optimisé la précision jusqu'à atteindre un niveau de 93 % et elle a aussi optimisé la spécificité et l'exactitude (99 %) sauf qu'une perte considérable de sensibilité a été remarquée de sorte que le nombre de vrais positifs a diminué, la sensibilité a diminué de 98 % à 93 %. Cette diminution est causée par quelques états de marche lente où l'animal se déplace doucement pendant une courte période et son énergie avait une valeur inférieure au seuil retourné par ce mélange.

Pour la classe comportementale couchée, des taux de classifications très élevés ont été

obtenus quand des accéléromètres sont attachés aux pattes des vaches (Se = 100 %, Pr = 99 %, exact = 99 %) [28, 156] et sont comparables à ceux obtenus dans le présent travail comme le montre le tableau 4.9. Cependant les taux de classification se détériorent quand des capteurs sont attachés aux cous [27, 29, 171], la tête et l'oreille [201]. Ces taux sont d'une sensibilité variant entre 77 % et 99 %, spécificité de 92 % jusqu'à 98 %, précision de 83 % jusqu'à 98 % et d'exactitude de 84 % jusqu'à 99 % selon les différents travaux (tableau 4.9).

Pour la classe comportementale debout, ce travail a montré un taux de classification de 100 % qui est considérablement plus grand comparé aux résultats obtenus si le capteur est attachés aux pattes (Se = 76 %, Pr = 68 %, Exact = 98 %) et reste plus élevé comparé au capteur attaché au cou, tête et oreille de l'animal (Se = 68 % jusqu'à 88 %, Sp = 84 % jusqu'à 94 %, Pr = 55 % jusqu'à 81 % et Exact = 84 % jusqu'à 98 %). Tous ces résultats sont représentés dans le tableau 4.9.

Pour la classification comportementale de la classe marche un taux élevé a été obtenu. Ce taux est supérieur à ceux obtenus par d'autres travaux [27, 28, 171, 172, 201] comme le montre le tableau 4.10. Lorsque le capteur est placé sur la tête, le cou, l'oreille ou le pied, les taux de classification obtenus sont les suivants: la sensibilité varie entre 69 % et 80 %, la spécificité varie entre 73 % et 79 %, la précision varie entre 77 % et 79 % et l'exactitude varie entre 18 % et 99 %. Mais bien que la vache laitière soit immobile, sa tête, son cou ou son oreille peuvent bouger (se retourner, se lécher, se frotter) et elle peut donner des coups de pied, ce qui explique ces taux obtenus. D'après nos résultats, le dos des vaches est un endroit plus stable et mieux adapté à cette discrimination.

Pour la classification comportementale de la classe transition debout/couchée une exactitude de 100% a été obtenue. Cette exactitude est plus élevée par rapport à celle obtenue par les autres travaux [197, 199] comme le montre le Tableau 4.11. Les taux de classification des transitions (debout/couchée) donnés dans [197, 199] en attachant un accéléromètre 3D au cou de l'animal sont inférieurs à ceux obtenus en plaçant l'IMU au dos de l'animal. Cela peut s'expliquer par le fait que le cou est un endroit affecté par plusieurs activités telles que l'alimentation, le toilettage, etc. Cependant, le dos n'est pas affecté par ces activités.

Dans notre travail, nous pouvons distinguer deux étapes principales : le développement du modèle de classification (arbres de décision) et sa mise en œuvre. Pour développer ce modèle, nous avons utilisé la carte arduino MEGA ADK rev3. Ce choix est motivé par le fait que le microcontrôleur de cette carte (ATmega2560) dispose de suffisamment de mémoire pour enregistrer et exécuter le programme d'acquisition de données par rapport aux autres microcontrôleurs tels que ATtiny85, ATmega168, ATmega328. Le programme d'acquisition nécessite un espace mémoire de 32 432 octets pour le stockage et 1401 octets pour son exécution.

Les résultats de cette étape montrent que seules trois variables (les axes x et y de l'accéléro-

Réf	Capteur	Méthode	Emp*	Couchée				Debout			
				Se	Sp	Pr	Exct	Se	Sp	Pr	Exct
<b>Notre</b>	Acc 3D	Arbre de decision	Dos	100	100	100	100	100	100	100	100
[201]	Acc 3D	Forêts aléatoires	Oreille	88	92	-	91	85	84	-	84
			En dessus de la tête	99	98	-	99	91	94	-	88
			Cou	94	98	-	97	92	87	-	90
[171]	Positionnement	SVM	Cou	-	-	-	86	-	-	-	89
[156]	Acc 3D	k plus proches voisins	Patte	100	-	99	-	76	-	68	-
			cou	96	-	83	-	68	-	81	-
[27]	Acc 3D	SVM	cou	80	-	83	84	80	-	65	87
[29]	Acc 3D	Arbre de decision	cou	77	-	98	-	88	-	55	-
[28]	Acc 3D	Arbre de decision	pattes	-	-	-	99	-	-	-	98

Emp : Emplacement du capteur

Tableau 4.9: Comparaison entre le présent travail et d'autres publications pour les classes comportementales debout et couchée.

Réf	Capteur	Méthode	Emp	Se	Sp	Pr	Exct
<b>Notre</b>	Acc Gyr 3D	Arbre de decision	Dos	96	99	91	98
[172]	Acc Mag 3D GPS	SVM; LR; RFE; KNN	Cou	69	-	77	-
[28]	Acc 3D	Arbre de decision	Pattes	-	-	-	67
[171]	Positionnement intérieur	SVM	cou	-	-	-	18
[201]	Acc 3D	Forêts aléatoires	Oreille	77	73	-	73
			tête	80	78	-	78
			Cou	79	79	-	79
[27]	Acc 3D	SVM	Cou	79	-	79	99

Tableau 4.10: Comparaison entre le présent travail et autres publications pour la classe comportementale marche.

mètre et l'axe des y du gyroscope) sont utilisées par ces arbres de décision à deux fréquences (1 Hz et 4 Hz). Comme l'espace mémoire nécessaire au stockage du programme permettant l'acquisition de ces trois variables est réduit à seulement 24 032 octets et à 1361 octets pour son exécution et que la carte arduino Mini Pro consomme nettement moins d'énergie que la carte arduino MEGA ADK. Nous avons remplacé cette dernière carte par la carte arduino Mini Pro. En utilisant une batterie de 2600 mA/h avec un facteur de décharge de 15 %, la durée de vie de notre système est estimée à environ trois ans lors de l'utilisation d'une fréquence de 1 Hz et un an lors de l'utilisation d'une fréquence de 4 Hz.

Réf	capteur	Méthode	Emp	Se	Sp	Pr	Exct
<b>Notre</b>	Acc 3D	Decision tree	Dos	100	100	100	100
[29]	Acc 3D	Decision tree	Cou	96	-	87	-
[27]	Acc 3D	SVM	Cou	71	-	29	100

Tableau 4.11: Comparaison entre le présent travail et autres publications pour la classe comportementale transition entre la position debout et couchée.

## 4.5 Conclusion

Le présent travail montre clairement que le dos de l'animal est un endroit efficace pour assurer le suivi des activités comportementales des vaches laitières. En effet, des taux de classification optimaux pour les classes : couchée (côté droit ou gauche), debout et transitions entre ces positions et des taux de classification élevés pour les classes marche et stationnaire sont obtenus.

Les modèles de mélanges finis présentent plusieurs avantages. Principalement, ils permettent de tenir compte des différences entre les animaux et les dispositifs individuels ce qui permet d'estimer des valeurs seuils pour chaque animal et dans des conditions environnementales inconnues et variables. De plus, ils offrent une efficacité de calcul adaptée à une grande quantité de données.

Cette approche peut être utilisée comme méthode de référence pour la classification comportementale. Elle réduit considérablement le temps et les coûts associés à la collecte de données par caméra utilisée pour réaliser l'annotation manuelle. Cependant, des recherches supplémentaires sur un grand nombre d'animaux et dans différentes zones géographiques et différents types de bâtiments d'élevage sont nécessaires.

Le positionnement de l'IMU sur le dos des vaches pourra remplacer les podomètres ou accéléromètres attachés aux pieds des animaux tout en ajoutant une meilleure valeur informative et cela en intégrant au microcontrôleur d'autres capteurs (RFID et GPS pour la localisation, détecteur de contraction musculaire pour la détection de vèlage, détecteur de pression pour les événements de chaleur). Cependant, la limite de cet emplacement se trouve dans sa capacité restreinte à assurer le suivi du comportement alimentaire, de l'abreuvement et de la rumination. Mais cette limite peut être surpassée par : 1. L'intégration d'un système de localisation permettant d'estimer le nombre et le temps des périodes d'alimentation et d'abreuvement et 2. L'ajout d'un deuxième accéléromètre qui sera cette fois attaché au cou de l'animal. Les deux capteurs fonctionneront simultanément pour assurer un suivi global et des taux optimisés de la classification. Comme on peut utiliser un système de reconnaissance de vache basé sur l'analyse d'images et le déployer pour le calcul du temps d'alimentation et d'abreuvement. Justement dans le prochain chapitre, nous présentons notre deuxième contribution qui se fo-

---

calise sur le monitoring du comportement alimentaire de vaches en utilisant une approche d'analyse d'images.



# Chapitre 5

## Analyse d'images pour l'identification individuelle et le suivi du comportement alimentaire de vaches laitières

### 5.1 Introduction

Le monitoring du comportement alimentaire et l'identification individuelle de vaches laitières ont un grand impact sur l'optimisation de la productivité et l'amélioration du suivi sanitaire. Le comportement alimentaire est considéré avec une grande attention en élevage de précision et en alimentation de précision. Dans l'alimentation de précision, il est utilisé pour analyser l'impact de la qualité et de la quantité de nourriture fournie sur l'ingestion, les performances de croissance et la productivité des animaux [202–204]. Il évalue également le rapport entre l'énergie consommée et l'énergie utilisée pour la croissance et la production [205, 206]. Dans l'élevage de précision, les changements dans le comportement alimentaire sont fréquemment utilisés pour prévoir les maladies [207]. En effet, la réduction du comportement alimentaire permet à l'organisme de l'animal d'améliorer sa résistance à la maladie et facilite le rétablissement [104]. De plus, cette réduction est aussi interprétée comme un moyen efficace pour réduire l'ingestion de certains micronutriments nécessaires à la multiplication des agents pathogènes [105]. En outre, le monitoring de l'activité alimentaire permet la détection précoce des carences alimentaires et aide à ajuster les rations [208]. Pour assurer un monitoring alimentaire, plusieurs paramètres tels que le temps passé à se nourrir, la quantité de nourriture ingérée, la rumination et la vitesse à laquelle la nourriture est ingérée, etc. peuvent être mesurés.

Le comportement alimentaire varie entre les animaux au sein d'un même troupeau et varie

au cours du temps pour un même animal. Néanmoins, plusieurs études [95, 104, 109, 111] ont prouvé l'impact des maladies bovines sur la modification du comportement alimentaire, de l'activité ou la production des vaches laitières. Dans [104, 111], des études ont été menées pour examiner le comportement des vaches laitières infectées par des mammites durant toute la période de leur maladie. Les résultats ont montré une diminution d'apport alimentaire de 2,5 kg/j [111] et une diminution du temps d'alimentation et de la quantité d'aliments ingérée [104]. Dans [95], les auteurs se sont intéressés à l'identification des changements de comportements alimentaires associés à l'apparition de troubles de santé (cétose, boiterie). Les résultats obtenus ont montré une diminution de l'apport alimentaire, du temps d'alimentation, du taux d'alimentation et du nombre de visites à la mangeoire. Dans [109], le but était de quantifier la diminution de l'apport alimentaire et les pertes de production de lait associées aux troubles de santé chez les vaches laitières. Les auteurs ont noté une diminution maximale de l'apport alimentaire à l'apparition de la cétose et de la mammite. Dans [110], les auteurs ont suivi l'alimentation individuelle des vaches laitières pour quantifier l'effet de la boiterie sur l'alimentation et la rumination. Les résultats ont montré qu'en général la variation intra individus était faible; soit 5 à 15 %, alors que la variation inter individus était beaucoup plus élevée, pouvant atteindre les 50 %. Les vaches boiteuses différaient des vaches non boiteuses dans trois aspects de leur comportement alimentaire : le temps, la fréquence et le taux d'alimentation. La boiterie réduit le temps quotidien d'alimentation de 46 % et la fréquence d'alimentation de 44 %.

Pour surveiller le comportement alimentaire, plusieurs approches et outils sont utilisés. La puce RFID [22], le système de localisation en temps réel (RTLS) [209] et le système à bande ultra large (UWB) [173, 210] sont utilisés pour localiser les vaches et estimer le temps passé dans la zone d'alimentation. Bien que ces systèmes aient la capacité d'estimer le temps passé dans la zone d'alimentation, ils ne peuvent pas préciser si l'animal mange réellement. Des capteurs de pression (RumiWatch Halter, ITIN + HOCH GmbH, Liestal, Suisse) sont attachés aux vaches laitières pour surveiller les mouvements de mastication [211, 212]. Ces capteurs ont donné de bons résultats en matière de surveillance. Mais, ils ne sont utilisés que dans la recherche scientifique en raison de leur coût très élevé. Un microphone détectant le son produit par les vaches laitières lorsqu'elles mangent est utilisé pour identifier l'activité d'alimentation [213, 214]. Un accéléromètre est utilisé pour la classification en temps réel de l'activité et du comportement alimentaire [156, 161]. Cependant, ces capteurs attachés présentent des limites telles que leur fragilité (arrachés lors du grattage), la gestion de leur énergie limitée et ils sont stressants pour les vaches. De ce fait, l'utilisation de caméras pour l'identification et le suivi du comportement alimentaire de vaches permet de surmonter ces limites. En effet, les caméras sont faciles à déployer, non stressantes et peu coûteuses.

L'objectif de ce présent travail est de développer un système d'analyse d'images qui permet d'identifier et de suivre le comportement alimentaire de vaches dans une ferme commerciale. Un capteur embarqué est développé et déployé pour l'acquisition et la sélection des images de vaches en période d'alimentation. Des modèles d'apprentissage ont été entraînés pour détecter la présence d'une vache dans la zone d'alimentation, déterminer le temps réel qu'elle passe à s'alimenter, identifier individuellement chaque vache, contrôler la disponibilité d'aliment dans la mangeoire et identifier le type de cet aliment.

Le système d'analyse d'images conçu permet : (1) En utilisant des algorithmes d'analyse d'images, les vaches laitières se trouvant dans la zone d'alimentation sont classées comme étant debout ou en train de se nourrir, (2) L'état d'alimentation n'est noté que si la vache laitière est présente dans la zone d'alimentation, sa tête est inclinée vers la mangeoire et que la nourriture est disponible dans la mangeoire, (3) L'utilisation de la partie supérieure de la tête comme région d'intérêt pour l'identification individuelle de la vache laitière, (4) L'utilisation d'une approche de segmentation sémantique pour la segmentation des images de la vache avant son identification, (5) Le prétraitement est intégré dans le capteur et fait en temps réel et (6) La surveillance de nouveaux paramètres de l'alimentation tels que la disponibilité et le type de la nourriture.

## 5.2 Matériel et méthode

La procédure du suivi du comportement alimentaire et d'identification de vaches est présentée dans la Fig 5.1. Tout d'abord, l'acquisition et la sélection en temps réel des images de vaches laitières sont effectuées par un système embarqué. En effet, ce système permet, en premier lieu, une acquisition et une sélection d'images comme le montre la Fig 5.1.a. Ensuite, la classification des images est réalisée en se basant sur un enchaînement de modèles d'apprentissage profond (CNNs et Xception) comme le montre la Fig 5.1.b. Les images sont classées par un premier réseau de neurones à convolution (CNN1) en deux classes (vache-présente et vache-absente dans la zone d'alimentation). Puis, les images de la classe vache-présente sont classées par un deuxième CNN (CNN2) en deux autres classes (debout et manger). La classe manger désigne que la vache est en période d'alimentation et la classe debout désigne que la vache n'est pas en train de manger. Les images de la classe vache-absente sont classées à l'aide d'un troisième CNN (CNN3) afin d'identifier le type et la disponibilité des d'aliments. Les images de vaches en période d'alimentation seront ensuite segmentées à l'aide d'une méthode d'apprentissage par transfert (Xception) pour supprimer l'arrière-plan. Trois approches de classification ont été développées pour identifier chaque vache. Ces approches se basent sur un réseau de neurones à convolution (CNN4), une machine à vecteurs de support (SVM)

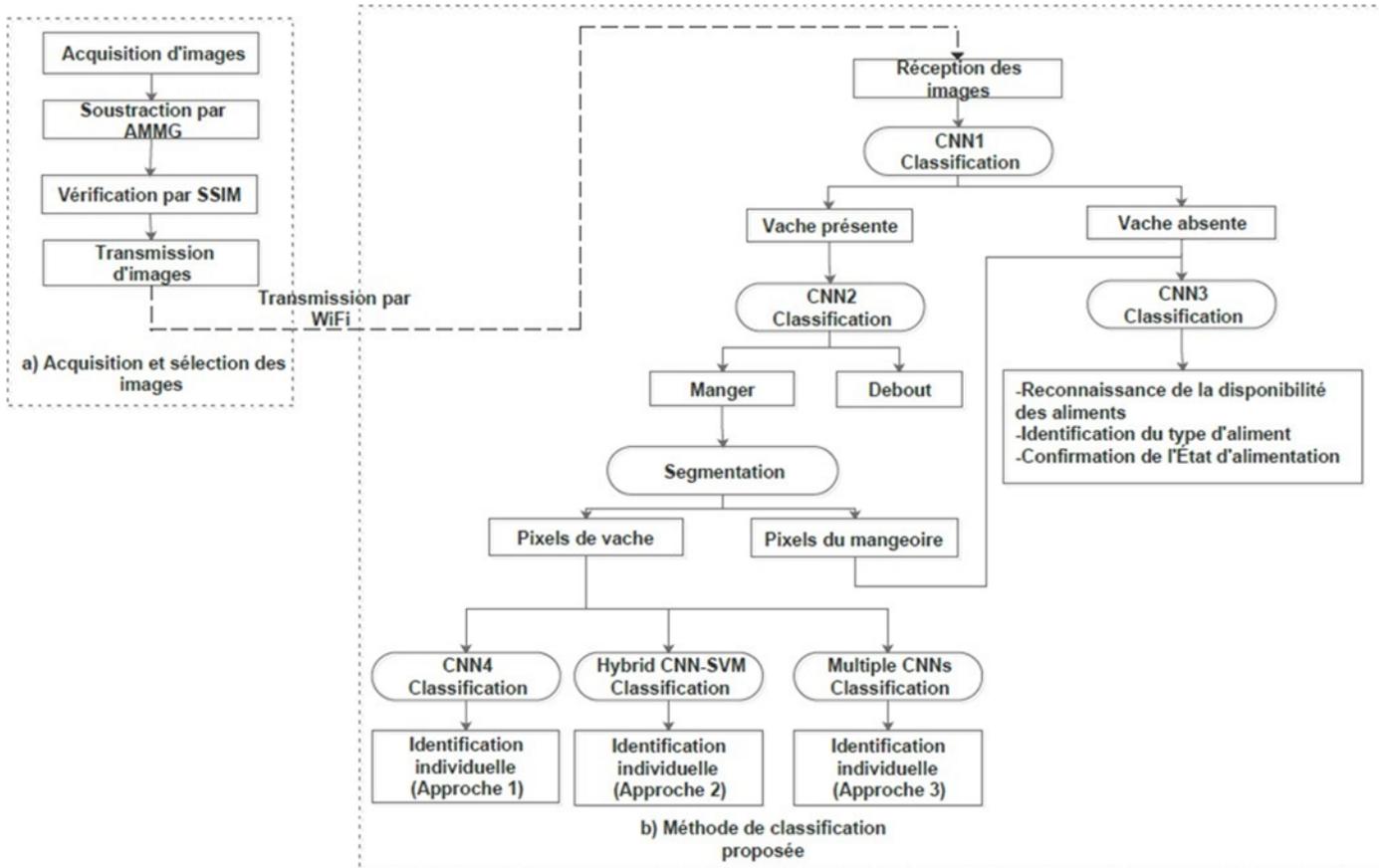


Fig 5.1: Procédure d'acquisition, de sélection et de classification d'images.

et une méthode de regroupement par une multitude de CNN.

### 5.2.1 Acquisition et sélection d'images

L'équipement d'acquisition illustré dans la Fig 5.2.a, se compose d'un module Raspberry Pi 3 Model B et d'une caméra USB (Yudanny USB HD Webcam) déployée pour collecter des images en temps réel. La caméra a été fixée au-dessus d'une entrée de la zone d'alimentation qui est délimitée par deux barrières d'une hauteur de 150 cm et de largeur de 30 cm et la zone de liberté utilisée entre deux barrières est de 80 cm comme l'illustre la Fig 5.2.b. Ce choix est fait pour assurer une répartition équitable des zones de liberté entre les vaches. L'éloignement de la caméra par rapport au sol était de 220 cm. La résolution de l'image était de 480 pixels (vertical) × 640 pixels (horizontal). La fréquence d'acquisition d'images est de 1 image par seconde. Le système est programmé à transmettre en Wi-Fi les images sélectionnées vers une borne centralisée, constituée d'un ordinateur (HP pavillon dv6).

L'ajout des barrières a été motivé par les travaux [177, 215, 216] qui ont montré que

les interactions sociales entre les animaux impactent négativement l'identification individuelle des animaux. Les expériences sur terrain se sont déroulées en décembre 2018 dans une ferme commerciale à stabulation libre pour vaches laitières située dans la région de Draâ Ben Khedda Wilaya de Tizi Ouzou. Dix-sept vaches de race HOLSTEIN ont été sélectionnées pour l'expérience. Le système d'acquisition exécute un programme permettant une sélection en temps réel d'images de vaches. Une méthode de sélection d'images qui se base sur la segmentation d'arrière-plan et avant-plan et le calcul d'un index de similitude est utilisée pour déterminer la dissimilitude entre les images acquises. Ensuite, les images dissimilaires sont transmises vers la borne qui se charge de classer le comportement alimentaire et d'identifier les vaches.

**La sélection d'images** l'algorithme permettant de sélectionner les images est donné par la Fig 5.3. Il permet de sélectionner toutes les images dissimilaires en temps réel. En effet, dans les images acquises trois états sont envisageables : (1) la vache n'est pas présente dans la zone d'alimentation, (2) la vache est présente et elle est statique, (3) la vache est présente et elle est en mouvement. Le système embarqué doit permettre de transmettre toutes les images dissimilaires de ces trois cas. Les images dissimilaires de la zone d'alimentation reflètent un changement dans le type ou la quantité d'aliment présent dans la mangeoire. Les images dissimilaires de la vache en mouvement reflètent un déplacement de la tête de la vache lors d'une période d'alimentation ou sa transition de l'état mangé vers l'état debout ou vice versa. L'image de vache statique est reflétée par la première et la dernière image acquise lors d'un mouvement de vache. Le système permet aussi la mise à jour d'un compteur qui donne le temps pendant lequel la vache était dans l'état reflété par l'image.

Nous avons utilisé l'algorithme d'extraction d'arrière-plans par Modèle de Mélanges de Gaussiennes Adaptatif (AMMG) [217]. Cet algorithme adapte automatiquement le nombre de lois normales utilisées pour modéliser un pixel donné. Son fonctionnement est automatique et il permet de réduire les besoins en mémoire. En utilisant l'algorithme AMMG, le contour de zone de variation de pixels entre deux images consécutives est sélectionné et la surface de ce contour est calculée. Pour implémenter cet algorithme, la librairie opencv a été utilisée. Le principe de fonctionnement de AMMG est que pour un nombre donné de B gaussiens, la probabilité qu'un nouvel échantillon de données  $x_t$  soit un pixel de l'arrière-plan est représentée comme suit :

$$p(x_t) = \sum_{i=1}^B \omega_{i,t} \cdot \eta(x_t, \mu_{i,t}, \sigma_{i,t}^2) \quad (5.1)$$

où  $\eta(x_t, \mu_{i,t}, \sigma_{i,t}^2)$  est la  $i$ ème fonction de densité de probabilité (gaussienne) avec son poids



(a) Equipement d'acquisition



(b) Zone d'alimentation

Fig 5.2: Zone et équipement d'acquisition d'images.

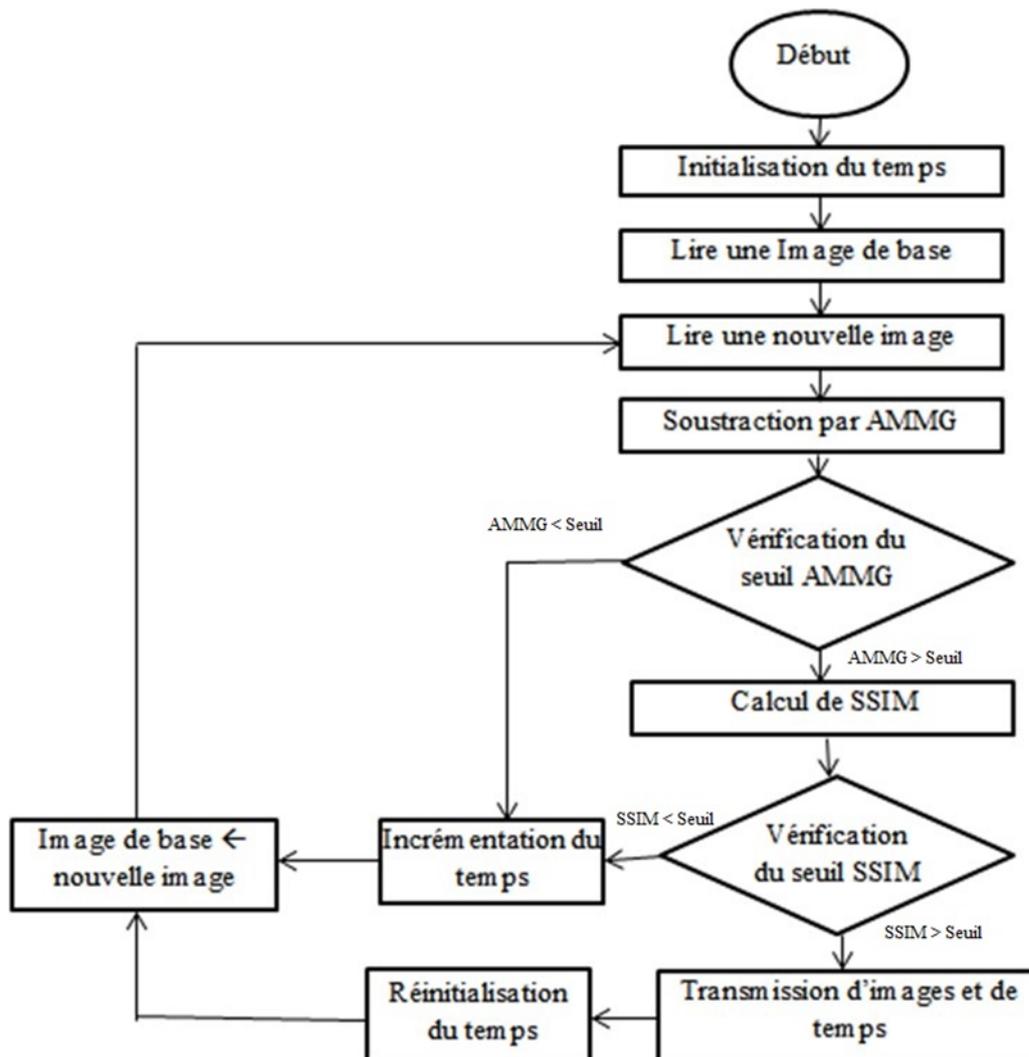


Fig 5.3: Organigramme de sélection d'images.

$\omega_{i,t}$  au temps  $t$ ,  $\sigma_{i,t}^2$  et  $\mu_{i,t}$  sont respectivement la valeur de variance et la valeur moyenne du  $i$ ème gaussien dans le mélange au temps  $t$ .

La matrice de covariance  $\sigma_{i,t}^2 I$  est supposée être diagonale et la matrice d'identité  $I$  a sa propre dimension. L'AGMM met à jour récursivement les paramètres du modèle de mélange  $\omega_{i,t}$ ,  $\mu_{i,t}$  et  $\sigma_{i,t}^2$  comme suit :

$$\omega_{i,t} = \omega_{i,t} + \alpha(o_{i,y} - \omega_{i,t}) \quad (5.2)$$

$$\mu_{i,t} = \mu_{i,t} + o_{i,t}(\alpha/\omega_{i,t})\delta_{i,t} \quad (5.3)$$

$$\sigma_{i,t}^2 = \sigma_{i,t}^2 + o_{i,t}(\alpha/\omega_{i,t})(\delta_{i,t}^T \delta_{i,t} - \sigma_{i,t}^2) \quad (5.4)$$

où  $\delta_{i,t} = x_t - \mu_{i,t}$ ,  $\alpha = 1/P$ ,  $P$  est un délai d'adaptation raisonnable. La propriété  $o_{i,t}$  est fixée à 1 pour la composante proche ayant la plus grande  $\omega_{i,t}$  et l'autre  $o_{i,t}$  est fixée à 0.

Si les composantes sont triées pour avoir des poids décroissants  $\omega_{i,t}$ , la valeur  $B$  est mise à jour comme suit :

$$B = \operatorname{argmin}_b \left( \sum_{i=1}^b \omega_i > (1 - cf) \right) \quad (5.5)$$

où  $cf$  est une mesure de la portion maximale des données qui peuvent appartenir aux objets du premier plan sans influencer l'arrière-plan [218].

L'extraction de l'arrière-plan permet de détecter tous les changements significatifs dans les images qui reflètent les états possibles. Dans le même sens et afin d'améliorer la sélection en temps réel des images, l'index de similitude structurelle a été calculé. Cet indice permet de confirmer la variation déjà estimée entre deux images. L'indice de similarité structurelle (SSIM) permet de mesurer la qualité d'une image en la comparant à une image de référence. Il est aussi utilisé pour mesurer la performance de la restauration d'une image. Les auteurs de [219] ont développé un indice de similarité structurelle (SSIM) qui calcule la différence entre une image déformée et une image de référence en utilisant diverses propriétés connues du système visuel humain. Pour réaliser ce calcul la librairie python (scikit-image) [220] a été utilisée. Cet indice de similarité est estimé selon l'équation (5.6).

$$SSIM(x,y) = [l(x,y)]^\alpha \cdot [c(x,y)]^\beta \cdot [s(x,y)]^\gamma \quad (5.6)$$

où  $l(x,y)$  représente la fonction d'adaptation à la luminosité,  $c(x,y)$  la fonction de comparaison de contraste et  $s(x,y)$  la fonction de comparaison de structure. Les paramètres positifs  $\alpha$ ,

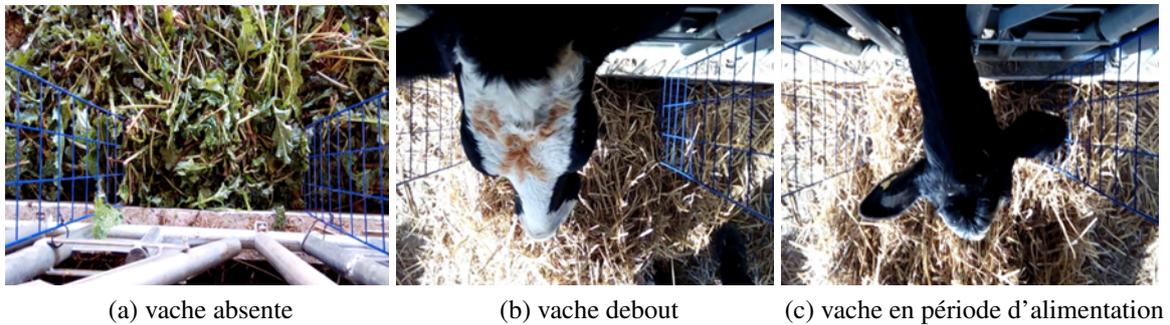


Fig 5.4: Images montrant les trois classes suivies.

$\beta$  et  $\gamma$  sont réglés pour équilibrer la contribution de chaque terme.

### 5.2.2 Choix des classes comportementales à suivre

L'étape d'acquisition de données a fourni une base de données composée de trois classes d'images dans lesquelles la vache est soit absente, soit présente et elle est en position debout ou présente et elle est en train de manger. Durant l'absence de la vache (Fig 5.4.a), le système a acquis des images de la zone d'alimentation car une variation considérable entre les images a été détectée. Cette variation est due soit à la distribution d'aliments par le personnel de la ferme ou à l'entrée/sortie des vaches à la zone d'alimentation. Lorsque la vache est présente dans la zone d'alimentation, sa tête peut se trouver dans deux différents états : debout ou incliné vers le bas. L'inclinaison de la tête vers le bas indique que la vache est probablement en période d'alimentation (Fig 5.4-c). Le soulèvement de la tête indique que la vache est debout et elle n'est pas en période d'alimentation (Fig 5.4-b).

Pour classer une vache en période d'alimentation trois critères sont définis. Ces critères sont la présence de la vache, la disponibilité d'aliments dans la mangeoire et l'inclinaison de sa tête vers le bas. Le comportement d'une vache en période d'alimentation est modélisé par un automate d'états finis déterministes (Fig 5.5). Cette période commence par un état initial qui désigne la présence de la vache. Ensuite, la vache peut être en deux états (debout, manger) et elle peut aussi transiter entre ces deux états. Au final, la sortie de la vache montre la fin de la période d'alimentation.

### 5.2.3 Partitionnement de la base de données

L'acquisition des données (images) a été effectuée pendant 5 jours (J1, J2, J3, J4 et J5). Les données de J1 et J2 sont utilisées comme données d'entraînement. Les données des jours J3, J4 et J5 sont utilisées comme données de test. Les deux jours d'entraînement (J1 et J2)

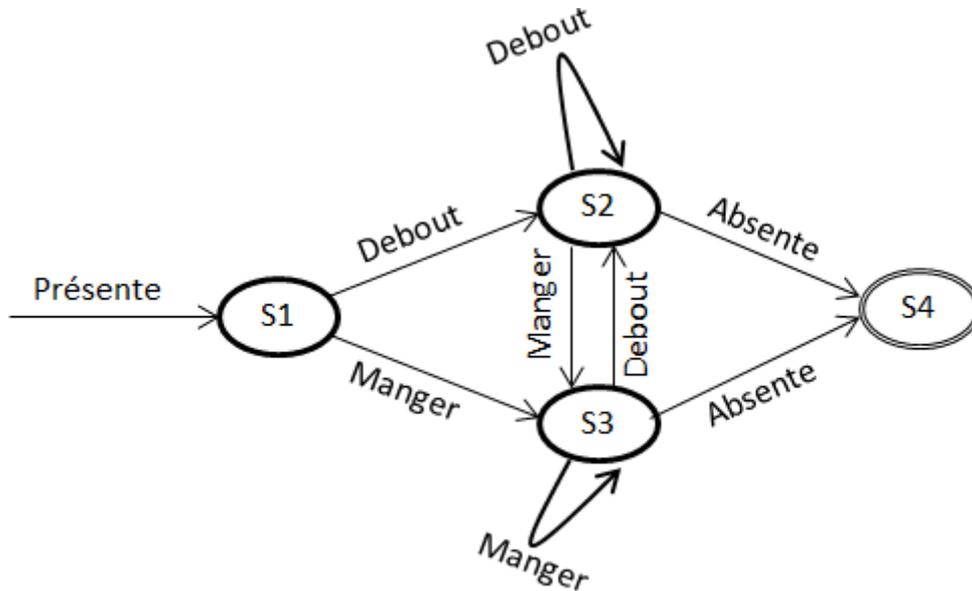


Fig 5.5: Automate d'états finis déterministes qui modélise une période d'alimentation.

sont consécutifs. La durée moyenne de la collecte des données d'entraînement était de cinq heures par jour (entre 8h00 et 13h00 le premier jour et entre 10h30 et 15h30 le deuxième jour). Cependant, les données de test ont été collectées pendant d'autres semaines et des jours non consécutifs (J3, J4 et J5). La durée moyenne de la collecte des données de test était de trois heures par jour (8h30-11h30 le jour 3, 10h00-13h00 le jour 4 et 13h30-16h30 le jour 5). Ce choix a été fait pour s'assurer que les images utilisées dans le test sont différentes des images d'entraînement. Lors de l'acquisition des images d'entraînement, le capteur a été installé dans une seule entrée de mangeoire pour les deux jours (J1 et J2). Cependant, pour avoir une grande variation entre les images d'entraînement et de test, le capteur a été installé à différents endroits pour les trois jours (J3, J4 et J5).

Tout au long de l'expérience, un observateur humain était présent. Il a noté l'identité de chaque vache et a filmé toutes les scènes avec une caméra. Ensuite, une annotation a été faite sur les moments où les vaches étaient présentes, absentes, debout et en train de se nourrir.

#### 5.2.4 Méthode de classification

Pour la classification des images, nous avons utilisé un algorithme basé sur les réseaux de neurones à convolution (CNNs) comme modèle d'apprentissage profond et supervisé. L'avantage des CNNs est que les caractéristiques sont extraites automatiquement en phase d'entraînement et que l'erreur de classification (loss) est minimisée afin d'optimiser les performances de classification et cela en ajustant les poids du réseau. Les CNNs donnent d'excellents résultats de

reconnaissance et peuvent être ré-entraînés dans le cadre de l'apprentissage par transfert.

La taille des images de la base de données a été redimensionnée de 640\*480 à 100\*100 pour avoir une grande vitesse d'exécution du modèle CNN. L'architecture du CNN utilisé dans ce travail est illustrée par la Fig 5.6. Elle est constituée d'un premier ensemble de couches d'extraction de caractéristiques suivi d'un second ensemble de couches de classification. Tout d'abord, quatre couches de convolution où chaque couche est suivie d'une couche de pooling sont utilisées pour l'extraction de caractéristiques. Ensuite, trois couches entièrement connectées (Inner Product), deux couches relues et deux couches dropout sont utilisées pour établir classification. Le CNN est implémenté à l'aide du Framework caffe [221]. Durant tout le travail, cette architecture sera gardée sauf le nombre de classes de sortie du réseau qui va être modifié en fonction de la classification à effectuer.

Deux sorties sont utilisées dans la couche de classification pour entraîner le premier et le second réseaux de neurones à convolution (CNN1 et CNN2). Le CNN1 réalise une classification binaire et crée un modèle qui détecte la présence de vaches dans la zone d'alimentation. Le CNN2 réalise la classification binaire du comportement des vaches en deux classes (debout et manger). Six sorties seront utilisées dans la couche de classification pour entraîner le troisième réseau (CNN3) qui identifie le type d'aliment et dix-sept sorties seront utilisées pour entraîner le réseau qui permet d'identifier les dix-sept vaches (CNN4). Les paramètres de ces deux réseaux (CNN3 et CNN4) seront discutés dans les sections suivantes.

Le CNN1 et CNN2 peuvent être vus comme des modèles de sélection d'images. Ils permettent de sélectionner les images de zone d'alimentation vide (sans vaches) utiles pour l'identification du type d'aliment et les images de vache en période d'alimentation utiles pour le suivi alimentaire et l'identification individuelle des vaches.

Chacun des CNNs utilise un ensemble de données d'apprentissage contenant des images pour l'entraînement et d'autres pour le test, comme le montre le tableau 5.1. Dans l'ensemble des images acquises et sélectionnées, nous avons compté 7801 images dont 7265 sont des images où des vaches sont présentes et seulement 536 sont des images où des vaches sont absentes. Comme la classe des images de vaches présentes était très dominante par rapport à celle des images de vaches absentes, seul un sous-ensemble d'images de vaches présentes a été retenu. Le nombre d'images de vaches présentes (7265 images) a été réduit à 536 images pour former un ensemble de données d'apprentissage de 1072 images (536 vaches présentes et 536 vaches absentes) pour le CNN1. À partir de cet ensemble de données, 60 % des images a été utilisé pour l'entraînement (643 images) et 40 % des images (429 images) a été utilisé pour le test. Une fois que le CNN1 est validé, il classe toutes les images (7801). Toutes les images classées comme vaches laitières présentes dans la zone d'alimentation (7265) sont utilisées comme un ensemble de données d'apprentissage pour CNN2 dont 4393 images pour

CNN	Données sélectionnées	Données D'apprentissage	Images d'entraînement	Images de test
CNN1	7801	1072	643	429
CNN2	7265	7265	4393	2872
CNN3	5411	1068	642	426
CNN4	4875	4875	2931	1944

Tableau 5.1: Nombre d'images utilisées dans l'entraînement et le test des CNNs

l'entraînement et le reste pour le test. Toutes les images classées par CNN2 comme vache laitière en état d'alimentation (4875 images) ont été segmentées par le modèle Xception et utilisées comme données d'apprentissage pour CNN4. Dans les images segmentées, les pixels de la zone d'alimentation sont remplacés par des pixels blancs. Parmi ces images segmentées, 2931 ont été utilisées pour l'entraînement et 1944 pour le test. L'ensemble de données d'apprentissage de CNN3 contient les images de vaches laitières absentes et un ensemble d'images segmentées. Dans ce cas, dans les images segmentées, les pixels de la vache laitière ont été remplacés par des pixels blancs. Comme les images segmentées étaient très dominantes par rapport aux images de vaches absentes (4875 images segmentées et 536 d'images de vaches absentes), le nombre d'images segmentées a été réduit à 532 images. Ainsi, le jeu de données d'apprentissage de CNN3 compte 1068 images (536 images de vaches laitières absentes et 532 images segmentées).

### 5.2.5 Segmentation d'images

L'étape précédente a permis de sélectionner toutes les images dans lesquelles les vaches étaient en train de s'alimenter. Chacune de ces images a été segmentée pour les subdiviser en deux sous-groupes d'images. Le premier groupe est celui qui représente l'avant-plan des images, il contient les pixels qui représentent les vaches. Le deuxième groupe est constitué de l'arrière-plan des images, il contient des informations sur la mangeoire. Les images des vaches sont utilisées par le CNN4 pour l'identification. Les autres images peuvent aussi être utilisées pour classer et indiquer automatiquement le type d'aliment présent dans la mangeoire.

Les méthodes de segmentation peuvent être classées en deux catégories : méthodes traditionnelles et méthodes d'apprentissage approfondi. La performance des algorithmes de segmentation traditionnels est fortement influencée par des facteurs tels que le mauvais éclairage, les ombres et les variations de fond causées par la présence des éleveurs, des camions et d'autres animaux dans les images [222]. Dans les méthodes d'apprentissage approfondi, les réseaux neuronaux convolutionnels (CNN) ont récemment été utilisés dans la segmentation des images des fermes et ont atteint de bonnes performances en termes d'exactitude et de

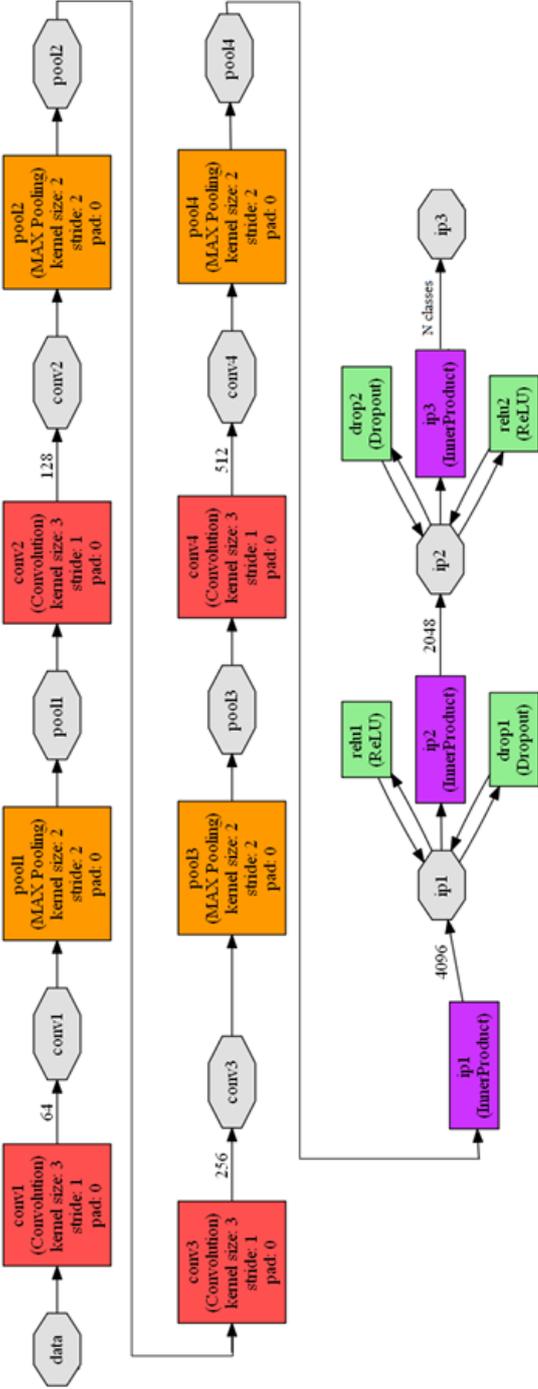


Fig 5.6: Architecture du réseau de neurones à convolution

précision [222–224] et peuvent surmonter les impacts environnementaux [223]. En effet, les caractéristiques apprises par ces réseaux comprennent plusieurs informations spatiales et sémantiques. Il existe deux types d'approches de segmentation utilisant l'apprentissage profond. La segmentation sémantique [223] et la segmentation d'instance [224]. La méthode de segmentation d'instance (Masque RCNN [222, 224]) a récemment été utilisée pour la segmentation des images de vaches et a atteint un taux de précision de 92% [222]. Cependant, par rapport aux approches de segmentation sémantique, les approches de segmentation d'instance sont plus lentes et ont un temps de calcul considérablement plus élevé [225]. De plus, la segmentation sémantique permet d'obtenir des taux de sensibilité plus élevés [226].

DeepLab [227] est l'une des techniques les plus prometteuses pour la segmentation sémantique d'images en utilisant des techniques d'apprentissage profond proposé sur tensorflow [228]. Cette segmentation est utilisée en imagerie pour regrouper les pixels qui appartiennent à la même classe; elle permet à la fois de segmenter et de reconnaître les classes. Deeplab peut être utilisé dans deux cadres d'apprentissage à savoir l'apprentissage profond et l'apprentissage par transfert. L'apprentissage profond permet de créer un nouveau modèle d'apprentissage. Ce type nécessite beaucoup de temps pour l'entraînement. Cependant, l'apprentissage par transfert permet d'utiliser un modèle déjà existant. Il sera ré-entraîné pour qu'il s'adapte à la nouvelle segmentation souhaitée. Deeplab fournit trois architectures de modèles de segmentation préentraînés qui sont mobilenet [229], Xception [230] et resnet [231]. Ces architectures ont été déployées pour améliorer les résultats des CNNs. Pour effectuer l'apprentissage par transfert sur ces architectures, la couche de sortie sera adaptée et le réseau sera ré-entraîné.

Par défaut, le critère d'arrêt de l'entraînement sur le Framework tensorflow est configuré par un nombre d'itérations prédéfini. Dans notre travail, nous avons apporté une modification à ce critère de façon à ce que l'entraînement continu jusqu'à atteindre une valeur de la fonction de perte inférieure ou égale à 0,2. En effet, lors de l'entraînement nous avons constaté que la fonction loss ne converge pas, pour cela nous avons choisi d'utiliser ce critère d'arrêt. Puis, nous avons choisi de ré-entraîné le modèle xception65\_coco\_voc\_trainval déjà pré-entraîné avec la base de données PASCAL VOC 2012 pour segmenter 21 classes. Comparer aux autres modèles pré-entraînés Xception65\_coco\_voc\_trainval a obtenu le meilleur taux de segmentation (87,80 % PASCAL mIOU). Dans notre cas, ce modèle a été entraîné avec une approche de transfert d'apprentissage où seul un sous-ensemble d'images a été sélectionné (90 images) pour accélérer le processus d'entraînement. Nous avons utilisé 72 images pour l'entraînement (12 images pour la mangeoire vide et 12 images pour chaque type de nourriture dans la mangeoire) et 18 images pour le test. Ensuite, les performances ont été calculées pour l'ensemble des images classées par CNN2 comme des vaches laitières en état d'alimentation (4875 im-

ages).

En plus du calcul de l'exactitude, de la sensibilité et de la précision pour évaluer la performance de la classification et de la segmentation, d'autres paramètres sont utilisés pour évaluer la performance de la segmentation tels que la moyenne d'intersection sur l'union (dite en anglais Mean Intersection Over Union (MIOU)) et la spécificité. En effet, le calcul de la valeur MIOU donne la similarité entre la région prédite et la réalité de terrain. Cette valeur est estimée selon l'équation (5.7). La spécificité représente la proportion de pixels de fond que la méthode a correctement réduit (correctement classés comme arrière-plan). Cette réduction de l'arrière-plan (spécificité) est estimée selon l'équation (5.8).

$$MeanIOU = \frac{AreaofOverlap}{AreaofUnion} \quad (5.7)$$

$$Background.reduction = 1 - \frac{Background.pixels.number.after.segmentation(segmented.image)}{Background.pixels.number.before.segmentation(original.image)} \quad (5.8)$$

La fonction de perte est calculée par la méthode d'entropie croisée Softmax [232]. Pour un échantillon donné avec une étiquette de classe  $y$ , la fonction de perte d'entropie est calculée par l'équation suivante :

$$entropie.croisée.Softmax(z,y) = -\log\left(\frac{\exp(z_y)}{\sum_{j=1}^C \exp(z_j)}\right). \forall y \in 1, 2, \dots, C \quad (5.9)$$

où  $C$  est le nombre total de classes et  $z$  la sortie prévue du modèle pour toutes les classes.

Des poids différents ont été attribués pour chaque classe lors de ce calcul. Le coefficient de l'avant-plan (le corps de la vache) vaut 3 et celui de l'arrière-plan (la zone d'alimentation) vaut 1. Ces valeurs ont été choisies pour deux raisons. La première est que la prévalence des pixels de l'arrière-plan est non équitable par rapport à celle de l'avant-plan; les pixels de l'arrière-plan sont plus nombreux. Les différentes valeurs de poids des classes lors du calcul de la fonction de perte permettent au réseau de ne pas se biaiser vers cette classe dominante. La deuxième est que le rôle de la segmentation est de réduire l'arrière-plan et surtout ne pas perdre aucune information de l'avant-plan. En effet, l'erreur de segmenter l'arrière-plan comme avant-plan est acceptable, tandis que la segmentation de l'avant-plan comme arrière-plan est inacceptable. La détection de l'avant-plan doit avoir une grande valeur de sensibilité (tous les pixels de l'avant-plan doivent être détectés) et la détection de l'arrière-plan doit être précise (tous les pixels détectés en tant qu'arrière-plan doivent être exacts).

## 5.2.6 Identification des vaches

L'étape de classification comportementale a permis de sélectionner automatiquement les images des vaches en train de s'alimenter. Puis, l'étape de segmentation a permis d'extraire les pixels qui représentent la vache dans chaque image. Cela, a permis de former une nouvelle base de données comptant que des pixels de vaches. Pour assurer l'identification individuelle des vaches, trois approches de classification sont utilisées. Les performances de ces approches sont comparées pour trouver la meilleure approche.

### 5.2.6.1 Approche 1 : Classification par réseau de neurones à convolution (apprentissage profond)

La première approche consiste à utiliser l'architecture CNN présentée dans la Fig 5.6 avec dix-sept classes de sortie. Chaque classe correspond à une vache.

### 5.2.6.2 Approche 2 : Extraction de caractéristiques par un réseau de neurones à convolution et classification par machine à vecteurs de support (SVM).

Cette approche propose de créer un modèle de classification hybride entre les réseaux de neurones à convolution et la machine à vecteurs de support (SVM) de sorte à utiliser le CNN pour la sélection et l'extraction de caractéristiques, puis réaliser la classification de ces caractéristiques par la machine à vecteurs de support (SVM). Cette approche est inspirée de [233] où les auteurs ont combiné un CNN et un SVM pour la reconnaissance de chiffres manuscrits.

Le modèle de classification entraîné dans l'approche 1 est utilisé pour sélectionner et extraire les caractéristiques de chaque image et former une nouvelle base de données contenant ces caractéristiques. Puis un classificateur SVM est entraîné et utilisé pour l'identification des vaches.

### 5.2.6.3 Approche 3 : CNN pour regroupement puis Multi CNN pour classification

Cette approche utilise un premier CNN pour séparer les vaches en deux catégories. Puis, utilise deux autres CNNs pour identifier les vaches présentes dans chaque catégorie comme le montre la Fig 5.7. Les vaches sont classées selon la distribution des couleurs en deux groupes. Le groupe 1 contient les vaches entièrement noires et le groupe 2 contient les vaches avec des caractéristiques blanches (vache noir et blanc).

Le nombre de classes (sorties) des CNNs utilisés pour l'identification dépend du nombre de vaches présentes dans chaque catégorie (12 vaches entièrement noires et 5 vaches avec des caractéristiques blanches).

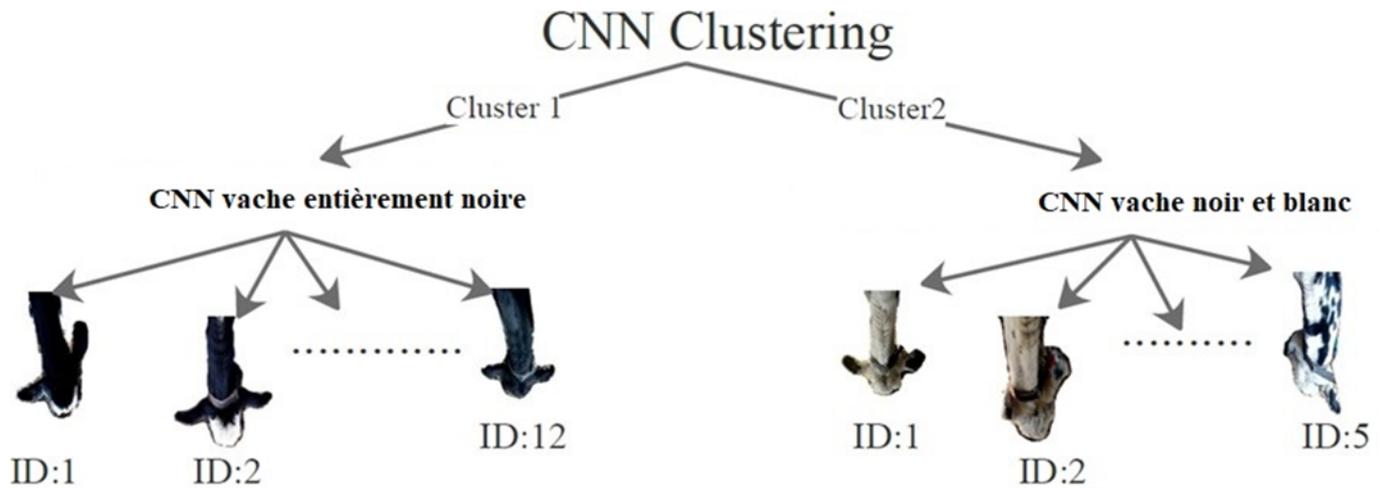


Fig 5.7: Schéma explicatif de la troisième approche.

### 5.2.7 Disponibilité de l'aliment et identification de son type

La dernière classification permet de détecter la présence d'aliments dans la mangeoire et de spécifier son type (foin, herbe fraîche mélange 1, herbe fraîche mélange 2, ensilage, concentré). Cette classification permet une meilleure gestion de la ferme. Elle permet essentiellement de :

- Contrôler les horaires de distribution d'aliment de manière automatique ;
- Offrir un meilleur pilotage alimentaire en calculant le temps d'alimentation associé à chaque type d'aliment ;
- Apporter de nouvelles informations à l'analyse comportementale. En effet, une vache n'est classée en période d'alimentation que si les aliments sont présents dans la mangeoire.

Pour réaliser cette classification, les images acquises par le capteur embarqué et classées comme étant zone d'alimentation vide par le CNN1 sont regroupées avec les images segmentées comme étant l'arrière-plan par le modèle Xception pour former une nouvelle base d'images. Un nouveau modèle (CNN3) avec six classes de sortie est utilisé. Une classe correspondra à la mangeoire sans aliment et les autres classes correspondent aux différents types d'aliments comme le montre la Fig 5.8.

Comme c'est indiqué dans le tableau 5.1, un total de 1068 images est utilisé comme données d'apprentissage. Ce nombre d'images est équitablement réparti en six classes (mangeoire vide et mangeoire avec l'un des cinq types d'aliments distribués) qui sont indiquées par la Fig. 5.8. Ainsi, il y a 178 images par classe (107 images sont utilisées pour l'entraînement de chaque classe et 71 images sont utilisées pour le test de chaque classe).

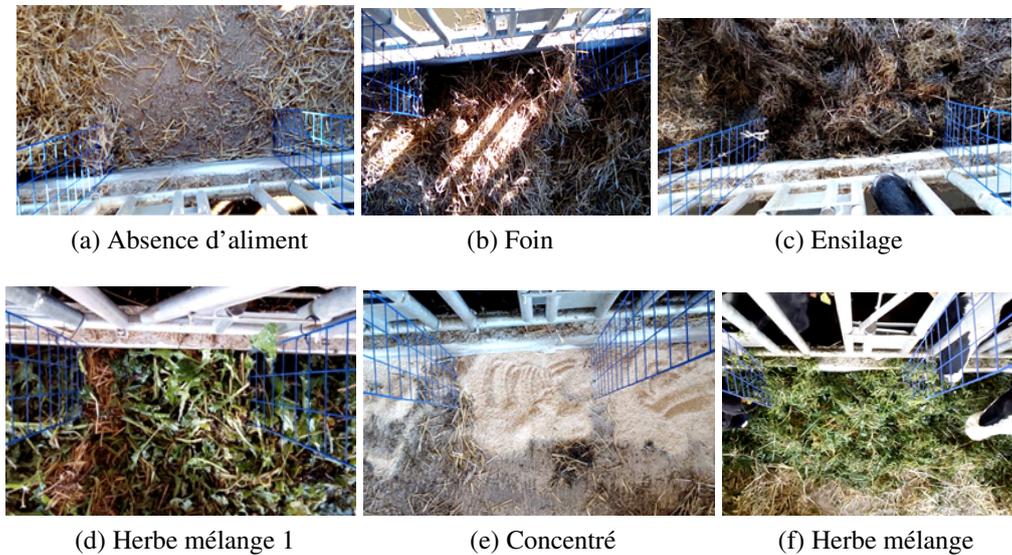


Fig 5.8: Les différents aliments distribués dans la mangeoire.

## 5.3 Résultats

### 5.3.1 Acquisition et sélection d'images

Le tableau 5.2 montre les différents résultats obtenus lors du déploiement du capteur embarqué pour l'acquisition et la sélection d'images. 19 409 images ont été acquises par le système avec une fréquence d'une image par seconde. La soustraction de l'arrière-plan par AGMM a permis de réduire ce nombre de 43,62 % et le calcul de l'index de similitude a permis de garder 37,43 % des images acquises. Dans l'ensemble des images sélectionnées 31,42 % des images représentent un état dans lequel les vaches sont debout (ne sont pas en état d'alimentation) et 68,85 % des images où les vaches sont en état d'alimentation (mange).

Des seuils élevés ont été utilisés lors du calcul de AGMM (40 000) et SSIM (0,8). Ce choix est effectué pour avoir une base de données ne contenant que des images consécutives avec une réelle variation dans les pixels. Une diminution de ces seuils augmentera le nombre d'images sélectionnées et engendrera une surcharge de traitement dans la méthode de classification. Par contre, les seuils élevés utilisés permettent de réduire le nombre d'images.

### 5.3.2 Détection de présence et suivi du comportement alimentaire (CNN1 et CNN2)

Le premier CNN a été entraîné pour réaliser une classification binaire et permettre la détection de la présence de la vache dans la zone d'alimentation. Pour cette classification,

Vache	Images	AMMG	SSIM	Mange	Debout
<b>1</b>	592	340	282	64	218
<b>2</b>	1467	809	583	476	107
<b>3</b>	1331	558	352	140	212
<b>4</b>	1367	787	516	307	209
<b>5</b>	822	600	447	271	176
<b>6</b>	1010	619	513	383	130
<b>7</b>	1212	914	397	394	3
<b>8</b>	1181	717	553	409	144
<b>9</b>	1060	637	349	254	95
<b>10</b>	1012	423	256	129	127
<b>11</b>	1033	622	415	175	240
<b>12</b>	1297	551	399	196	203
<b>13</b>	1271	694	415	254	161
<b>14</b>	1317	860	707	658	49
<b>15</b>	1138	615	264	225	39
<b>16</b>	1197	435	215	169	46
<b>17</b>	1102	763	602	478	124
<b>ALL</b>	19409	10944	7265	4982	2283

Tableau 5.2: Nombre d'images acquises pour chaque vache avant et après la sélection.

le taux d'exactitude est optimal (100%). Il a été obtenu après seulement quelques étapes d'entraînement. Le CNN1 distingue facilement l'état de présence de la vache dans la zone d'alimentation et cela en reconnaissant la présence de certaines parties de la vache (les yeux, le museau, le cou. etc). Il trouve rapidement les meilleures caractéristiques qui séparent les deux classes (présence et absence de la vache). La Fig 5.9 (b) montre le résultat de la fonction perte et de l'exactitude de la méthode en fonction des itérations durant l'entraînement. Le modèle engendré a été testé sur l'ensemble des images disponibles dans l'échelle de données et il a montré un résultat optimal.

Le deuxième CNN a été entraîné pour réaliser une classification binaire et permettre de suivre le comportement alimentaire des vaches. Le taux de classification de ce CNN est élevé (92,61 %). Il a été atteint grâce à la grande différence qui existe entre les caractéristiques de la vache lorsqu'elle est en train de manger et lorsqu'elle est debout. Lorsque la vache est en période d'alimentation son cou est incliné vers le bas et permet au réseau de suivre seulement trois caractéristiques de forme qui sont la forme du cou, la forme des oreilles et la forme du haut de la tête. Cependant, lorsque le cou de la vache est soulevé des modifications importantes vont apparaître. Ces modifications sont la présence de nouvelles parties comme les yeux, le museau, etc. et un grand changement dans la forme des caractéristiques déjà

N°itérations			Classe prédite		Exactitude
			Manger	Debout	
4200	Classe réelle	Manger	1694	264	86,52
		Debout	50	864	94,53
		Précision	97,13	76,59	89,06
5800	Classe réelle	Manger	1862	97	95,05
		Debout	115	798	87,40
		Précision	94,18	89,16	92,61

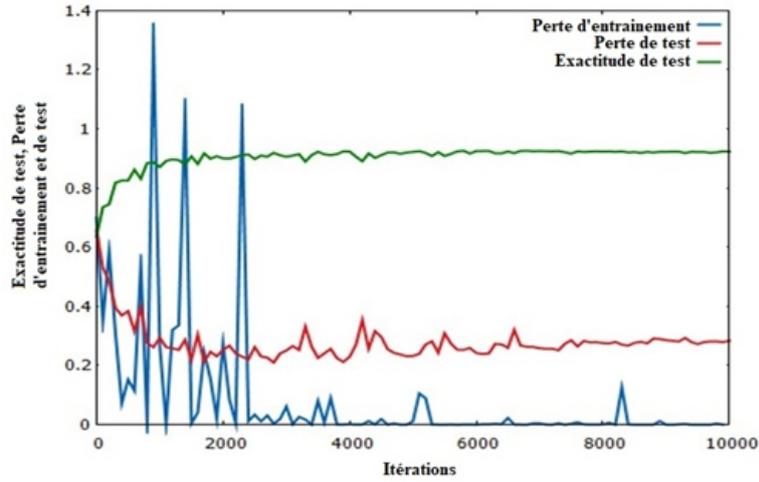
Tableau 5.3: Matrice de confusion, exactitude et précision du CNN2.

suivies lorsque le cou est incliné vers le bas (forme de la tête, du cou et des oreilles). La Fig 5.10 montre un exemple de la map de caractéristiques extraites de l'avant-plan et de l'arrière-plan par le CNN2 dans les deux classes suivies. Les régions rouges sont définies comme celles qui contiennent les caractéristiques les plus importantes suivies par le réseau. Dans les deux cas (vache debout ou vache en train de se nourrir) la région la plus importante était le haut de la tête. CNN2 a déterminé que cette région était la plus appropriée pour séparer les deux classes. Les régions en vert indiquent des caractéristiques moins importantes. Ces régions se trouvent le plus souvent au niveau de la zone d'alimentation. L'explication de l'activation de ces régions est que lorsque la vache est debout, elle est plus proche de la caméra que lorsqu'elle se nourrit. Cet éloignement est utilisé comme une caractéristique secondaire par CNN2. La Fig 5.9-a permet de visualiser le résultat des différentes fonctions de perte et d'exactitude pour les données d'entraînement et de test du CNN2.

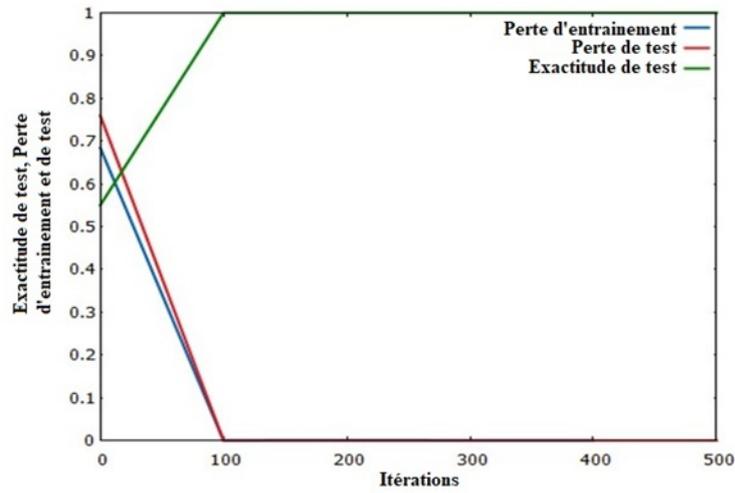
En fonction du nombre d'itérations du réseau, différents taux de classification ont été atteints. La meilleure exactitude a été atteinte dans l'itération 5800 avec un taux de 92,61 % et la meilleure précision a été atteinte dans l'itération 4200. Le tableau 5.3 montre les résultats obtenus lors du test de ces deux modèles (générés en itérations 5800 et 4200). Le modèle le plus précis est utilisé pour la sélection automatique d'images afin de construire la base de données d'identification.

### 5.3.3 La segmentation sémantique

Le but de cette étape est de réaliser une segmentation binaire des images en avant-plan et arrière-plan. Cela est effectué en réentraînant le modèle Xception pour séparer les pixels des vaches de ceux de la zone d'alimentation. La Fig 5.11 montre les résultats de la fonction Loss en fonction du nombre d'itérations pendant l'entraînement. On remarque que cette fonction ne converge pas mais elle diminue jusqu'à atteindre une valeur de 0,2 après 1144 itérations. Le tableau 5.4 donne les performances de classification pour des valeurs différentes de la fonction

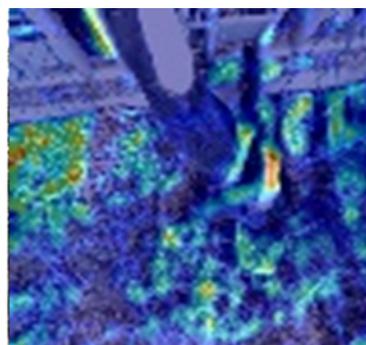


(a) Comportement alimentaire

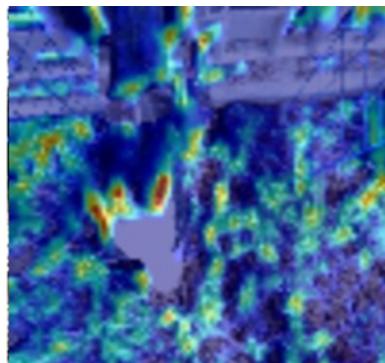


(b) Détection de présence

Fig 5.9: Valeurs de la fonction de perte et d'exactitude durant l'entraînement et le test du CNN2 (a) et du CNN1 (b).



(a) classe : manger



(b) classe : debout

Fig 5.10: Caractéristiques extraites par la dernière couche de convolution du CNN2.  
(a) : vache en période d'alimentation, (b) : vache debout.

Loss	0,6	0,5	0,4	0,3	0,2
Itération	119	134	267	361	1144
Mean IOU	0,50	0,67	0,71	0,76	0,78
Réduction de pixels	0,76	0,88	0,91	0,93	0,94

Tableau 5.4: Résultats obtenus lors du réentraînement du modèle Xception pour la segmentation des images

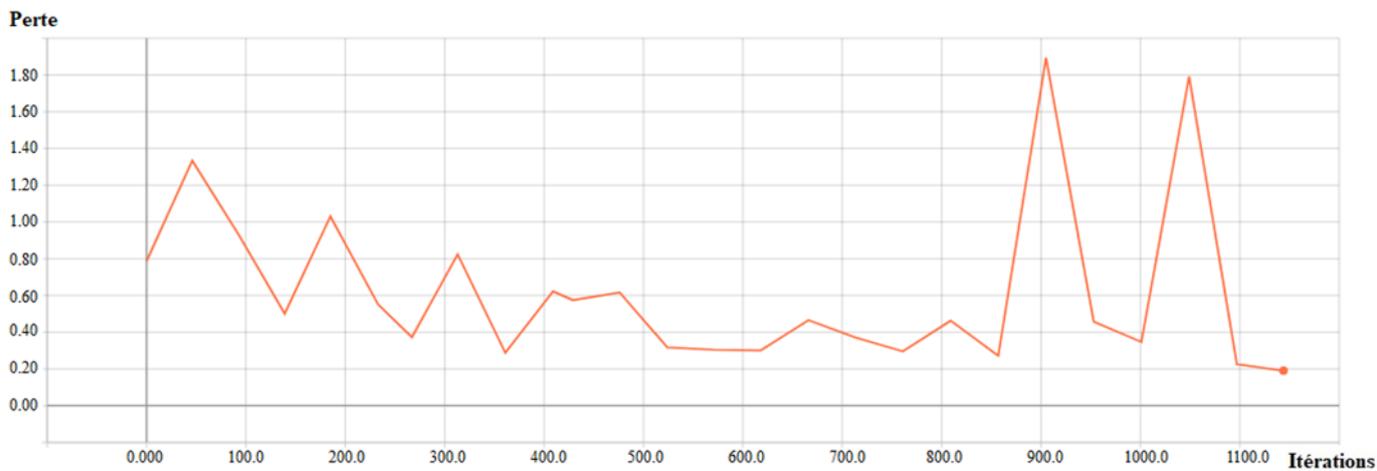


Fig 5.11: Résultat de la fonction de perte (Loss) lors du réentraînement du modèle Xception pour la segmentation des images.

Loss. On remarque que le modèle de classification permet de réduire jusqu'à 94 % de pixels de l'arrière-plan. De plus, le modèle permet de sauvegarder tous les pixels de l'avant-plan. Avant l'étape de segmentation, les proportions des pixels de l'avant-plan et ceux de l'arrière-plan étaient respectivement de 19 % et 81%. Puis, la segmentation a permis d'avoir une nouvelle distribution de pixels (avant-plan 78 % et arrière-plan 22 %).

Le modèle de segmentation obtenu a été utilisé pour segmenter toutes les images de vaches en période d'alimentation. Le même taux de réduction de l'arrière-plan a été observé (94 %). Le modèle a permis de préserver tous les pixels de l'avant-plan dans 99,8 % des images de l'ensemble de données. La Fig 5.12 montre des exemples de cette segmentation sémantique. Des résultats différents peuvent être trouvés en modifiant les poids des classes dans le calcul de la fonction Loss. Pour homogénéiser la prévalence de nos classes, les meilleures valeurs de poids sont 1 pour l'arrière-plan et 2,63 pour l'avant-plan. Cependant, nous avons choisi d'utiliser le poids 3 pour l'avant-plan pour forcer l'entraînement à préserver ses pixels.



Fig 5.12: Résultats de la segmentation d'images de huit différentes vaches en appliquant le modèle Xception réentraîné.

		Estimation																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Classe réelle	1	32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2	4	187	1	0	2	0	0	0	0	1	1	1	0	0	0	0	0
	3	0	0	57	0	2	0	0	0	0	0	0	0	0	0	0	0	0
	4	0	0	0	70	0	6	0	0	0	0	0	0	0	0	0	0	0
	5	1	0	0	0	114	0	8	0	0	0	0	1	0	0	0	1	0
	6	0	0	0	3	0	131	0	0	0	0	0	0	0	0	0	0	0
	7	0	0	0	0	0	0	157	0	0	0	0	0	0	0	0	0	0
	8	0	0	0	0	0	0	0	162	4	0	4	0	0	0	0	0	0
	9	0	0	0	0	0	0	0	2	91	1	1	1	4	0	0	0	0
	10	0	0	0	0	0	0	0	0	0	41	3	0	0	0	0	0	0
	11	0	0	0	0	0	0	0	0	0	0	62	3	0	0	0	0	0
	12	0	0	0	0	0	0	0	0	0	0	1	73	1	2	0	0	0
	13	0	0	0	0	0	0	0	0	0	0	0	1	101	1	0	0	0
	14	0	0	0	0	0	0	0	0	0	0	0	1	0	261	0	2	0
	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	81	0	0
	16	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	67	0
	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	190

Tableau 5.5: Matrice de confusion de la première approche.

### 5.3.4 Identification des vaches

Le tableau 5.5 représente la matrice de confusion obtenue lors du test de la première approche (classification par réseau de neurones à convolution). Un taux total d'exactitude (96,55 %) a été obtenu. Cette classification a permis de classer les vaches : 1, 7, 15 et 17 avec un taux d'exactitude optimal (100 %). L'exactitude de classification des autres vaches dépasse 91 % ; les vaches 3, 6, 8, 11, 13, 14 et 16 ont été classées avec une exactitude élevée comprise entre 95 % et 99 % et les vaches 2, 4, 5, 9, 10 et 12 ont eu un taux de classification au-dessus de la moyenne (entre 91 % et 94 %).

Nous avons utilisé la méthode Grad-CAM [234] pour visualiser les régions suivies par le modèle d'identification entraîné par la première approche. La Fig 5.13 montre la Map de caractéristiques du CNN4 et permet de comprendre les caractéristiques suivies par ce réseau pour distinguer les vaches. Cette MAP montre les différentes régions extraites par la troisième et la dernière couche à convolution chez quatre vaches (id: 02, 05, 07, 17). Les régions en rouge montrent les caractéristiques les plus importantes suivies par le CNN4. Il est à noter que pour les vaches 2, 7 et 17, beaucoup de caractéristiques ont été suivies par le réseau. Le cou, l'oreille et le haut de la tête étaient souvent en rouge. Les vaches 2, 7 et 17 sont des vaches noires avec des motifs blancs. Il y a donc une grande variation de textures entre ces vaches, c'est pourquoi le réseau à utiliser toutes ces textures lors de l'extraction des caractéris-

tiques. Cependant, pour la vache 5, le CNN4 s'est concentré sur un nombre plus restreint de caractéristiques. Le CNN4 n'utilise que de petites zones (comme une oreille) pour identifier les vaches. Ce résultat peut être interprété comme suit : la vache 5 est complètement noire donc elle ne possède aucune texture. Pour cela, le CNN4 utilise la différence qui existe dans la forme d'une partie de la tête de la vache.

L'utilisation de la deuxième approche (CNN pour l'extraction des caractéristiques et SVM pour la classification) a permis d'apporter une légère augmentation (de 0,17 %) dans le taux d'exactitude globale de classification. L'explication de cette augmentation est que par rapport au CNN seul, la combinaison CNN-SVM a une plus grande capacité de généralisation [235] et permet d'augmenter les taux de classification [235].

Le regroupement des vaches en cluster par la troisième approche a eu une exactitude de 100 %. Le taux d'exactitude de la classification des vaches dans le cluster des vaches noires est de 96,24 % et celui de la classification dans le cluster des vaches noires avec des motifs blancs est de 98,89 %. L'exactitude globale de cette troisième approche est de 97,06 %. Le tableau 5.6 montre les différentes performances de classification des trois approches.

D'après le tableau 5.6 nous constatons que l'approche 3 est plus précise que les deux autres approches. Au lieu d'identifier toutes les vaches par un seul CNN, l'approche 3 crée plusieurs groupes et utilise donc un CNN pour classer un petit nombre de vaches laitières dans chaque groupe. Le résultat de cette approche montre que l'utilisation de plusieurs CNN (un sur chaque groupe) améliore le taux de précision global de 0,51 % tout en réduisant le nombre de vaches laitières à classer par chaque CNN. Dans les grandes exploitations agricoles, l'identification individuelle sera certainement plus complexe en raison du grand risque de trouver des vaches laitières très similaires. Pour faire face à cette complexité, cette approche peut être très utile. En effet, le grand nombre de vaches sera divisé en plusieurs groupes. Ces groupes seront constitués en fonction de différents critères tels que la répartition des couleurs sur les vaches laitières (vaches laitières entièrement noires, vaches laitières noires avec des motifs blancs, et vaches laitières entièrement blanches), la présence de cornes, de collier et même en fonction des races de vaches (dans les exploitations laitières de croisement). Lorsque des vaches semblent être très semblables dans un même groupe, on peut leur attacher des colliers de différentes couleurs ou utiliser des tatouages différents pour les distinguer.

### **5.3.5 Disponibilité de l'aliment et identification de son type**

L'utilisation du réseau de neurones à convolution pour détecter la présence d'aliment et spécifier son type a donné une exactitude optimale (100 %). Les aliments distribués se caractérisent par une grande variation dans leurs formes, couleurs et textures. Ce taux a été obtenu après seulement 200 itérations. Les résultats montrent la grande capacité de compréhension de la

<b>Id Vache</b>	<b>Exactitude Approche 1</b>	<b>Exactitude Approche 2</b>	<b>Exactitude Approche 3</b>
<b>1</b>	100,00	96,87	90,62
<b>2</b>	94,92	98,48	97,46
<b>3</b>	96,61	86,66	100,00
<b>4</b>	92,11	88,31	96,05
<b>5</b>	91,20	97,61	91,30
<b>6</b>	97,76	97,76	96,99
<b>7</b>	100,00	99,36	100,00
<b>8</b>	95,29	98,83	96,47
<b>9</b>	91,00	98,00	90,91
<b>10</b>	93,18	95,55	100,00
<b>11</b>	95,38	81,81	98,46
<b>12</b>	94,81	97,43	92,21
<b>13</b>	98,06	91,34	95,15
<b>14</b>	98,86	97,72	99,24
<b>15</b>	100,00	98,78	100,00
<b>16</b>	95,71	100	98,55
<b>17</b>	100,00	100	100,00
<b>ALL</b>	96,55	96,72	97,06

Tableau 5.6: Taux d'exactitude des trois approches

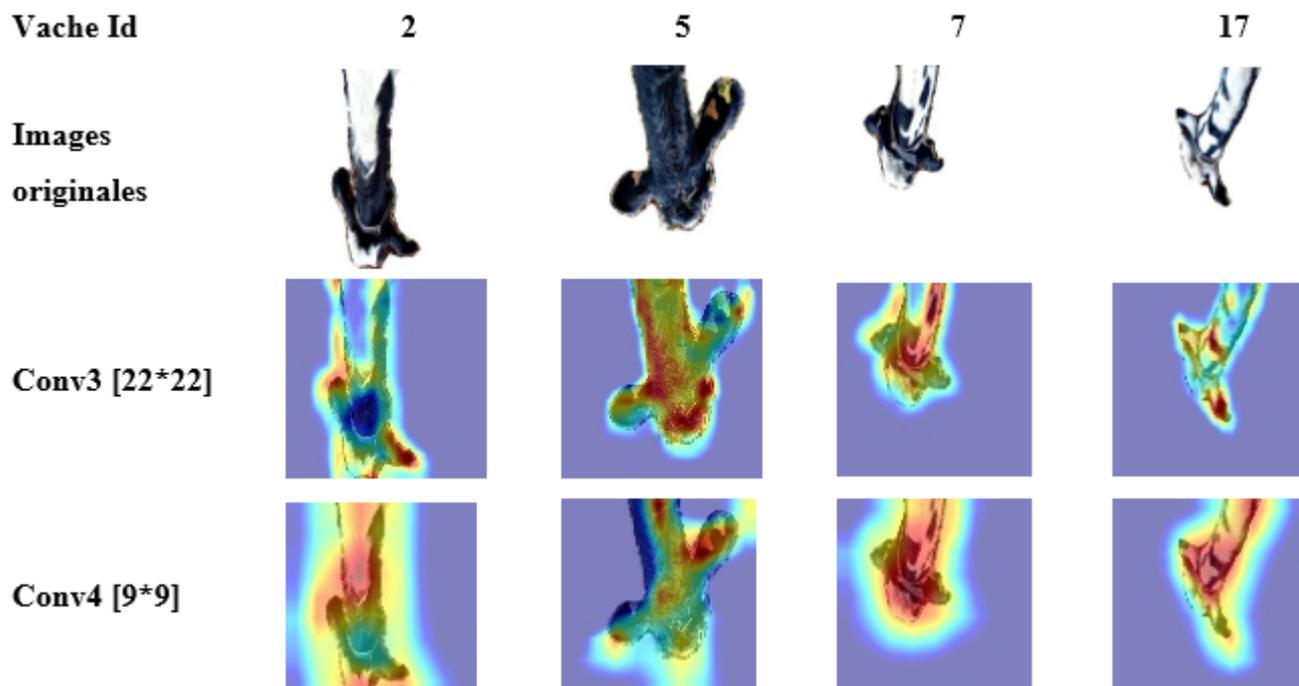


Fig 5.13: Caractéristiques extraites par la troisième et la quatrième couche de convolution du CNN4.

scène de l'élevage par l'analyse d'images permettant ainsi d'apporter une nouvelle valeur informative au pilotage alimentaire.

## 5.4 Comparaison des résultats

La méthode proposée dans ce travail est flexible. Elle permet de réaliser un enchaînement de méthodes de classification pour assurer la détection de présence, d'identification et de suivi du comportement alimentaire de vaches. La comparaison entre les performances obtenues et celles des autres travaux de recherche est présentée dans les tableaux 5.7, 5.8 et 5.9.

La première fonctionnalité assurée par la méthode proposée est la détection de la présence des vaches. La méthode d'analyse d'images proposée par [236] permet une détection simultanée de la présence de plusieurs vaches laitières dans la zone d'alimentation et atteint une sensibilité de 87 % (voir le tableau 5.7). Cependant, dans notre travail, une sensibilité optimale (100%) a été atteinte. Cette performance élevée est due à deux critères de conception. Le premier critère est l'éloignement de la caméra de la mangeoire. En effet, dans [236] la caméra a été installée en haut du bâtiment d'élevage. Par contre, dans le travail courant la caméra est fixée dans une entrée de la mangeoire ce qui facilite la reconnaissance. Le second

Travail	Equipment	emplacement	performance	Taux
Notre	Caméra	Haut de la mangeoire	Exactitude	100 %
			Sensibilité	100 %
Porto et al. [236]	Caméra	Haut du bâtiment	Sensibilité	87 %

Tableau 5.7: Comparaison des performances de classification pour la détection de la présence des vaches.

Travail	Capteur	Emplacement	Performance	Taux
notre	Camera	Vue de haut	Exactitude	92,61 %
			Précision	97.13 %
[161]	Accéléromètre	Cou	Exactitude	94,44 %
[156]	Accéléromètre	Cou	Précision	92 %

Tableau 5.8: Comparaison des performances de classification, entre notre travail et différents travaux de recherche sur le suivi alimentaire.

critère est que contrairement au travail de [236], nous avons utilisé des barrières entre les mangeoires pour empêcher l'interaction sociale et limiter les mouvements des vaches laitières, ce qui facilite le processus de détection de présence.

Avec la méthode que nous avons proposée, nous avons obtenu une bonne performance de détection du comportement alimentaire de vaches et cela en classant l'état du cou en étant debout (hors période d'alimentation) ou en train de manger (en période d'alimentation). Dans [161], une méthode statistique est utilisée pour discriminer les activités debout et manger. Cette méthode utilise des données d'accéléromètre attaché au cou, le taux d'exactitude obtenu est de 94 %. Dans [156], l'activité alimentaire est suivie en attachant un accéléromètre aux cous des vaches. Les auteurs ont obtenu un taux d'exactitude de 92 %. Les résultats obtenus par notre système sont comparables en termes d'exactitude à ceux obtenus par [161] et en termes de précision à ceux obtenus par [156]. Cependant, l'attachement des accéléromètres [156, 161] sur les vaches produit un stress important. L'utilisation de caméras dans notre travail élimine ce stress et garantit le bien-être des vaches.

Dans [177], les auteurs ont utilisé le dos comme région d'intérêt pour l'identification individuelle de 10 vaches. Ils ont obtenu une exactitude de 99,7 %. Cependant, l'étape de sélection d'images a été effectuée manuellement et la région d'intérêt utilisée ne permet pas de suivre le comportement alimentaire. Dans [180], la méthode proposée a atteint un faible taux d'exactitude. L'échèle de données se limite juste à 12 images de la face des vaches. Cependant, afin de vérifier les performances de la méthode les images ont été transformées en changeant la luminosité, la distorsion, le bruit et les angles de capture des images. Dans [183], des approches d'apprentissage profond ont été testées (VGG-16, ResNet-50, DenseNet-121,

Travail	Animal	Nbr	Images	Entraînement (%)	Emplacement	Méthodes	Exactitude
Notre	vache	17	4875	60	vue de haut	CNN	97,06
[215]	Porc	10	1553	60	Face	CNN	96,7
[177]	Vache	10	1965	85	Dos	SVM	99,7
[237]	vache	500	5000	60	Museau	F-KNN	96,74
[180]	vache	12	12	-	Face	ANMA	80
[183]	vache	105	1433	70	Vue de côté	CNN	96,65

Tableau 5.9: Comparaison des performances de classification entre notre travail et différents travaux de recherche sur l'identification.

AlexNet) pour identifier 105 vaches. La vue de côté a été utilisée comme région d'intérêt. Un taux d'exactitude élevé a été atteint par le réseau Alexnet. Néanmoins, la sélection d'images a été effectuée manuellement. Pour chaque vache, seulement 5 images ont été sélectionnées. De plus, l'utilisation de cette méthode dans le pilotage alimentaire permet d'apporter des informations uniquement sur la présence des vaches dans la mangeoire. Dans [215], un CNN a été déployé pour extraire des caractéristiques et entraîner un modèle permettant l'identification et le suivi de l'abreuvement de 10 porcs. La face des porcs est utilisée comme région d'intérêt. Ce travail montre la grande aptitude des CNN dans l'identification. Cependant, l'emplacement et le mouvement de la tête des animaux est restreint à l'endroit de la source d'eau. Par contre, dans notre travail, la tête de la vache a une plus grande zone de liberté de mouvement pendant la période d'alimentation. Dans [237], plusieurs approches de classification (K-NN, FKNN, arbre de décision, GMM, PNN, MLP et réseaux bayésiens naïfs) ont été évaluées pour identifier des vaches en utilisant le museau comme région d'intérêt, un taux d'exactitude élevé a été obtenu. Cependant, ce résultat doit être vérifié en utilisant des images de museau extraites automatiquement de la scène d'élevage.

La méthode proposée dans ce travail a permis d'atteindre un taux d'exactitude de 97,06 % pour l'identification. Ce taux est plus élevé de : 0,36 % par rapport à [215], 0,32 % par rapport à [237], 17,06 % par rapport à [180] et 0.41% par rapport [183].

Tout au long de cette étude, nous avons développé un système d'identification adapté aux petites fermes. Ce type de fermes est très répandu dans plusieurs pays [238, 239]. Notre système permet de réduire au minimum le temps d'exécution. Cela en réduisant le nombre d'images sélectionnées en utilisant l'AGMM et le SSIM et en réduisant la taille des images à 100 pixels × 100 pixels. Malgré ces restrictions, le système atteint des taux de classification élevés et permet une identification rapide. Cela encourage son utilisation dans les petites fermes. En outre, notre méthode peut être adaptée à un grand nombre de vaches car l'approche 3 permet de répartir les vaches en différents groupes et ainsi l'identification se fera toujours sur un petit nombre de vaches.

L'approche utilisée dans la segmentation basée sur le modèle Xception a obtenu d'excellents résultats. La meilleure performance a été obtenue avec un MIOU de 78%. Une spécificité élevée de 94% a été obtenue, ce qui montre la capacité de la méthode à réduire l'arrière-plan. Une sensibilité de 99% a été aussi obtenue, ce qui reflète la capacité de la méthode à préserver les pixels de l'avant plan. La méthode a préservé tous les pixels de la vache dans 99,8% des images (4875 images). L'unique erreur de segmentation était liée aux faux positifs (les pixels de l'arrière-plan ont été détectés comme étant l'avant-plan). L'exactitude et la précision de la méthode ont été estimées à 96% et 91% respectivement. La méthode proposée a permis de réduire considérablement le nombre de pixels de la zone d'alimentation tout en conservant tous les pixels de vaches comme le montre la Fig 5.12.

Le tableau 5.10 présente une comparaison de nos résultats avec ceux obtenus dans d'autres études [222, 224, 240]. Ces études ont utilisé différentes approches de segmentation pour extraire plusieurs parties du corps des vaches (dos, pattes, etc.). On remarque que les résultats obtenus sont très comparables en termes de précision, de spécificité et de scores F1. Cependant, notre méthode a obtenu un taux de sensibilité plus élevé. Ce taux de sensibilité élevé est dû à l'utilisation des différents coefficients qui favorisent la préservation des pixels de l'avant-plan lors du calcul de la fonction Loss (voir la section 5.2.5).

Globalement, par rapport aux autres travaux, les avantages de la méthode proposée sont :

- (1) Le fonctionnement automatique de l'acquisition jusqu'à l'identification.
- (2) La flexibilité qui se définit par sa capacité à assurer plusieurs classifications avec une même échelle de données.
- (3) Son aspect non stressant (capteur non attaché).

Après la comparaison des résultats, nous estimons le coût de notre système. Comme notre système est composé d'un Raspberry Pi 3B et d'une caméra USB, son prix est de 7400DA par zone d'alimentation puisque le Raspberry Pi 3B coûte 6400DA et la caméra USB 1000DA. Pour réduire ce coût, nous avons testé la capacité du système à gérer plusieurs caméras en même temps. Nous avons constaté qu'avec un taux d'acquisition d'une image par seconde, la Raspberry Pi 3B peut gérer jusqu'à 5 caméras simultanément. Cela permet de réduire le coût à 11400DA pour 5 zones d'alimentation. Par conséquent, environ 2280DA est le prix requis par zone d'alimentation.

## 5.5 Conclusion

Durant ce chapitre, nous avons conçu un système d'identification automatique de vaches laitières et de surveillance de leur comportement alimentaire sans utiliser de capteurs attachés pour améliorer le confort des animaux. La caméra a le potentiel de comprendre la scène de

Étude	Partie du corps des vaches	Modèle	Critère d'évaluation	Performance
<b>Notre</b>	haut de la tête de la vache	Xception	MIOU	78 %
			scores de spécificité moyenne (réduction de fond)	94 %
			scores de sensibilité moyenne (préservation de l'avant plan)	99 %
			scores de précision moyens	91 %
			scores d'exactitude moyens	96 %
			F1 score	93 %
<b>[224]</b>	Détection de l'apparence complète (pâturage)	Mask R-CNN	précision moyenne Scores	94 %
	Détection de la tête (pâturage)			84 %
	Détection de l'apparence complète (mangeoire)			90 %
<b>[222]</b>	Détection de l'apparence complète	DeepMask	Exactitude moyenne	53 %
		SharpMask		82 %
		Mask R-CNN		92 %
<b>[240]</b>	Vue latérale de vache	CNNs	F1 score	71 %
	Patte			59 %

Tableau 5.10: Comparaison des performances de la segmentation de notre méthode et celles des travaux de recherche existants.

l'élevage et d'offrir des informations en temps réel. Plusieurs méthodes de classification et de segmentation d'images utilisant l'apprentissage profond ont été testées. Les résultats de ces tests ont montré que :

1. L'utilisation de réseaux de neurones à convolution est très pratique dans des environnements complexes tels que les fermes. De bons résultats ont été obtenus malgré la grande dissimilitude présente entre les images d'entraînement et de test.
2. L'utilisation de DeepLab comme approche de segmentation sémantique a donné un grand avantage en termes de rapidité d'exécution et de sensibilité. Ce résultat montre que la méthode proposée peut améliorer les performances et la praticabilité de l'identification.
3. Les résultats obtenus sont intéressants. En effet, pour la détection de la disponibilité de la nourriture et la détermination de son type l'exactitude est de 100%, pour la classification des états debout et en train de manger l'exactitude est de 92 % et enfin pour l'identification l'exactitude est de 97 %. Ces résultats pourraient fournir une base précieuse pour la surveillance automatique des maladies, la reproduction et la gestion de l'alimentation des vaches laitières.
4. Parmi les trois approches utilisées pour identifier les vaches laitières, l'approche de classification par multiples CNN a atteint la meilleure exactitude. Bien que notre système a été testé dans une petite ferme, cette approche peut être utilisée dans une grande ferme en séparant les vaches en plusieurs groupes (en fonction de la couleur des vaches, de la présence de corne, etc.). Dans ce sens, nous envisageons d'utiliser cette approche pour l'identification individuelle dans une grande ferme où l'identification sera faite séparément dans chaque groupe.
5. Les caméras ont le potentiel de comprendre la scène d'élevage, elles permettent d'offrir de nouvelles informations telles que la reconnaissance des aliments distribués. De plus, d'autres tâches telles que la mesure de la quantité d'aliment présent dans une mangeoire ou l'évaluation de la répartition des aliments entre les différentes mangeoires peuvent être réalisées par les caméras. Les informations offertes peuvent apporter des solutions pour l'alimentation de précision ou d'élevage de précision et permettent une meilleure gestion du troupeau et de la ferme.

## Conclusion générale

La compréhension du comportement de vaches laitières et sa modélisation est une étape importante dans le développement de systèmes de monitoring comportemental. Ces systèmes doivent contribuer à améliorer et à favoriser les processus de surveillance sanitaire, du bien être, de reproduction, etc. Ces systèmes de monitoring permettent de détecter des changements subtils avant même que des signes cliniques évidents ne s'affichent. Cela, incite à intégrer de nouveaux capteurs à l'échelle du troupeau, d'un ensemble d'animaux ou d'un animal sur une ou plusieurs parties de son corps (oreilles, dos, pied, etc.) pour la collecte de données.

Dans cette thèse, après avoir proposé un état de l'art sur l'élevage de précision, les maladies des vaches laitières et leur impact sur le comportement de ces dernières, nous avons proposé deux systèmes de suivi comportemental (activité et alimentaire).

Pour le premier, nous avons tout d'abord développé un système embarqué doté, entre autre, d'une unité de mesure inertielle (IMU) permettant l'acquisition de données. Ce système est ensuite déployé dans la ferme commerciale SOFLAIT de Draa Ben Khedda pour collecter des données comportementales des vaches (activité). Ces données acquises sont, par la suite, analysées et classées en sept différentes classes où chaque classe représente un comportement (debout, couchée vers le côté droit, couchée vers le côté gauche, se lever, se coucher, marcher, stationnaire). Le modèle de classification conçu est basé sur des modèles de mélange fini et des arbres de décision. Les transitions entre les comportements debout et couchée sont validées par des contraintes imposées par un automate d'états finis déterministes. Pour optimiser les résultats de classification, nous avons utilisé une approche de fusion de données. Les résultats obtenus [35] sont intéressants, notamment en termes d'exactitude qui est de 99% et surpassent les résultats obtenus par les travaux de recherche existants [27–29, 156, 171, 172, 201].

Les capteurs utilisés dans le monitoring des activités ont une grande aptitude à classer les comportements liés directement à leurs emplacements (oreille, pied, cou, etc). Cependant, ces emplacements restent limités en termes de surveillance de certains autres phénomènes comme les contractions musculaires, le rythme cardiaque, la température des mamelles, l'acceptation de chevauchement et la mesure de la circonférence. Dans notre cas, à l'aide d'une ceinture, nous avons choisi d'attacher le système d'acquisition de données au dos de la vache. À travers les résultats obtenus nous pouvons conclure que le dos est un emplacement stable et adéquat pour suivre les activités des vaches. En plus, cet emplacement permet d'intégrer plus de capteurs pour suivre plus d'événements sanitaires, reproductifs et morphologiques.

Pour le deuxième, nous avons utilisé un système embarqué doté de caméras, développé et déployé pour l'acquisition et la sélection d'images en temps réel. Ensuite, à l'aide d'un ensemble de méthodes de classification d'images telles que les réseaux de neurones à convolution (CNN) et de segmentation d'images (Xception), notre système permet la détection de

la présence de vaches en zone d'alimentation, la classification de leur comportement alimentaire, l'identification individuelle et la reconnaissance de la disponibilité et le type d'aliment présent dans la mangeoire. Les résultats de ce travail [241], en termes de classification du comportement alimentaire et de l'identification individuelle, comparés à ceux présentés par d'autres travaux [156, 161, 177, 180, 183, 236, 237, 242] sont très intéressants.

Nous tenons à souligner que la plupart des travaux de recherche portant sur l'identification des animaux et sur le suivi alimentaire utilisent des capteurs attachés (RFID, accéléromètre). Cependant, ces derniers causent du stress aux animaux et encourent le risque de se détacher. Par conséquent, nous pensons qu'il est plus intéressant d'utiliser des capteurs non attachés comme les caméras à chaque fois que c'est possible. Donc, les capteurs attachés seront utilisés seulement dans des situations où les capteurs non attachés ne le permettent pas.

En guise de perspectives à court terme, nous envisageons :

- D'ajouter la détection d'autres comportements tels que le toilettage, les interactions sociales et les chevauchements (les chaleurs) à notre système de monitoring d'activité.
- De réduire la taille du capteur du premier système afin de pouvoir le placer sous la peau des vaches, éliminant ainsi le stress et le risque de détachement.
- D'ajouter au système de monitoring du comportement alimentaire un mécanisme permettant de calculer de nouvelles variables telles que la quantité d'aliments présents dans la mangeoire et la quantité d'aliments ingérés par les vaches.
- De mettre à l'échelle le système de monitoring du comportement alimentaire dans le contexte d'un plus grand nombre de vaches.

En guise de perspective à long terme, nous envisageons:

- D'intégrer les deux systèmes que nous avons développés dans les outils de suivi sanitaire et cela en menant des études permettant d'établir des corrélations entre les comportements suivis et les différentes maladies.



# Bibliographie

- [1] A. haring, organic dairy farms in the eu: Production systems, economics and future development, *livestock production science*, vol. 80, no. 1-2, pp. 89-97, 2003. doi: 10.1016/s0301-6226(02)00308-1.
- [2] Editorial: Food security, complex emergencies and longer-term programming, disasters, vol. 29, pp. s1-s4, 2005. doi: 10.1111/j.0361-3666.2005.00281.x.
- [3] C. desire, l'élevage de précision : changements dans l'organisation du travail et la gestion de données dans des exploitations laitières, porcines et avicoles ,inra, 2015.
- [4] M michel, t muriel et h nathalie, élevage et pâturage de précision : l'animal sous surveillance électronique, *courrier de l'environnement de l'inra*, 2013. 63:13–24, 08.
- [5] Bewley j.m. et russell r.a. reasons for slow adoption rates of precision dairy farming technologies: evidence from a producer survey. in: *Proceedings of the first north american conference on precision dairy management, toronto, canada, 30-31. 2010.*
- [6] Berckmans d. precision livestock farming technologies for welfare management in intensive livestock systems. *rev. sci. tech.* 33, 189-196. 2014.
- [7] I. halachmi, precision livestock farming applications making sense of sensors to support farm management, wageningen academic publishers, 2015.
- [8] Florence kling eveillard, automatisation et relation homme animal, 6ème séminaire sur les enjeux de l'èlevage pour les doctorants de l'inra, 2017.
- [9] D. berckmans, automatic on-line monitoring of animals by precision livestock farming catholic university leuven, laboratory for agricultural buildings research, kasteelpark arenberg 30, b-3001 leuven, belgium,2004.
- [10] Berckmans, d. general introduction to precision livestock farming. *animal frontiers*,2017. doi: 10.2527/af.2017.0102.

- [11] C. j. rутten, a. g. j. velthuis, w. steeneveld et h. hogeveen, sensors to support health management on dairy farms, department of farm animal health, faculty of veterinary medicine, utrecht university, wageningen ,2013.
- [12] D. berckmans, value creation through precision livestock farming systems , model and manage bio responses, ku leuven catholic university leuven, belgium,2015.
- [13] A. helwatkar, d. riordan et j. walsh., sensor technology for animal health monitoring ,imar technology gateway institute of technology ,tralee, ireland ,2014.
- [14] Awasthi, a., awasthi, a., riordan, d. et walsh, j. non-invasive sensor technology for the development of a dairy cattle health monitoring system. computers, 2016.
- [15] T. f. mottram, h. r. whay, s. g. vass et b. l. nielsen., apparatus for animal health monitoring , us patent, 2001.
- [16] S. viazzi, c. bahr, a. schlageter tello, t. v. hersem, c. e. b. romanini, a. pluk, i. halachmi, c. lokhorst et d. berckmans, analysis of individual classification of lameness using automatic, american dairy science association, 2013 .
- [17] T. godsk et m. b. kjaergaard, high classification rates for continuous cow activity recognition using low cost gps positioning sensors and standard machine learning techniques. advances in data mining. applications and theoretical aspects, 2011.
- [18] Georg h, ude g, schwalm a, wenderdel b. investigation on temperature sensing injectable transponders for electronic animal identification an evaluation of suitable injection sites with bull calves. landbauforschung voelkenrode , 2009.
- [19] L. ruiz-garcia et l. lunadei, the role of rfid in agriculture: Applications, limitations and challenges, computers and electronics in agriculture, vol. 79, no. 1, pp. 42-50, 2011.
- [20] T. pizzuti et g. mirabelli, future technology in tracing animals on the food chain, woodhead publishing series in food science, technology and nutrition, advances in food traceability techniques and technologies, woodhead publishing, 2016.
- [21] A. awad, from classical methods to animal biometrics: A review on cattle identification and tracking, computers and electronics in agriculture, vol. 123, pp. 423-435, 2016.
- [22] W. hong-da, application of radio frequency identification (rfid) in dairy information management, journal of northeast agricultural university (english edition), vol. 19, no. 1, pp. 78-81, 2012.

- [23] N hostiou, c allain, s chauvat, a turlot et j fagon, conséquences de l'élevage de précision sur le travail des éleveurs. inra, 2014.
- [24] J. rushen, n. chapinal, a. de passillé. automated monitoring of behavioural-based animal welfare indicators. *anim. welf.*, 2012.
- [25] U. braun, t. tschoner. evaluation of eating and rumination behaviour in cows using a noseband pressure sensor, 2014.
- [26] K. yoda, k. sato, y. niizuma, m. kurita, c. bost, l. maho, et y. naito, precise monitoring of porpoising behaviour of adélie penguins determined using acceleration data loggers., 1999.
- [27] P. martiskainen, m. jarvinen, j.p. skonb, j. tiirikainen, m. kolehmainen, et j. mononen, cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines, 2009.
- [28] B. roberta, b. white, d. renter, et r. larson. evaluation of three-dimensional accelerometers to monitor and classify behavior patterns in cattle, 2009.
- [29] Vazquez diosdado, j., barker, z., hodes, h., amory, j., croft, d., bell, n. and codling, e. classification of behaviour in housed dairy cows using an accelerometer-based activity monitoring system. *animal biotelemetry*, 2015.
- [30] B. wolfger, e. timsit, e. pajor, n. cook, h. barkema, et k. orsel, technical note: Accuracy of an ear tag-attached accelerometer to monitor rumination and feeding behavior in feedlot cattle, 2015.
- [31] W. clapham, j. feddersa, k. beeman, et j. neela, acoustic monitoring system to quantify ingestive behavior of free-grazing cattle, 2011.
- [32] Ito k, von keyserlingk mag, leblanc sj et weary dm. lying behavior as an indicator of lameness in dairy cows. *journal of dairy science*, 2010. doi: <http://dx.doi.org/10.3168/jds.2009-2951>.
- [33] R.j. kilgour. a review of the behavior of cattle at pasture. *appl. anim. behav. sci.*, 2012. doi: <http://dx.doi.org/10.1016/j.applanim.2011.12.002>.
- [34] Ringgenberg n, bergeron r et devillers n. validation of accelerometers to automatically record sow postures and stepping behaviour. *applied animal behaviour science*, 2010. doi: <http://dx.doi.org/10.1016/j.applanim.2010.09.018>.

- [35] Brahim achour, malika belkadi, rachida aoudjit, mourad laghrouche, unsupervised automated monitoring of dairy cows behavior based on inertial measurement unit attached to their back, computers and electronics in agriculture, volume 167, 2019. doi: <https://doi.org/10.1016/j.compag.2019.105068>.
- [36] Dutta, r., smith, d., rawnsley, r., bishop-hurley, g., hills, j., timms, g., et henry, d. dynamic cattle behavioural classification using supervised ensemble classifiers. computers and electronics in agriculture, 2015.
- [37] Williams, m.l., james, w.p., rose, m.t. fixed time data segmentation and behavior classification of pasture-based cattle: Enhancing performance using a hidden markov model. j. comput. electron. agric, 2017.
- [38] Gonzalez , l., bishop hurley, g., handcock, r., et crossman, c. behavioral classification of data from collars containing motion sensors in grazing cattle. computers and electronics in agriculture, 2015.
- [39] Eckelkamp, e. invited review: Current state of wearable precision dairy technologies in disease detection. applied animal science, 2019. doi: 10.15232/aas.2018-01801.
- [40] M. madeleine et e. pottier, suivi automatisé de la santé et du bien-être en élevage bovin laitier en france : état des lieux et perspectives, thèse pour le doctorat vétérinaire, la faculte de medecine de creteil janvier 2016.
- [41] M. saint dizier et s. chasant maillard, methods and on farm devices to predict calving time in cattle. the veterinary journal, 2015.
- [42] Rapport: Détection des vèlages , fiche santé viande bovine. 2013.
- [43] A. chanvallon et s. coyrat castel, comparison of three devices for the automated detection of estrus in dairy cows, theriogenology, 2014.
- [44] D. streyl, c. sauter louis et h. zerbe. establishment of a standard operating procedure for predicting the time of calving in cattle. journal of veterinary science, 2010.
- [45] A. maltz. a practical way to detect approaching calving of the dairy cow by a behavior sensor. precision livestock farming , 2007.
- [46] R. borchers, y. m. chang et j. bewley. predicting impending calving using automatic activity and rumination measures in dairy cattle. adsa-asas-csas joint annual meeting, 2014.

- [47] J. cooper prado et n. m. long. relationship of ruminal temperature with parturition and estrus of beef cows. *journal of animal science*, 2010.
- [48] Surveillance automatisée des vêlages en élevage bovin laitier. institut de l'élevage, 2013.
- [49] M. hanna, m. m. michael et s. cockramb. changes in the behaviour of dairy cows during the 24 h before normal calving compared with behaviour during late pregnancy. *applied animal behaviour science*, 2011.
- [50] S. reith et s. hoy. relationship between daily rumination time and estrus of dairy cows , *journal of dairy science*, 2012.
- [51] C. pahl, e. hartung. a. grothmann, k. mahlkow nerger et a. haeussermann. rumination activity of dairy cows in the 24 hours before and after calving. *journal of dairy science*, 2014.
- [52] M.titler, m.g.maquivar, s.bas, .j.rajala schultz, e.gordon, k.mccullough, p.federico et g.m.schuenemann . prediction of parturition in holstein dairy cattle using electronic data loggers. *journal of dairy science*, 2015.
- [53] J. roelofs, f. lopez gatiusc, r. hunterd, f. v. eerdenburge et a. c. hanzenf,. when is a cow in estrus? clinical and practical aspects. *theriogenology*, 2010.
- [54] R. jonsson, detection of oestrus and lameness in dairy cows. phd thesis, technical university of denmark, 2011.
- [55] R. firsk et e. stamer, automation of oestrus detection in dairy cows: a review. *livestock production science*, volume 75, issue 3, 2002.
- [56] Détection automatisée des chaleurs en élevage bovin laitier : quel outil choisir ?, institut de lelevage, 2012.
- [57] M. saint dizier et s. chasant maillard, review article : Towards an automated detection of oestrus in dairy cattle. *reproduction in domestic animals*, 2012.
- [58] F. mazeris. delaval herd navigator proactive herd management. the first north american conference on precision dairy management, 2010.
- [59] J. delage, j. poly et b. vissac. étude de l'efficacité relative des diverses formules de barymétrie applicables aux bovins. *Annales de zootechnie, inra/edp sciences*, 1955.

- [60] G. bretschneder, a. cuatrin, d. arias et d. vottero. estimation of body weight by an indirect measurement method in developing replacement holstein heifers raised on pasture. *archivos de medicina veterinaria*, 2014.
- [61] S. i. s. tasdemir et a. urkmez, .a fuzzy rule based system for predicting the live weight of holstein cows whose body dimensions were determined by image analysis. *turkish journal of electrical engineering and computer sciences*, 2011.
- [62] I. s. tasdemira et a. urkmez b . determination of body measurements on the holstein cows using digital image analysis and estimation of live weight with regression analysis. *computers and electronics in agriculture* volume 76, issue 2, may 2011.
- [63] S. ozkaya. the prediction of live weight from body measurements on female holstein calves by digital image analysis. *the journal of agricultural science*, 2012.
- [64] J. mekpariyup, k. saithanu, et n. arunkeeree, estimation of body weight of holstein-friesian cattle with multiple regression analysis, 2013.
- [65] C gerardo, j.j, ghirardi, h marta, m maria et f bocquier. utilisation des bolus électroniques pour la traçabilité des ruminants : état de la technique, mise en place et évaluation en ovins et bovins. *rencontres recherches ruminants*, 2006.
- [66] Gooden, j. m. beach, et a. d. purchas, measurement of sub- cutaneous backfat depth in live lambs with an ultrasonic probe, 1980.
- [67] V. fisher, a review of the technique of estimating the composition of livestock using the velocity of ultrasound, 1997.
- [68] Mirielle theriault, utilisation des mesures aux ultrasons pour estimer la composition de la carcasse chez l'agneau lourd, université laval québec, 2005.
- [69] G. azzaro, m. caccamo, j. ferguson, s. battiato, g. farinella, g. guarnera, g. puglisi, r. petriglieri, et g. licitra, objective estimation of body condition score by modeling cow body shape from digital images, 2011.
- [70] J. bewley, a. peacock, o. lewis, r. boyce, d. roberts, m. coffey, s. kenyon, et m. schutz, potential for estimation of body condition scores in dairy cattle, 2008.
- [71] S. battiato, g. m. farinella, g. c. guarnera, g. puglisi, g. azzaro, m. caccamo, g. licitra, et j. d. ferguson, estimation of cows body condition score from images, 2010.

- [72] Marilyn krukowski, automatic determination of body condition score of dairy cows from 3d images. thèse de master, stockholm, sweden, 2009.
- [73] T banhazi, h lehr, j black, h crabtree, c schofield, m tscharke, d berckmans. precision livestock farming: An international review of scientific and commercial aspects. journal of agricultural and biological engineering. 2012.
- [74] C kamphuis ,w steeneveld, h hogeveen. economic modelling to evaluate the benefits of precision livestock farming technologies. precision livestock farming applications 2015. dio: 10.3920/978-90-8686-815-59.
- [75] Daniela lovarelli, jacopo bacenetti, marcella guarino, a review on dairy cattle farming: Is precision livestock farming the compromise for an environmental, economic and social sustainable production, journal of cleaner production, volume 262,2020.
- [76] Van asseldonk. economic evaluation of information technology applications on dairy farms. wageningen agricultural university, wageningen, the netherlands, 1999.
- [77] Banhazi t m, lewis b, tscharke m. the development and commercialisation aspects of a practical feed intake measurement instrumentation to be used in livestock buildings. australian journal of multi-disciplinary engineering, 2011.
- [78] hostiou, n., allain, c., chauvat, s., turlot, a., pineau, c., et fagon, j. l'élevage de précision : quelles conséquences pour le travail des éleveurs ?. inrae productions animales, 2014.
- [79] S. shields, g. orme-evans the impacts of climate change mitigation strategies on animal welfare. animals, 2015.
- [80] G. zhang, p. kai, c. zong. precision ventilation to improve indoor air quality and reduce ammonia emission of pig production housing, precision livestock farming, the 6th european conference on precision livestock farming, 2013.
- [81] E. vranken, j.-f. cabaraux, d. berckmans. development of a simplified measuring procedure for ammonia emission from agricultural constructions. precision livestock farming, 2013.
- [82] T. bartzanas, d.k. papanastasiou, i. skoufos, a. tzora, d. fidaros, c. baxevanou, et al. environmental friendly poultry production using novel precision livestock techniques. precision livestock farming 2015.
- [83] Y. blaise, a. andriamandroso, b. heinesch, y. beckers, e. castro munoz, f. lebeau, et al. influences of feeding behaviour and forage quality on diurnal methane emission dynamics of grazing cows. precision livestock farming, 2017.

- [84] White, r.r., capper, j.l., precision diet formulation to improve performance and profitability across various climates: modeling the implications of increasing the formulation frequency of dairy cattle diets. *j. dairy sci.* 2014.
- [85] Emanuela tullo, alberto finzi, marcella guarino, review: Environmental impact of livestock farming and precision livestock farming as a mitigation strategy, *science of the total environment*, volume 650, part 2, 2019.
- [86] P. llonch, m. haskell, r. dewhurst, s. turner. current available strategies to mitigate greenhouse gas emissions in livestock systems: an animal welfare perspective. *animal*, 2017.
- [87] M.t.m. king, s.j. leblanc, e.a. pajor, t.c. wright, t.j. devries. behavior and productivity of cows milked in automated systems before diagnosis of health disorders in early lactation. *j. dairy sci.*, 2018.
- [88] M.j. bell, p.c. garnsworthy, a.w. stott, j.e. pryce. effects of changing cow production and fitness traits on profit and greenhouse gas emissions of uk dairy systems *j. agric. sci.* 2014.
- [89] P.j. gerber, b. henderson, h.p. makkar. mitigation of greenhouse gas emissions in livestock production a review of technical options for non-co2 emissions, *fao*, 2013.
- [90] W. chen, e. white, n.m. holden. the effect of lameness on the environmental performance of milk production by rotational grazing. *j. environ. manag.*, pp. 143-150. 2016.
- [91] A. hospido, u. sonesson. the environmental impact of mastitis: a case study of dairy herds *sci. total environ.*, pp. 71-82. 2005.
- [92] J. spilke et r. fahr, decision support under the conditions of automatic milking systems using mixed linear models as part of a precision dairy farming concept, *efita*, 2003.
- [93] belkheir b. ghozlane, benidir m. benahmed n. bousbia a. dépistage de mammites subcliniques chez des vaches en zone montagnaise de tizi-ouzou (algérie). *renc. rech. ruminants*, 2016.
- [94] belkheir b. ghozlane f. benidir m. bousbia a. benahmed n. agguini s. typologie du lait de mélange des exploitations bovines en zone de montagne de tizi- ouzou (algérie), *renc. rech. ruminants*, 2014, 21.
- [95] L. gonzalez, b. tolkamp, m. coffey, a. ferret et i. kyriazakis, changes in feeding behavior as possible indicators for the automatic monitoring of health disorders in dairy cows, *journal of dairy science*, vol. 91, no. 3, pp. 1017 1028, 2008.

- [96] J.m. huzzey, d.m. veira, d.m. weary, m.a.g. von keyserlingk. prepartum behavior and dry matter intake identify dairy cows at risk for metritis, *journal of dairy science*, volume 90, issue 7, 2007.
- [97] M. stangaferro, r. wijma, l. caixeta, m. al abri and j. giordano, use of rumination and activity monitoring for the identification of dairy cows with health disorders: Part i. metabolic and digestive disorders, *journal of dairy science*, vol. 99. 2016.
- [98] T. halasa, k. huijps, et h. hogeveen, economic effects of bovine mastitis and mastitis management: A review, *veterinary quarterly*, vol. 29, no. 1, pp. 18 31, 2007. dio: 10.1080/01652176.2007.9695224.
- [99] R. bicalho et g. oikonomou, control and prevention of lameness associated with claw lesions in dairy cows, *livestock science*, vol. 156, no. 1 3, pp. 96 105, 2013. dio: 10.1016/j.livsci.2013.06.007.
- [100] S. walker, r. smith, j. routly, d. jones, m. morris and h. dobson, lameness, activity time budgets, and estrus expression in dairy cattle, *journal of dairy science*, vol. 91, no. 12, pp. 4552 4559, 2008. dio: 10.3168/jds.2008 1048.
- [101] A. van nuffel et al. lameness detection in dairy cows: Part 1. how to distinguish between non lame and lame cows based on differences in locomotion or behavior, *animals*, vol. 5, no. 3, pp. 838 860, 2015. dio: 10.3390/ani5030387.
- [102] A. van nuffel et al. lameness detection in dairy cows: Part 2. use of sensors to automatically register changes in locomotion or behavior, *animals*, vol. 5, no. 3, pp. 861 885, 2015. dio: 10.3390/ani5030388.
- [103] S. leblanc, monitoring metabolic health of dairy cattle in the transition period, *journal of reproduction and development*, vol. 56, no. pp. s29 s35, 2010. dio: 10.1262/jrd.1056s29.
- [104] K. fogsgaard, c. rantved, p. sarensen et m. herskin, sickness behavior in dairy cows during *escherichia coli* mastitis, *journal of dairy science*, vol. 95, no. 2, pp. 630 638, 2012.
- [105] J. siivonen et al. impact of acute clinical mastitis on cow behaviour, *applied animal behaviour science*, vol. 132, no. 3 4, pp. 101 106, 2011. dio: 10.1016/j.applanim.2011.04.005.
- [106] A. schlageter tello et al. manual and automatic locomotion scoring systems in dairy cows: A review, *preventive veterinary medicine*, vol. 116, no. 1 2, pp. 12 25, 2014. dio: 10.1016/j.prevetmed.2014.06.006.

- [107] V. thorup, l. munksgaard, p. robert, h. erhard, p. thomsen et n. friggens, lameness detection via leg mounted accelerometers on dairy cows on four commercial farms, *animal*, vol. 9, no. 10, pp. 1704 1712, 2015. dio: 10.1017/s1751731115000890.
- [108] H. whay, locomotion scoring et lameness detection in dairy cattle, in *practice*, vol. 24, no. 8, pp. 444 449, 2002. dio: 10.1136/inpract.24.8.444.
- [109] N. bareille, f. beaudeau, s. billon, a. robert and p. faverdin, effects of health disorders on feed intake and milk production in dairy cows, *livestock production science*, vol. 83, no. 1, pp. 53 62, 2003. dio: 10.1016/s0301 6226(03)00040 x.
- [110] V. thorup et al. lameness affects cow feeding but not rumination behavior as characterized from sensor data, *frontiers in veterinary science*, vol. 3, 2016. dio: 10.3389/fvets.2016.00037.
- [111] K. fogsgaard, t. bennedsgaard et m. herskin, behavioral changes in freestall housed dairy cows with naturally occurring clinical mastitis, *journal of dairy science*, vol. 98, no. 3, pp. 1730 1738, 2015. dio: 10.3168/jds.2014 8347.
- [112] R.m. de mol , j. van dijk, m.h. troost, a. sterk, r. jorritsma et p.h. hogewerf , early detection of metabolic disorders in dairy cows by using sensor data , *precision livestock farming applications* , 2015.
- [113] A. itle, j. huzzey, d. weary et m. von keyserlingk, clinical ketosis and standing behavior in transition cows, *journal of dairy science*, vol. 98, no. 1, pp. 128 134, 2015. dio: 10.3168/jds.2014 7932.
- [114] S. garcia, s. ramirez gallego, j. luengo, j. benitez et f. herrera, big data preprocessing: methods and prospects, *big data analytics*, vol. 1, no. 1, 2016. dio: 10.1186/s41044-016-0014-0.
- [115] O. maimon et l. rokach, *data mining and knowledge discovery handbook*. new york: Springer, 2010.
- [116] A. kamilaris, a. kartakoullis et f. prenafeta-boldu, a review on the practice of big data analysis in agriculture, *computers and electronics in agriculture*, vol. 143, pp. 23-37, 2017. dio: 10.1016/j.compag.2017.09.037.
- [117] J. han, m. kamber et j. pei, *data mining*. waltham, ma: Morgan kaufmann/elsevier, 2012.
- [118] D. lindsay et d. pearce, *reproduction in sheep*. cambridge: Cambridge university press, 2011.
- [119] S. garcia, *data preprocessing in data mining*. springer international pu, 2016.

- [120] T. gueta et y. carmel, integrating data-cleaning with data analysis to enhance usability of bio-diversity big-data, proceedings of tdwg, vol. 1, p. e20244, 2017. dio: 10.3897/tdwgproceedings.1.20244.
- [121] H. cheng, d. feng, x. shi et c. chen, data quality analysis and cleaning strategy for wireless sensor networks, eurasip journal on wireless communications and networking, vol. 2018, no. 1, 2018. dio: 10.1186/s13638-018-1069-6.
- [122] Uddin, muhammad fahim, noise. removal and structured data detection to improve search for personality features, proceedings of the 2016 ieee/acm international conference on advances in social networks analysis and mining , 2016.
- [123] B. arputhamary et l. arockiam, data integration in big data environment, bonfring international journal of data mining, vol. 5, no. 1, pp. 01-05, 2015. dio: 10.9756/bijdm.8001.
- [124] Y. peng, y. zhang, y. tang et s. li, an incident information management framework based on data integration, data mining, and multi-criteria decision making, decision support systems, vol. 51, no. 2, pp. 316-327, 2011. dio: 10.1016/j.dss.2010.
- [125] X. wang, f. wu et l. xing, recognition and reduction of traffic flow redundant data, journal of computer applications, vol. 29, no. 4, pp. 1110-1113, 2009. dio: 10.3724/sp.j.1087.2009.01110.
- [126] S. ramirez gallego, b. krawczyk, s. garcia, m. woiniak et f. herrera, a survey on data preprocessing for data stream mining: Current status and future directions, neurocomputing, vol. 239, pp. 39-57, 2017. dio: 10.1016/j.neucom.2017.01.078.
- [127] A. azar and a. hassanien, dimensionality reduction of medical big data using neural-fuzzy classifier, soft computing, vol. 19, no. 4, pp. 1115-1127, 2014. dio: 10.1007/s00500-014-1327-4.
- [128] Hamidzadeh, r. monsefi et h. sadoghi yazdi, irahc: Instance reduction algorithm using hyperrectangle clustering, pattern recognition, vol. 48, no. 5, pp. 1878-1889, 2015. dio: 10.1016/j.patcog.2014.11.005.
- [129] N. vafaei, r. ribeiro et l. matos, data normalisation techniques in decision making: case study with topsis method, international journal of information and decision sciences, vol. 10, no. 1, p. 19, 2018. dio: 10.1504/ijids.2018.090667.
- [130] S. gowtham et s. karuppusamy, review of data mining classification techniques, bonfring international journal of software engineering and soft computing, vol. 9, no. 2, pp. 8-11, 2019. dio: 10.9756/bijsesc.9013.

- [131] Sagar s. nikam, a comparative study of classification techniques in data mining algorithms, international journal of modern trends in engineering and research, vol. 4, no. 7, pp. 58-63, 2017.
- [132] A saxena, mukesh prasad, akshansh gupta, neha bharill, om prakash patel, aruna tiwari, meng joo er, weiping ding, chin-teng lin, a review of clustering techniques and developments. neurocomputing. dio: 10.1016/j.neucom.2017.06.053.
- [133] R. rubina and p. verma, various techniques of clustering: A review, iosr journal of computer engineering, vol. 18, no. 05, pp. 23-28, 2016. dio: 10.9790/0661-1805032328.
- [134] N. sarika and dr.v. viji vinod, survey based on clustering techniques for healthcare system, journal of advanced research in dynamical and control systems, vol. 11, no. 0009-, pp. 978-981, 2019. dio: 10.5373/jardcs/v11/20192659.
- [135] H. izakian and a. abraham, fuzzy c-means and fuzzy swarm for fuzzy clustering problem, expert systems with applications, vol. 38, no. 3, pp. 1835-1838, 2011. dio: 10.1016/j.eswa.2010.07.112.
- [136] W. wang, linear mixed function-on-function regression models, biometrics, vol. 70, no. 4, pp. 794-801, 2014. dio: 10.1111/biom.12207.
- [137] Peter c. austin, ewout w. steyerberg, the number of subjects per variable required in linear regression analyses, journal of clinical epidemiology, volume 68, issue 6, ,pages 627-636, 2015.
- [138] M. soysal and e. schmidt, machine learning algorithms for accurate flow-based network traffic classification: Evaluation and comparison, performance evaluation, vol. 67, no. 6, pp. 451-467, 2010. dio: 10.1016/j.peva.2010.01.001.
- [139] A munther, o rozmie razif , a mosleh, amohammed, y shahrul. a preliminary performance evaluation of k-means, knn and em unsupervised machine learning methods for network flow classification. . international journal of electrical and computer engineering (ijece). 6. 778-784.2016. dio:10.11591/ijece.v6i1.8909.
- [140] Hossin, mohammad et m.n, sulaiman. a review on evaluation metrics for data classification evaluations. international journal of data mining and knowledge management process. (2015).
- [141] Perlovsky, leonid, neural networks for sensor fusion and adaptive classification, neural networks, vol. 1, p. 42, 1988. dio: 10.1016/0893-6080(88)90084-6.
- [142] S. sra, s. nowozin et s. wright, optimization for machine learning, neural information processing series, 2016.

- [143] M. kumar et m. d.c, a survey on improving classification performance using data pre processing and machine learning methods on nsl-kdd data, international journal of engineering and computer science, 2016. dio: 10.18535/ijecs/v5i4.17.
- [144] Henrik brink, joseph richards, et mark fetherolf. 2016. real-world machine learning (1st ed.). manning publications co., greenwich, ct, usa. 2016.
- [145] T. mitchell, machine learning. new york: Mcgraw hill, 2017.
- [146] D. smadja et al., detection of subclinical keratoconus using an automated decision tree classification, american journal of ophthalmology, vol. 156, no. 2, pp. 237-246.e1, 2013. dio: 10.1016/j.ajo.2013.03.034.
- [147] D. bhukya et s. ramachandram, decision tree induction: An approach for data classification using avl-tree, international journal of computer and electrical engineering, pp. 660-665, 2010. dio: 10.7763/ijcee.2010.v2.208.
- [148] Filippo amato, alberto lopez, eladia maria p mindez, petr vahara, aleia hampl, josef havel. artificial neural networks in medical diagnosis, journal of applied biomedicine, volume 11, issue 2, pages 47-58. 2013.
- [149] Tatt hee oong et n. isa, adaptive evolutionary artificial neural networks for pattern classification, ieee transactions on neural networks, vol. 22, no. 11, pp. 1823-1836, 2011. available: 10.1109/tnn.2011.2169426.
- [150] F. narter, diagnostic classification of cystoscopic images using deep convolutional neural networks, journal of urological surgery, vol. 6, no. 3, pp. 264-264, 2019. dio 10.4274/jus.galenos.2019.06.017.
- [151] W. yanling and w. jixia, the em algorithm for the finite mixture of exponential distribution models, international journal of contemporary mathematical sciences, vol. 9, pp. 57-64, 2014. dio: 10.12988/ijcms.2014.312133.
- [152] Reynolds douglas , gaussian mixture models , encyclopedia of biometrics, springer us,2015.
- [153] P. nie, l. zang et l. liu, application of multi class classification algorithm based on twin support vector machine in intrusion detection, journal of computer applications, vol. 33, no. 2, pp. 426-429, 2013. available: 10.3724/sp.j.1087.2013.00426.
- [154] Jan luts, fabian ojeda, raf van de plas, bart de moor, sabine van huffel, johan a.k. suykens, a tutorial on support vector machine based methods for classification problems in chemometrics, analytica chimica acta, volume 665, issue 2,2010.

- [155] Shen yin, jiapeng yin, tuning kernel parameters for svm based on expected square distance ratio, information sciences, volumes 370, 2016.
- [156] Said benaissa, frank a.m. tuytens, david plets, toon de pessemier, jens trogh, emmeric tanghe, luc martens, leen vandaele, annelies van nuffel, wout joseph, bart sonck . on the use of on cow accelerometers for the classification of behaviours in dairy barns. research in veterinary science, 2017.
- [157] Shen weizheng, cheng fei, zhang yu, wei xiaoli, fu qiang, zhang yonggen, automatic recognition of ingestive related behaviors of dairy cows based on triaxial acceleration, information processing in agriculture, 2019.
- [158] M. williams, n. mac parthalin, p. brewer, w. james et m. rose. a novel behavioral model of the pasture based dairy cow from gps data using data mining and machine learning techniques, journal of dairy science, vol. 99, no. 3, pp. 2063 2075, 2016.
- [159] Y. guo, g. poulton, p. corke, g. bishop hurley, t. wark et d. swain. using accelerometer, high sample rate gps and magnetometer data to develop a cattle movement and behaviour model, ecological modelling, vol. 220, no. 17, pp. 2068 2075, 2009.
- [160] A. andriamandroso et al., development of an open source algorithm based on inertial measurement units (imu) of a smartphone to detect cattle grass intake and ruminating behaviors, computers and electronics in agriculture, vol. 139, pp. 126 137, 2017.
- [161] C. arcidiacono, s. porto, m. mancino et g. cascone, development of a threshold based classifier for real time recognition of cow feeding and standing behavioural activities from accelerometer data, computers and electronics in agriculture, 2017.
- [162] Lars relund nielsen, asger roer pedersen, mette s. herskin, lene munksgaard, quantifying walking et standing behaviour of dairy cows using a moving average based on output from an accelerometer, applied animal behaviour science, volume 127. 2010.
- [163] J barwick, d lamb, r dobos, m welchet m trotter. categorising sheep activity using a tri axial accelerometer. computers and electronics in agriculture. 2018. dio: 10.1016/j.compag.2018.01.007.
- [164] Joan bryce burla, anic ostertag, heike schulze westerath, edna hillmann, gait determination and activity measurement in horses using an accelerometer, computers and electronics in agriculture, volume 102, 2014.

- [165] Leifeng guo, mitchell welch, robin dobos, paul kwan, wensheng wang, comparison of grazing behaviour of sheep on pasture with different sward surface heights using an inertial measurement unit sensor, *computers and electronics in agriculture*, volume 150, 2. 2018.
- [166] B. white, j. coetzee, d. renter, a. babcock, d. thomson and d. andresen, evaluation of two dimensional accelerometers to monitor behavior of beef calves after castration, *american journal of veterinary research*, vol. 69, no. 8, pp. 1005 1012, 2008.
- [167] Z.e. barker, j.a. viazquez diosdado, e.a. codling, n.j. bell, h.r. hodge, d.p. croft, j.r. amory. use of novel sensors combining local positioning and acceleration to measure feeding behavior differences associated with lameness in dairy cattle, *journal of dairy science*, volume 101, issue 7, 2018.
- [168] said benaissa, frank a.m. tuytens, david plets, hannes cattrysse, luc martens, leen vandaele, wout joseph, bart sonck . classification of ingestive related cow behaviours using rumiwatch halter and neck mounted accelerometers, *applied animal behaviour science*, volume 211, 2019.
- [169] Md. sumon shahriar, daniel smith, ashfaqur rahman, mark freeman, james hills, richard rawnsley, dave henry, greg bishop hurley. detecting heat events in dairy cows using accelerometers and unsupervised learning, *computers and electronics in agriculture*, 2016.
- [170] T. hammond, d. springthorpe, r. walsh and t. berg kirkpatrick. using accelerometers to remotely and automatically characterize behavior in small animals, *the journal of experimental biology*, vol. 219, no. 11, pp. 1618 1624, 2016.
- [171] J behmann, k hendriksen, u miller, wbuescher, l plimer. support vector machine and duration aware conditional random field for identification of spatio temporal activity patterns by combined indoor positioning and heart rate sensors. *geoinformatica*. 2016.
- [172] D smith, a rahman, h bishop, h james, s shahriar, d henry, r rawnsley. behavior classification of cows fitted with motion collars: Decomposing multi class classification into a set of binary problems. *computers and electronics in agriculture*. 2016.
- [173] Matti pastell, lilli frondelius, a hidden markov model to estimate the time dairy cows spend in feeder based on indoor positioning data, *computers and electronics in agriculture*, volume 152, 2018.
- [174] F. yang and l. zhang, real time human activity classification by accelerometer embedded wearable devices, 2017 4th international conference on systems and informatics (icsai), hangzhou, pp. 469 473, 2017.

- [175] M. shoaran, b. a. haggi, m. taghavi, m. farivar and a. emami neyestanak, energy efficient classification for resource constrained biomedical applications, in *ieee journal on emerging and selected topics in circuits and systems*, vol. 8, no. 4, 2018.
- [176] Y. tang and n. verma, energy efficient pedestrian detection system: Exploiting statistical error compensation for lossy memory data compression, in *ieee transactions on very large scale integration (vlsi) systems*, vol. 26, no. 7, july 2018.
- [177] Wenyong li, zengtao ji, lin wang, chuanheng sun, xinting yang, automatic individual identification of holstein dairy cows using tailhead images, *computers and electronics in agriculture*, volume 142, part b, 2017.
- [178] W. andrew, c. greatwood and t. burghardt, visual localisation and individual identification of holstein friesian cattle via deep learning, 2017 *ieee international conference on computer vision workshops (iccvw)*, venice, 2017.
- [179] Tarek gaber, alaa tharwat, aboul ella hassanien, vaclav snasel. biometric cattle identification approach based on weber's local descriptor and adaboost classifier. *computers and electronics in agriculture*, volume 122, 2016.
- [180] Kim h, ikeda y, choi h. the identification of japanese black cattle by their faces. *asian australas j anim sci*, 2005.
- [181] Q. yang, d. xiao, s. lin feeding behavior recognition for group-housed pigs with the faster r-cnn *journal of computers and electronics in agriculture*, 2018), dio: 10.1016/j.compag.2018.11.002.
- [182] M.f. hansen, m.l. smith, l.n. smith, m.g. salter, e.m. baxter, et al. towards on-farm pig face recognition using convolutional neural networks *journal of computers in industry*, 2018, dio: 10.1016/j.compind.2018.02.016.
- [183] W. shen, h. hu, b. dai, x. wei, j. sun, et y. sun. individual identification of dairy cows based on convolutional neural networks. *journal of multimedia tools and applications*, 2019.
- [184] P. sepulveda varas, k. l. proudfoot, d. m. weary et m. a. v. keyserlingk, changes in behaviour of dairy cows with clinical mastitis, *applied animal behaviour science*, 2016.
- [185] F. llario, s. sendra, l. parra et j. lloret, detection and protection of the attacks to the sheep and goats using an intelligent wireless sensor network, *ieee international conference on communications workshops (icc2013)*, budapest, hungary, pp. 1015 1019, jun. 2013.

- [186] S. sendra, f. llario, l. parra et j. lloret, smart wireless sensor network to detect and protect sheep and goats, to wolf attacks, recent advances in communications and networking technology, vol. 2, pp. 91 101, 2013.
- [187] A. buerkert et e. schlecht, performance of three gps collars to monitor goats grazing itineraries on mountain pastures, computers and electronics in agriculture, vol. 65, pp. 85 92, 2009.
- [188] M. pastell, j. hietaoja, j. yun, j. tiusanen et a. valros, predicting farrowing of sows housed in crates and pens using accelerometers and cusum charts, computers and electronics in agriculture, vol. 127, pp. 197 623, 2016.
- [189] S. t. ahmed, h. s. mun, m. islam, h. yoe et c. j. yang, monitoring activity for recognition of illness in experimentally infected weaned piglets using received signal strength indication zigbee based wireless acceleration sensor, asian australasian j. animal sci, vol. 29, pp. 149 156, 2015.
- [190] W. huang, w. zhu, c. ma, y. guo et c. chen, identification of group housed pigs based on gabor and local binary pattern features, journal of biosystems engineering, vol. 166, pp. 90 100, february 2018.
- [191] P. loreti, a. catini, m. d. luca, l. bracciale, g. gentile et c. d. natale, ultra low power wireless sensor network for pink iguanas monitoring, multidisciplinary digital publishing institute proceedings, 2018.
- [192] T. norton, d. piette, v. exadaktylos et d. berckmans, automated real time stress monitoring of police horses using wearable technology, applied animal behaviour science journal, vol. 198, pp. 67 74, 2018.
- [193] S. o. h. madgwick, a. j. l. harrison et r. vaidyanathan, estimation of imu and marg orientation using a gradient descent algorithm, 2011 ieee international conference on rehabilitation robotics, zurich, 2011.
- [194] G. mclachlan et d. peel, finite mixture models., john wiley and sons publishers., 2000.
- [195] T. benaglia, d. chauveau, d. r. hunter et d. s. young, mixtools: An r package for analyzing finite mixture, journal of statistical software, 2009.
- [196] B. gabriel et b. valner, sensor fusion methods for reducing false alarms in heart rate monitoring, journal of clinical monitoring and computing, 2016.
- [197] Q. t. huynh, u. d. nguyen, l. b. irazabal, n. ghassemian et b. q. tran, optimization of an accelerometer and gyroscope based fall detection algorithm, journal of sensors , 2015.

- [198] B. anda, s. baglio, c. o. lombardo et v. marletta, a multisensor data fusion approach for adl and fall classification, *iee transactions on instrumentation and measurement*, vol. 65, sept. 2016.
- [199] P. tsinganos et a. skodras, on the comparison of wearable sensor data fusion to a single sensor machine learning technique in fall detection, *sensors* , 2018.
- [200] L. scrucca, m. fop, t. murphy et a. raftery, mclust 5: Clustering, classification and density estimation using gaussian finite mixture models, *the r journal*, 2016.
- [201] K. m. abell, m. e. theurer, r. l. larson, b. j. white, d. k. hardin et r. f. randle, predicting bull behavior events in a multiple sire pasture with video analysis, accelerometers, and classification algorithms, *computers and electronics in agriculture*, vol. 136, 2017.
- [202] Moore, j. e., brant, m. h., kunkle, w. e., et hopkins, d. i. effects of supplementation on voluntary forage intake, diet digestibility and animal performance. *journal of animal science*, 1999. doi:10.2527/1999.77suppl\_2122x.
- [203] Coleman, s. w., et moore, j. e. feed quality and animal performance. *field crops research*, 2003. doi:10.1016/s0378-4290(03)00138-2.
- [204] Forbes, j. m. voluntary food intake and diet selection in farm animals. cambridge, ma: Cabi north american office. 2007.
- [205] Harder, i., stamer, e., junge, w., et thaller g. estimation of genetic parameters and breeding values for feed intake and energy balance using pedigree relationships or single-step genomic evaluation in holstein friesland cows. *journal of dairy science*. 2019. dio: <https://doi.org/10.3168/jds.2019-16855>.
- [206] Neeteson-van nieuwenhoven, a.-m., merks, j., bagnato, a., et liinamo, a.-e. sustainable transparent farm animal breeding and reproduction. *livestock science*, 2006. dio: doi:10.1016/j.livsci.2006.05.016.
- [207] K. schirmann, d.m. weary, w. heuwieser, n. chapinal, r.l.a. cerri, m.a.g. von keyserlingk, short communication: Rumination and feeding behaviors differ between healthy and sick dairy cows during the transition period, *journal of dairy science*, volume 99, issue 12,2016.
- [208] M.r. borchers, y.m. chang, i.c. tsai, b.a. wadsworth, j.m. bewley, a validation of technologies monitoring dairy cow feeding, ruminating, and lying behaviors, *journal of dairy science*, volume 99, issue 9, 2016.

- [209] Shane, d. d., white, b. j., larsen, r. l., amrine, d. e., et kramer, j. l. probabilities of cattle participating in eating and drinking behavior when located at feeding and watering locations by a real time location system. *journal of computers and electronics in agriculture*, 2016. dio: doi:10.1016/j.compag.2016.07.005.
- [210] Porto, s. m. c., arcidiacono, c., giummarra, a., anguzza, u., et cascone, g. localisation and identification performances of a real-time location system based on ultra wide band technology for monitoring and tracking dairy cow behaviour in a semi-open free-stall barn. *journal of computers and electronics in agriculture*, 2014. dio: doi:10.1016/j.compag.2014.08.001.
- [211] Zehner, n., umstatter, c., niederhauser, j. j., et schick, m. system specification and validation of a noseband pressure sensor for measurement of ruminating and eating behavior in stable-fed cows. *journal of computers and electronics in agriculture*, 2017. doi:10.1016/j.compag.2017.02.021.
- [212] Ruuska, s., kajava, s., mughal, m., zehner, n., et mononen, j. validation of a pressure sensor-based system for measuring eating, rumination and drinking behaviour of dairy cattle. *journal of applied animal behaviour science*, 2016. doi:10.1016/j.applanim.2015.11.005.
- [213] Vanrell, s. r., chelotti, j. o., galli, j. r., utsumi, s. a., giovanini, l. l., rufiner, h. l., et milone, d. h. a regularity-based algorithm for identifying grazing and rumination bouts from acoustic signals in grazing cattle. *journal of computers and electronics in agriculture*, 2018. doi:10.1016/j.compag.2018.06.021.
- [214] Deniz, n. n., chelotti, j. o., galli, j. r., planisich, a. m., larripa, m. j., leonardo rufiner, h., et giovanini, l. l. embedded system for real-time monitoring of foraging behavior of grazing cattle using acoustic signals. *computers and electronics in agriculture*, 2017. doi:10.1016/j.compag.2017.04.024.
- [215] L. blanger et a. panisson, a face recognition library using convolutional neural networks, *international journal of engineering research and science*, vol. 3, no. 8, pp. 84 92, 2017.
- [216] A. mcnamee, m. keane, d. kenny, a. moloney, f. buckley et e. o riordan, beef production from holstein friesian, norwegian red holstein friesian and jersey holstein friesian male cattle reared as bulls or steers, *livestock science*, vol. 173, 2014.
- [217] Z. zivkovic et f. van der heijden, efficient adaptive density estimation per image pixel for the task of background subtraction, *pattern recognition letters*, vol. 27, no. 7, pp. 773 780, 2006. dio: 10.1016/j.patrec.2005.11.005.

- [218] Zivkovic, z. improved adaptive gaussian mixture model for background subtraction. 17th international conference on pattern recognition, icpr, 2004. doi:10.1109/icpr.2004.1333992.
- [219] Z. wang, a. bovik, h. sheikh et e. simoncelli, image quality assessment: From error visibility to structural similarity, *ieee transactions on image processing*, vol. 13, no. 4, pp. 600 612, 2004. dio: 10.1109/tip.2003.819861.
- [220] S. van der walt et al., scikit image: image processing in python, *peerj*, vol. 2, p. e453, 2014. dio: 10.7717/peerj.453.
- [221] Y. zhang, x. cui, y. liu and b. yu, tire defects classification using convolution architecture for fast feature embedding, *international journal of computational intelligence systems*, vol. 11, no. 1, p. 1056, 2018. dio: 10.2991/ijcis.11.1.80.
- [222] Yongliang qiao, matthew truman, salah sukkarieh, cattle segmentation and contour extraction based on mask r-cnn for precision livestock farming, *computers and electronics in agriculture*, volume 165, 2019.
- [223] Dihua wu, xuqiang yin, bo jiang, mei jiang, zhenyu li, huaibo song, detection of the respiratory rate of standing cows by combining the deeplab v3+ semantic segmentation model with the phase-based video magnification algorithm, *biosystems engineering*, 2020.
- [224] B xu , w wensheng, f greg, k paul, g leifeng, c guipeng, t amy, s derek. automated cattle counting using mask r-cnn in quadcopter vision system., *computers and electronics in agriculture*. 2020.
- [225] Petrovai, a., nedeveschi, s. efficient instance and semantic segmentation for automated driving. 2019 *ieee intelligent vehicles symposium (iv)*. 2019. doi: 10.1109/ivs.2019.8814177.
- [226] Jesus ruiz-santaquiteria, gloria bueno, oscar deniz, noelia vallez, gabriel cristobal, semantic versus instance segmentation in microscopic algae detection, *engineering applications of artificial intelligence*, volume 87,2020.
- [227] chen, g. papandreou, i. kokkinos, k. murphy et a. yuille, deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, *ieee transactions on pattern analysis and machine intelligence*, vol. 40. 2016.
- [228] M. abadi, tensorflow: learning functions at scale, *acm sigplan notices*, vol. 51, no. 9, pp. 1 1, 2016. dio: 10.1145/3022670.2976746.
- [229] Andrew g. howard et al , mobilenets: Efficient convolutional neural networks for mobile vision applications, *corr*, 2017.

- [230] Francois chollet, xception: Deep learning with depthwise separable convolutions, corr,2016.
- [231] Kaiming he et al , deep residual learning for image recognition ,corr,2015.
- [232] Pasumarthi, r. k., bruch, s., wang, x., li, c., bendersky, m., najork, m., pfeifer, j., golbandi, n., anil, r., et wolf, s. . tf-ranking: Scalable tensorflow library for learning-to-rank. in proceedings of the 25th acm sigkdd international conference on knowledge discovery and data mining (kdd '19). acm, new york, ny, usa, 2019.
- [233] X. niu et c. suen, a novel hybrid cnn svm classifier for recognizing handwritten digits, pattern recognition, vol. 45, no. 4, pp. 1318 1325, 2012. dio: 10.1016/j.patcog.2011.09.021.
- [234] Ramprasaath r. selvaraju et al , grad cam: Why did you say that? visual explanations from deep networks via gradient based localization, corr ,2016.
- [235] Haifeng wu, qing huang, daqing wang, lifu gao, a cnn-svm combined model for pattern recognition of knee motion using mechanomyography signals, journal of electromyography and kinesiology, volume 42, 2018.
- [236] S. porto, c. arcidiacono, u. anguzza and g. cascone, "the automatic detection of dairy cow feeding and standing behaviours in free stall barns by a computer vision based system", biosystems engineering, vol. 133, pp. 46 55, 2015. available: 10.1016/j.b.
- [237] S. kumar et s. singh, automatic identification of cattle using muzzle point pattern: a hybrid feature extraction and classification paradigm, multimedia tools and applications, vol. 76, no. 24, pp. 26551 26580, 2016. dio: 10.1007/s11042 016 41.
- [238] Barkema, h., von keyserlingk, m., kastelic, j., lam, t., luby, c., roy, j. et al. invited review: Changes in the dairy industry affecting dairy cattle health and welfare. journal of dairy science, 2015. doi: 10.3168/jds.2015-9377.
- [239] Scientific opinion on the assessment of dairy cow welfare in small-scale farming systems. efsa journal. 2015. doi: 10.2903/j.efsa.2015.4137.
- [240] He liu, amy r, reibman, jacquelyn p, boerman a. cow structural model for video analytics of cow health. arxiv, 2020.
- [241] Brahim achour, malika belkadi, idir filali, mourad laghrouche, mourad lahdar, image analysis for individual identification and feeding behaviour monitoring of dairy cows based on convolutional neural networks (cnn), biosystems engineering, volume 198, pages 31-49, 2020, dio: <https://doi.org/10.1016/j.biosystemseng.2020.07.019>.

- [242] J. bikker et al., technical note: Evaluation of an ear attached movement sensor to record cow feeding behavior and activity, journal of dairy science, vol. 97, no. 5, pp. 2974 2979, 2014.