ENHANCED AUTOMATED FRAMEWORK FOR CATTLE TRACKING AND CLASSIFICATION

BELLO ROTIMI-WILLIAMS

UNIVERSITI SAINS MALAYSIA

2022

ENHANCED AUTOMATED FRAMEWORK FOR CATTLE TRACKING AND CLASSIFICATION

by

BELLO ROTIMI-WILLIAMS

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

September 2022

ACKNOWLEDGEMENT

My undiluted appreciation goes to God Almighty the creator of the universe who in His infinite mercy has made ways for me where seems to be no way.

My sincere appreciation goes to my Supervisors; Dr. Ahmad Sufril Azlan Mohamed and Prof. Abdullah Zawawi Talib for their availability, guidance, constructive supervision and feedback on this thesis. Their availability and comments have in no small measure reshaped the structure of the whole thesis, thereby making the research a huge success.

To the Dean (both past and present) of the School of Computer Sciences, Universiti Sains Malaysia for providing an enabling environment that reshaped my research mentality, I express my gratitude.

To the School of Computer Sciences' Journal Club, for providing the platform for scholars to share ideas and published papers, I say thank you.

My appreciation also goes to the Institute of Postgraduate Studies, Universiti Sains Malaysia for the Graduate Assistant position I was offered throughout my candidature.

I was filled with emotional pains when departing from my family especially my wife and my children to study for the Ph.D. overseas; therefore, I am indebted to them for their understanding, endurance and perseverance throughout the period.

I am grateful to my late parents whose love for my education knows no boundaries. Also, I am indebted to my siblings, my relatives-in-law, my academic advisers, and my dependable friends for their love for me.

I gratefully acknowledge the support of others too numerous to mention who have immensely contributed to the success of my academic journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADE	Average Distance Error
AHDS	Animal Human Detection System
AlexNet	Alex Network
AP	Average Precision
ASIFT	Affine Scale-Invariant Feature Transform
AUC	Area under Curve
BP	Back Propagation
CCTV	Closed-Circuit TeleVision
CLAHE	Contrast Limited Adaptive Histogram Equalisation
CNN	Convolutional Neural Networks
C-SIFT	Colour-Scale Invariant Feature Transform
DELF	Deep Local Feature
DNN	Deep Neural Networks
DPFF	Deep Parts Features Fusion
DT	Decision Tree
FCIS	Fully Convolutional Instance-Aware Semantic Segmentation
FCN	Fully Convolutional Network
FC6	Fully Connected Layer 6
FC7	Fully Connected Layer 7
FERET	Face Recognition Technology
FFT	Fast Fourier Transforms
F-KNN	Fuzzy-K Nearest Neighbour
FLANN	Fast Learning Artificial Neural Network
FPN	Feature Pyramid Networks
FPS	Frames Per Second
GCFD	Generalised Colour Fourier Descriptors
GFD	Generalised Fourier Descriptor
GMM	Gaussian Mixture Model
GoogLeNet	Google Networks
GPU	Graphic Processing Unit
HOG	Histogram of Oriented Gradients

HoOG	Haar of Oriented Gradients
HSV	Hue, Saturation, Value
HueSIFT	Hue, Saturation, Value-Scale Invariant Feature Transform
IDE	Integrated Development Environment
ImageNet	Image Network
IOU	Intersection over Union
KBPS	Kilobits Per Second
KHz	KiloHertz
K-NN	K-Nearest Neighbour
LDA	Latent Dirichlet Allocation
MAP	Mean Average Precision
Mask R-CNN	Mask Region-based Convolutional Neural Network
MB	Mega Byte
MLP	Multi-Layer Perceptron
MPA	Mean Pixel Accuracy
$PF_{tm}M$	Particle $Filter_{temporary\ memory}$ Mean-Shift
MPII	Max-Planck Institute for Informatics
MS COCO	Microsoft Common Objects in Context
NBCM	Naive Bayes Classification Models
OpenCV	Open Source Computer Vision
ORB	Oriented-Fast and Rotated Brief
PF	Particle Filter
PF_{tm}	Particle $Filter_{temporary\ memory}$
PNN	Probabilistic Neural Network
RBF	Radial Basis Function
R-CNN	Region-Based Convolutional Neural Networks
ResNet	Residual Network
R-FCN	Region-Based Fully Convolutional Networks
RGB	Red, Green, Blue
RGB-D	Red, Green, Blue-Depth
rgSIFT	$_{redgreen}$Scale Invariant Feature Transform
RNN	Recurrent Neural Network
ROI	Region of Interest
ROIAlign	Region of Interest-Alignment

RPN	Region Proposal Network
SFU	Simon Fraser University
SIFT	Scale-Invariant Feature Transform
SLAM	Simultaneous Localisation And Mapping
SPPNet	Spatial Pyramid Pooling Networks
SSD	Single-Shot ‘Multibox’ Detection
SSWC	Sum of Squared Within Cluster
SURF	Speeded Up Robust Feature
SUN	Scene Understanding
SVM	Support Vector Machine
UAV	Unmanned Aerial Vehicle
VGGNet	Visual Geometry Group-Network
VOC	Visual Object Classes
YOLO	You Only Look Once
ZFNet	Zeiler and Fergus Networks
2D	Two Dimensions
3D	Three Dimensions

LIST OF APPENDICES

KERANGKA AUTOMATIK YANG DIPERTINGKATKAN UNTUK PENJEJAKAN DAN KLASIFIKASI LEMBU

ABSTRAK

Penggunaan kaedah berasaskan penglihatan komputer untuk memantau lembu secara individu merupakan antara usaha yang kini giat dilakukan oleh para penyelidik. Kaedah berasaskan penglihatan komputer boleh digunakan untuk memantau setiap individu lembu. Ketepatan kaedah dan kerangka kerja sedia ada adalah di bawah jangkaan dalam mengendalikan tugasan ini. Lebih-lebih lagi, kaedah-kaedah berkenaan masih boleh ditambah baik untuk mencapai keputusan yang lebih baik dan lebih tepat. Matlamat penyelidikan ini adalah untuk menyediakan kerangka bagi sistem penjejakan dan klasifikasi lembu yang lebih baik. Algoritma penjejakan objek yang dipertingkatkan ($PF_{tm}M$) yang menyepadukan algoritma penapis zarah (PF_{tm}) dengan penjejak anjakan min (M) dicadangkan dan digunakan sebagai langkah pertama untuk menangani masalah yang timbul akibat berlakunya oklusi dan pergerakan bukan linear objek lembu dalam video. Penyepaduan penapis zarah dengan penjejak anjakan min mengambil kira teknik berikut: (1) ingatan sementara untuk menjejaki objek lembu yang dioklusi; (2) menambah setiap kelemahan algoritma dengan kekuatan algoritma yang lain untuk menjejaki pergerakan bukan linear. Kekuatan penapis zarah (PF) ialah sifat bukan linearnya yang digunakan untuk menjejak pergerakan bukan linear objek tetapi, dengan masa pengiraan yang tinggi dan julat carian sebagai kelemahannya. Kekuatan ingatan sementara (tm) ialah keupayaannya untuk menjejak oklusi penuh dengan pengurangan masa pengiraan dan julat carian. Kekuatan anjakan min ialah kepekaannya terhadap pergerakan objek dan penyebaran warna dengan

menggunakan fungsi persamaan tetapi, dengan ketidakupayaan untuk mengesan pergerakan bukan linear objek dan oklusi penuh sebagai kelemahannya. Langkah kedua melibatkan peningkatan model Mask R-CNN untuk klasifikasi lembu kepada jenisnya. Model Mask R-CNN yang dipertingkatkan mempertimbangkan kaedah berikut: (1) pra-peningkatan imej dengan memasukkan GCFD untuk mengurangkan masalah penukaran warna; (2) menyepadukan lapisan bersambung sepenuhnya Mask R-CNN dengan subrangkaian untuk klasifikasi lembu mengikut jenisnya. Di samping itu, kerangka automatik yang lebih baik yang menggabungkan semua kaedah yang dicadangkan dan algoritma lain yang sesuai untuk pengesanan dan klasifikasi lembu direka bentuk. Kaedah pengesanan dan pengelasan lembu yang dipertingkatkan menghasilkan ketepatan 89% dan purata ketepatan purata (mAP) sebanyak 0.93 pada data input (Muturu dan Keteku) dan 0.90 pada set data MS COCO lembu (Holstein dan Friesian). Kerangka kerja yang dicadangkan merekodkan ketepatan purata 95.97. Secara keseluruhannya, hasil utama penyelidikan ini boleh digunakan dalam bidang penternakan haiwan untuk penternakan jitu dalam memantau lembu secara individu termasuk kebajikan dan prestasi mereka.

ENHANCED AUTOMATED FRAMEWORK FOR CATTLE TRACKING AND CLASSIFICATION

ABSTRACT

Employing computer vision-based methods in monitoring individual cows has become what researchers are striving for. Computer vision-based methods could be used to monitor each individual cows. The accuracy of the existing methods and frameworks is below expectation in handling these tasks. Moreover, they can still be improved to achieve better and more accurate results. The goal of this research is to provide a framework for better cattle tracking and classification systems. An enhanced object tracking algorithm ($PF_{tm}M$) that integrates enhanced particle filter algorithm (PF_{tm}) with mean-shift tracker (M) is proposed and deployed as first step to address the problems arise due to occurrence of occlusion and non-linear movement of cow objects in video. The integration of particle filter with mean-shift tracker considers the following techniques: (1) temporary memory for keeping tracks of occluded cow objects; (2) supplementing each algorithm's weakness by the strength of the other for tracking non-linear movement. Strength of particle filter (PF) is its non-linearity property which it uses to track object's non-linear movement but, with high computational time and search range as its weakness. Temporary memory (tm) strength is its ability to track full occlusion with reduced computational time and search range. Mean-shift strength is its sensitivity to object's movement and colour distribution by using similarity function but, with inability to track object's non-linear movement and full occlusion as its weakness. The second step involves enhancing the Mask R-CNN model for effective classification of cows to their types. The enhanced Mask R-CNN model considers the following methods: (1) enhancing the images by incorporating GCFD to mitigate colour conversion problem; (2) integrating the fully connected layers of Mask R-CNN with a subnetwork for classification of cows to their types. In addition, a better automated framework incorporating all the proposed methods and other suitable algorithms for cattle tracking and classification is designed. The proposed enhanced cattle tracking and classification methods produced a precision of 89% and a mean average precision (mAP) of 0.93 on input data (Muturu and Keteku) and 0.90 on MS COCO cow (Holstein and Friesian) dataset respectively. The proposed framework recorded an average precision of 95.97. As a whole, the main output of this research could be employed by animal husbandry for precision livestock farming in monitoring individual cows including their welfare and performance.

CHAPTER 1

INTRODUCTION

1.1 Background

Object tracking means the enablement to monitor an object or group of objects during all phases of the object's life. Object tracking algorithm centres on the detection of an object in order to identify and get more facts in tracking the object to a particular location. Object tracking algorithm focuses on detecting the location of an object in order to further obtain other information (Haroun *et al.*, 2019; Fukunaga *et al.*, 2015; Homburger *et al.*, 2014). Several popular object tracking algorithms include particle filter algorithm (Iswanto & Li, 2017), mean-shift algorithm (Zhi-Qiang *et al.*, 2014) and convolutional neural network-based methods (Li *et al.*, 2021).

Particle filter algorithm steps such as initialisation step (prediction), sampling step (update) and selection step (resample) are improved by eliminating bounding box proposals and the subsequent pixel or feature resampling stage, and by encapsulating all computation in a single network (Liu *et al.*, 2016b; Redmon *et al.*, 2016). This improvement includes the use of a small convolutional network to predict object category and offsets in bounding box locations and using separate predictors (filters) for different aspects ratio detections. Applying these filters to multiple feature maps forms the later stages of a network in order to perform detection at multiple scales (Liu *et al.*, 2016b). However, the accuracy of this process can still be increased significantly for object tracking algorithms.

Mean-shift is among the widely employed methods of visual object tracking due to its easy algorithm which enables fast computation and is lusty against partial occlusion, thereby making it more efficient and effective for tracking applications (Zhi-Qiang *et al.*, 2014). However, it has a notable limitation which is its inability to track fully occluded and fast-moving object. In an effort to overcome the existence of occlusion, non-linear motion and other difficulties involved in object tracking, a combination of mean-shift and other tracking algorithms was proposed by Iswanto and Li (2017). The reason for combining these tracking algorithms is to use the strength of one algorithm to supplement the weakness of the other algorithm.

Mean-shift is employed as a master tracker when there is no occurrence of occlusion to the object being tracked. When there is an occurrence of occlusion or the tracking output of mean-shift is not convincing, the integrated algorithms become the master tracker in improving the tracking outputs. The experimental outputs of their proposed method show a desirable performance. However, their method only employed colour as the tracking feature, thereby making tracking of objects with similar colour distribution difficult. Therefore, this method can be further improved in order to obtain better accuracy for object tracking algorithms.

The Convolutional Neural Network-based methods for object tracking are limited (Li *et al.*, 2021). Among the common methods are 1) Two-stream Network (Simonyan & Zisserman, 2014); 2) Long-term Recurrent Convolutional Networks (LRCN) (Donahue *et al.*, 2015); 3) Generic Object Tracking Using Regression Networks (GOTURN) (Held *et al.*, 2016); and 4) SlowFast Network (Feichtenhofer *et al.*, 2019). Other methods such as Simple Online and Real-time Tracking (SORT), Hungarian Algorithm (HA), Munkres Variant of the Hungarian Assignment Algorithm (MVHAA), Spatial-aware Temporal Response Filter (STRF), and Channel and Spatial Reliability Discriminative Correlation Filter Tracker (CSRDCF) were utilised by object detection models, such as Faster R-CNN, FCN, SSD, VGG and YOLO for the detection and tracking of animals in images using their geometric features in continuous frames (Li *et al.*, 2021). Nevertheless, the accuracy of object tracking algorithms can still be significantly increased.

These days, object classification that involves locomotive living organisms such as mammals and non-locomotive living organisms such as plants has gained much acceptance by researchers (Pereira *et al*., 2019; Kumar *et al*., 2018; Trnovszký *et al*., 2017). The salient reason for object classification is for the identification of individual objects from a group of objects. Different types of cow objects possess and display different characteristics such as behaviours, grazing and feeding patterns. The use of spatial positioning extraction or specific movement pattern extraction from sequence of cow images improves the classification of cow objects (Song *et al*., 2019; Hansen *et al*., 2018; Zhao *et al*., 2018a; Valletta *et al*., 2017).

Machine learning methods have a significant impact on the field of Artificial Intelligence. Convolutional Neural Network (CNN), Region-based Convolutional Neural Network (R-CNN), Fast R-CNN, Faster R-CNN, Recurrent Neural Network (RNN) and other Convolutional Neural-based Networks have demonstrated their effectiveness and achieved better results that surpass the performance of the human level in many learning tasks end-to-end. However, machine learning applications are not yet widespread in the classification of cows (Bezen *et al*., 2020; Kumar & Singh, 2019; Kumar *et al*., 2018).

The accuracy of the classification process for object classification methods can be further improved. Moreover, it is necessary to provide a better automated method that can perform tracking and classification of cows to their types. Therefore, this research attempts to improve cattle tracking and classification methods for identifying the cow type.

1.2 Research Motivation

Cows are of different breeds, and an individual cow is being reared for different purposes. For example, some are reared for farm labour, some for their milk, some for their meat and some for their skin to make leather materials. So, regular monitoring of the cows is needed but, it is not easy to monitor the cows all the time.

Object tracking and classification all over the world are still considered as a herculean task to accomplish even though several algorithms have been proposed. Human efforts cannot be compared to computational methods of tracking when it comes to the tracking of objects. The level of information which can be obtained about the activity of the cows based on movements is minimal.

Not much research had been recorded on addressing cow rustling (stealing) and other challenges confronting cow business. Stolen and non-registered cows are being transported across the border, herds are being swapped during grazing, ownership disputes and false insurance claims are increasing, and arresting these kinds of illegal activity has been a difficult task for cow breeders and herders.

Moreover, since humans and animals do not possess the same classifications and characteristics, human identifiers cannot be used for animal identification. There are so many processing tasks that involve images which existing image segmentation methods can no longer capable of handling; some segmentation tools produced are no longer applicable to cater for the present segmentation problems. Therefore, the main motivation of this research work is the need to proffer a solution to the above challenges.

1.3 Problem Statement

The application of single algorithms such as particle filter algorithm (Rodriguez *et al.*, 2018), mean-shift (Zhi-Qiang *et al.*, 2014) and Kalman filter algorithm (Kim *et al.*, 2011) have notable issues of heavy computational time, colour misjudgement and non-linear tracking respectively. Also, the application of combined algorithms such as mean-shift-particle-Kalman filter algorithm (Iswanto & Li, 2017), mean-shift-Kalman filter algorithm (Iraei & Faez, 2015), mean-shift-particle filter algorithm (Qiao & Yu, 2014) and particle-Kalman filter algorithm (Yin *et al.*, 2011) are inefficient in attacking the above-mentioned issues of single algorithms due to their individual limitations as investigated in this study.

Particle filter algorithm major limitation is in the substantial number of particles required as samples which results in heavy computational time and reduces the performance of tracking speed, most especially when multiple-object tracking is involved where an immense number of particles is employed for each target object. Mean-shift utilises only colour as the tracking feature, thereby making tracking of objects with similar colour distribution difficult.

Most movements in the real world are mostly non-Gaussian movements which affect Kalman filter-based visual object tracking performance when the target object suddenly changes the velocity from linear movement to non-linear movement. Mean-shift-Kalman filter could only deal with linear moving objects since the Kalman filter is employed as the predictor when the object is occluded.

Mean-shift particle filter requires substantial number of particles as samples which results in heavy computational time and low performance. Moreover, other objects with similar colour distribution may occlude the target object due to inability of mean-shift to distinguish objects with similar colour distribution, thereby rendering the system useless. Particle-Kalman filter takes immense time to implement the linearity of the Kalman filter for every particle to obtain a more condensed particle to deal with the non-linearity problem.

Mean-shift-particle-Kalman filter algorithm employs only colour as the tracking feature, and takes immense time to implement the linearity of the Kalman filter for every particle to obtain a more condensed particle to deal with the non-linearity problem. Moreover, enhancement of the existing cow tracking algorithms based on the field of object tracking is greatly affected by significant variations (Achour *et al.*, 2020; Achour *et al.*, 2019).

To address the problem of tracking multiple animals, different approaches have been developed (Xu *et al.*, 2020). A notable attempt to keep the identities correct over time is the work by Pérez-Escudero *et al.* (2014). The method extracts characteristic fingerprints from each animal that are matched with the trajectories in the video and uses a re-segmentation stage, which optimises for particular shapes that reduce the number of occlusions.

However, this approach is very computationally heavy and not suitable for automated applications. Therefore, enhancement of these tracking methods is necessary in order to obtain a more accurate method that is computationally light and suits the tracking that involves occlusion and non-linear movement of cow objects in videos.

The task involves in cow image segmentation is a complex task as so many removals of unwanted background objects may be involved in the image segmentation process (Williams *et al.*, 2019; Ter-Sarkisov *et al.*, 2017). A lot of researchers have worked on cow classification methods but most concentrated on static images and put less effort into motion images (Jingqiu *et al.*, 2017; Duyck *et al.*, 2015).

Not much work has been done in improving the existing cow image segmentation methods which is important especially when there is a need to classify multiple objects of the same characteristics in the same video frame using their unique representation features (Liu *et al.*, 2020).

Moreover, the existing instance segmentation methods such as Mask R-CNN (He *et al.*, 2020; He *et al.*, 2017), SOLO (Wang *et al.*, 2021; Wang *et al.*, 2020) and FCIS (Li *et al.*, 2017), and semantic segmentation methods such as FCN (Long *et al.*, 2015) and Deeplab V3+ (Wu *et al.*, 2020) are not effective in detecting and classifying the cow objects to their types due to their individual limitations.

Mask R-CNN major limitations are 1) the detect-then-segment strategy it uses to predict mask for each instance reduces the accuracy of its performance; 2) the Simultaneous Localisation And Mapping (SLAM) algorithm it employs in extracting image features affects the quality of the segmentation; 3) the fully connected layers (FCLs) it uses in classifying objects can only categorise by class and not by type.

Although SOLO (Segmenting Objects by LOcations) eliminates the need for the bounding box detection, ROI operations, and grouping post-processing attributed to the mainstream approaches, identification of individual objects by their types is not considered in the method.

FCIS (Fully Convolutional Instance-aware Semantic segmentation) being an extension of instance mask proposal and FCNs, automatically inherits their notable limitations among which are 1) reduction in the accuracy of performance due to the detect-then-segment strategy they use in predicting mask for each instance; 2) inability to classify objects by their types but by their class. Deeplab V3+ being a semantic segmentation model that was developed using the ResNet-101 framework automatically inherits the segmentation issue of the SLAM algorithm which is its inability to completely extract object from an image background.

Furthermore, the existing characterisation methods are not accurate enough as the image of a cow is affected by some external factors which cause the image to appear blurry such as in the case where illumination variation causes image patches which distort the images and contribute to the increasing difficulty of the detection process (Kumar *et al.,* 2018*).*

Efforts to address the above-mentioned challenges were made in the review works of Qiao *et al*. (2021) and Li *et al.* (2021). From their reviews, they anticipated the development of intelligent perception for precision cows farming through automated technologies combined with deep learning technologies. Therefore, there is a need to further enhance the existing image segmentation methods to suit scenarios which is effective for classification of cows to their types.

There are so many human identifiers such as MPII (Andriluka *et al.*, 2014), COCO for human skeleton (Lin *et al.*, 2014), DeepPose for human body parts detection using images (Toshev & Szegedy, 2014), Stackedhourglass network (Newell *et al.*, 2016), ArtTrack (Insafutdinov *et al.*, 2017), OpenPose (Cao *et al.*, 2019; Cao *et al.*, 2017) and Deepcut (Pishchulin *et al.*, 2016) but, they are capable of identifying only humans, and they cannot be used for animal identification without modification. The existing cow detection and classification systems such as MaskSplitter (Ter-Sarkisov *et al.*, 2018) and Mask R-CNN (Salau & Krieter, 2020; Xu *et al.*, 2020; Qiao *et al.*, 2019) exhibit systematic errors on overlapping instances and create spurious edges.

In an effort to improve the rate of tracking and classification of cows most especially in an automated environment, Qiao *et al*. (2021) presented a framework for intelligent perception for precision livestock farming. They analysed important existing methods employed in precision cow farming so that research can be facilitated and development of related areas promoted. However, there is no available framework that includes either modified or built for identifying animal types.

Therefore, availability of an automated framework that can track and identifying animal type is necessary. Hence, there is a need to design an enhanced framework for object tracking and classification for tracking and classifying cows to their types. Therefore, the research questions are:

- How can the existing object tracking methods be enhanced to suit the occurrence of occlusion and non-linear movement of cow in video?
- What improvements are required to the existing image segmentation methods for effective classification of cows to their types?
- What better automated framework can be designed for tracking and classification of cows to their types?

1.4 Research Objectives

The goal of this research is to provide a framework for better cattle tracking and classification systems. Therefore, for this goal to be achieved, this research has the following objectives:

- To enhance the existing object tracking methods to suit the occurrence of occlusion and non-linear movement of cow in video.
- To improve the existing image segmentation methods for effective classification of cows to their types.
- To design a better automated framework for tracking and classification of cows to their types.

1.5 Research Contributions

The contributions of this research are as follows:

- An enhanced object tracking algorithm ($PF_{tm}M$) which integrates particle filter algorithm (PF) with mean-shift tracker (M) that addresses the problems arise due to occlusion and non-linear movement of cow objects in video. The integration of particle filter with mean-shift tracker considers the following techniques:
  - Temporary memory for keeping tracks of occluded cow objects.
  - Supplementing each algorithm's weakness by the strength of other for tracking non-linear movement. Strength of particle filter (PF) is its non-linearity property which it uses to track object's non-linear movement but, with high computational time and search range as its weakness. Temporary memory (tm) strength is its ability to track full occlusion with reduced computational time and search range. Mean-shift strength is its sensitivity to object's movement and colour distribution by using

similarity function but, with inability to track object's non-linear movement and full occlusion as its weakness.

- An improved image segmentation method for effective classification of cows to their types. The enhanced Mask R-CNN method considers the following methods:
  - o Enhancing the images by incorporating GCFD to mitigate colour conversion problem.
  - o Integrating the fully connected layers of Mask R-CNN with a subnetwork for classification of cows to their types.
- A better automated framework that combines the algorithms of tracking and classification methods for the tracking and classification of cows to their types. Each of the combined algorithms of the framework has the following contributions:
  - o Enhanced object-tracking algorithm which includes temporary memory for keeping tracks of multimodal objects.
  - o Improved image segmentation method which removes non-Gaussian noise from an image.

Additionally, an adapted Grabcut method is integrated into the framework for contour extraction of the classified cow objects.

1.6 Scope and Limitation

In this research, the focus is on automated cattle tracking and classification methods for precision livestock monitoring. The focus consists of tracking and classifying cows to their types by deploying tracking and classification methods. The input data (sequence of 1000 images) employed for the tracking and classification experiment was acquired using a cow capturing system in a cow's ranch containing two types of cows, namely Keteku and Muturu, and other complicated background objects.

The only cow dataset employed for this research was acquired from the Microsoft Common Objects in COntext (MS COCO) dataset (Lin *et al.,* 2014) in order to have a wider range of datasets that contain different densities of cows for evaluating the classification performance of the proposed methods.

The total number of cow image datasets in the MS COCO datasets is 2071 images; while there are 1986 images for training, 85 images are for validation and testing. MS COCO cow dataset is used for the evaluation because it is widely and publicly used for different computer vision competitions, and different researchers make use of it in this area of research. Holstein and Friesian cows are the two types of cow in MS COCO cow dataset.

The body patterns (colour) as unique texture features of the cows will be considered for feature extraction. A better automated framework for cattle tracking and classification will be designed. This research will be done up to the point where the cow types are identified. Therefore, a high-speed computer system that can execute all the processes will be used.

1.7 Benefits of the Research

This research benefits mainly the livestock farming/animal husbandry. The benefits are as follows:

- The research will be beneficial to the animal husbandry especially in Nigeria and the world in general in monitoring the welfare and performance of cattle.

- The research is a contributory step towards intelligent and precision livestock farming, from which further studies could be extended.

1.8 Thesis Organisation

The rest of this research is structured as follows:

Reviewed in Chapter 2 are the several existing methods used for object tracking and classification. Since the goal of this research is to provide a framework for better cattle tracking and classification systems, a more specific review is presented on ungulate animal tracking and classification frameworks.

Chapter 3 describes the research methodology of the proposed work, and it consists of three stages as follows: (1) enhancing the existing object tracking methods ($PF_{tm}M$) which includes integrating particle filter algorithm (PF) with mean-shift tracker (M); (2) improving the existing image segmentation methods which include enhancing Mask R-CNN for object classification; (3) designing a better automated framework of a complete cattle tracking and classification system.

The chapter also describes the dataset collection methods. Also explained in this chapter are the performance measures and evaluation methodologies of the proposed enhanced object tracking methods, improved image segmentation methods, and framework of a complete cattle tracking and classification system.

The implementation of the complete framework for cattle tracking and classification is presented in Chapter 4. The chapter also describes the process involves in training and testing the models for classifying cows to their types.

Presented in Chapter 5 are the results and discussions which include the evaluation results of the enhanced object tracking methods, the improved image segmentation methods, and the proposed framework of a complete cattle tracking and classification system. Finally, Chapter 6 concludes the research.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter reviews several existing methods used in tracking and classifying objects. Since the goal of this research is to provide a framework for better cow tracking and classification systems, a more specific review is presented on ungulate animal tracking and classification frameworks.

Figure 2.1 is an illustration of the global scheme of the research. The scheme consists of three different approaches, namely object tracking, object classification, and ungulate animal tracking and classification frameworks. In order to study the current challenges involved in these approaches, a review of the individual approach is carried out.

Object tracking is reviewed in order to proffer solutions to object-tracking related issues whereby existing object tracking methods can be improved majorly in the cow object-tracking domain. Object classification is reviewed in order to proffer solutions to object-classification related issues whereby existing object classification methods can be improved majorly in the cow classification domain.

A review and explanation of the existing works related to cattle tracking and classification are also carried out.

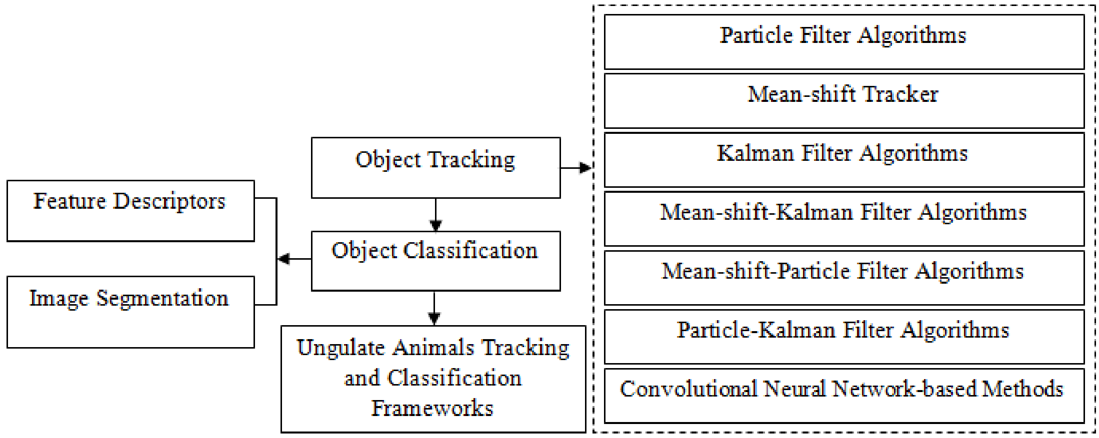

Figure 2.1: Global Scheme of Research

2.2 Object Tracking

The various characteristics and attributes possessed by an object that often leads to changes in posture, shape, size, movement, speed and behaviour have made object tracking a big and interesting research area for researchers even though this comes with a lot of challenges.

It is generally accepted that the changes exhibited by an object determine the method of tracking such object. It is not easy to track non-stationary objects such as human beings or animals; the herculean task involved in object tracking is the appearance of multiple objects mainly with the same characteristics in the same frame.

This is the main reason why enhancement of tracking algorithms and methods is a continuous process. Each of the tracking algorithms and methods has its merits and demerits depending on the data used and problem solved. More importantly, in order to identify the techniques and characteristics used in tracking objects, various popular object tracking methods were studied.

2.2.1 Particle Filter Algorithms

Particle filter algorithms are visual object tracking algorithms that are able to deal with the non-Gaussian problem by using weighted distribution approach to track the trajectory patterns and location of an object based on the object with the highest weight at each time-step during weight distribution which is propagated with time with the use of Bayesian filtering equations.

Particle filters are used in object tracking because it is easy to integrate the algorithms in a divergent environment and they have a simple and fast implementation (Iswanto *et al*., 2019; Iswanto and Li, 2017; Zhang *et al*., 2017). Figure 2.2 shows Simultaneous Localisation And Mapping (SLAM) process based on particle filter.

However, the major limitation of particle filter is the number of particles required as samples. This substantial number of particles results in heavy computational time and reduces the performance of tracking speed, mostly when multiple-object tracking is involved where an immense number of particles is employed for each target object.

In the implementation process of the particle filter algorithms, three notable steps are involved, namely prediction (initialisation step), update (sampling step) and resample (selection step). Initially, particles are scattered in the image across a region of interest and for every time frame, the moving object is identified by the model of its colour and motion.

Higher weight is given to the particles around the moving object and this is achieved by the continuous accumulation of particles around the moving object. However, the overlapping of multiple objects is an issue to particle filter as this causes inaccuracy in tracking an object.

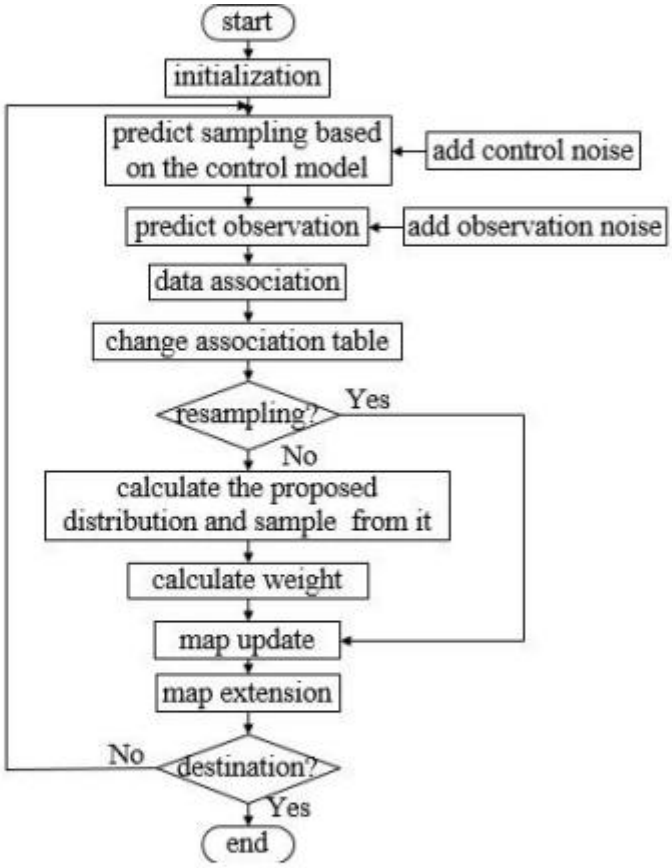

Figure 2.2: Simultaneous Localisation and Mapping Based on Particle Filter (Zhang *et al*., 2017)

2.2.2 Mean-shift Tracker

Mean-shift tracker is among the widely employed methods of visual object tracking. It is widely employed due to its simplicity which enables fast computation and lustiness against occlusion that is partially light, making it more efficient and effective for tracking application (Zhi-Qiang *et al*., 2014). However, it has some limitations which include inability to handle full occlusion and track object that moves at high speed.

Iswanto *et al*. (2019) and Iswanto and Li (2017) in an effort to overcome the existence of illumination variation, non-rigid object, occlusion, non-linear motion and other difficulties involved in automated implementation of tracking in computer vision, proposed a combination of mean-shift (Zhi-Qiang *et al*., 2014) and particle-Kalman filter as a tracking algorithm.

The reason for combining these tracking algorithms is to take the strength of each algorithm as a supplement to the weakness of each algorithm as shown in Figure 2.3. Figure 2.3 shows the different tracking results for full occlusion using MKF (Mean-shift Kalman Filter), MPF (Mean-shift Particle Filter) and the proposed method. Mean-shift is employed as a master tracker when there is no occurrence of occlusion to the object being targeted.

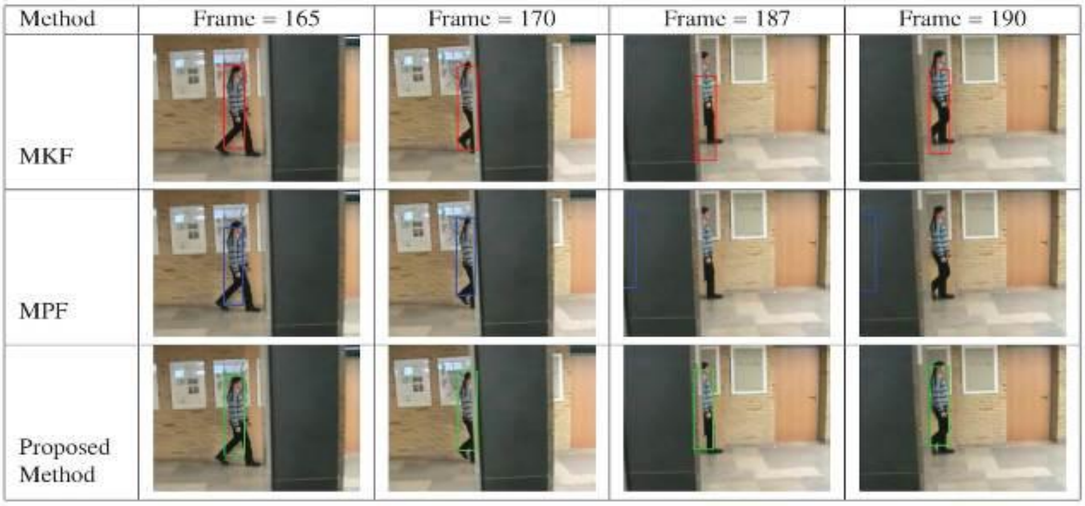

Figure 2.3: Tracking Results for Full Occlusion (Iswanto & Li, 2017)

When there is an occurrence of occlusion or the tracking output of mean-shift is not convincing, the particle-Kalman filter becomes the master tracker in order to improve the tracking outputs. The experimental outputs of their proposed method have shown that the performance is desirable in tracking objects under the existence of illumination variation, non-rigid object, occlusion and non-linear motion.

However, the tracking may not be up to expectation when there is occlusion of the target object by another object of similar colour distribution because the method only employed colour as the tracking feature.

2.2.3 Kalman Filter Algorithms

Kalman filters are visual object tracking algorithms employed by many researchers because they are easy to formulate and implement (Zhang *et al.,* 2017; Rastorguev, 2015). Moreover, their prediction equation is often employed in predicting the location of linearly moving target object occluded by other objects.

Figure 2.4 shows simultaneous localisation and mapping (SLAM) process based on Kalman filter. However, most movements in the real world are mostly non-Gaussian movements which might affect Kalman-based visual object tracking performance when the target object suddenly changes its velocity from linear movement to non-linear movement.

Rodriguez *et al.* (2018) proposed a recursive filtering algorithm such as the Kalman filter for the tracking of objects. The algorithm uses the animal's previous position to estimate the next, thus reducing the search range and the risk of identifying foreign objects as the tracked animals. Also, with the Kalman filter, multiple-object tracking is allowed, and knowledge of the animal shape is not required.

However, the Kalman filter at all times follows the known object that is closest to the predicted position, meaning that it is easily responsive to false positives and false negatives. Besides, the Kalman filter is not reliable to preserve the identity of multiple animals where there is occlusion.

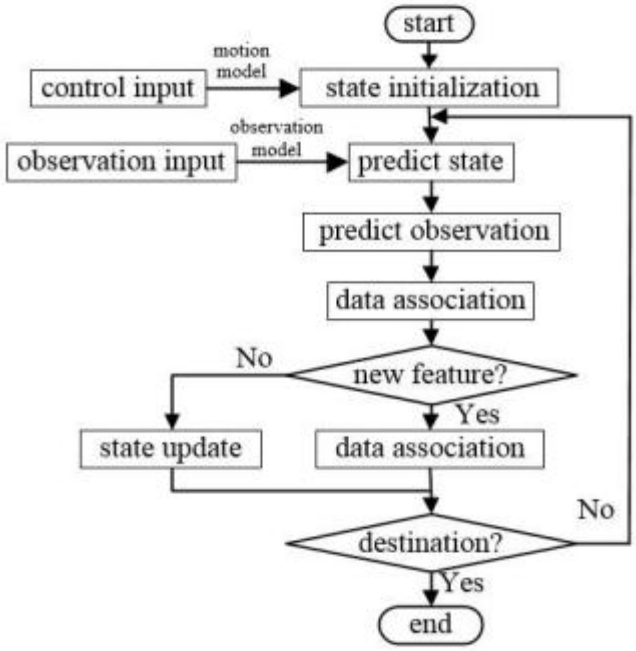

Figure 2.4: Simultaneous Localisation and Mapping Based on Kalman Filter (Zhang *et al.*, 2017)

2.2.4 Mean-shift-Kalman Filter Algorithms

Mean-shift and Kalman filter algorithms were combined in tracking of occluded objects in Zhou and Yan (2014) and in Iraei and Faez (2015). If occlusion should occur, the Kalman filter algorithm is employed to update the estimated location of the target object as shown in Figure 2.5.

The results of their experiment show the robustness of combining mean-shift and Kalman filter in dealing with occlusion. However, the system could only deal with linear moving objects since the Kalman filter is employed as the predictor when the object is occluded.

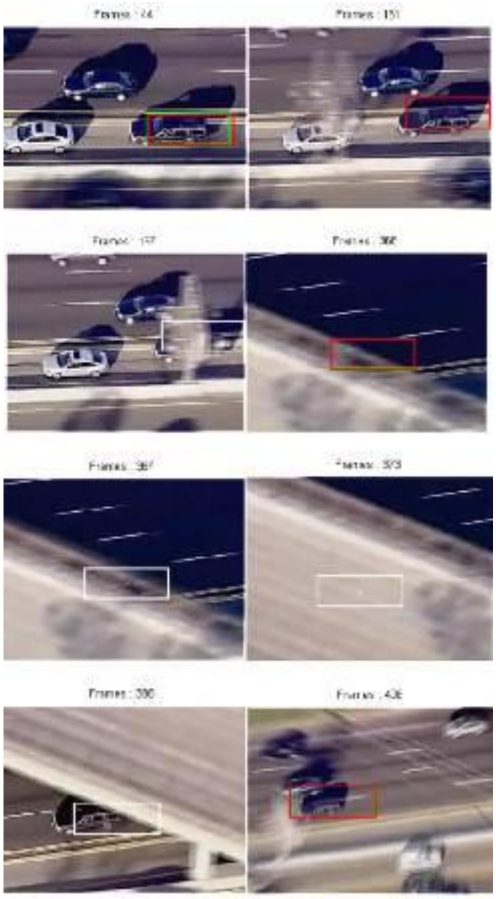

Figure 2.5: Result of the Experience in Presence of Partial and Full Occlusion (Iraei & Faez, 2015)

2.2.5 Mean-shift-Particle Filter Algorithms

Iswanto *et al*. (2019) and Iswanto and Li (2017) proposed a tracking algorithm that can overcome several problems encountered in object tracking such as tracking non-stationary objects, non-linear motion and occlusion. They achieved this by combining mean-shift and particle filter as shown in Figure 2.6.

The purpose of combining these methods is to utilise the strength points possess by each algorithm. In their proposed method, mean-shift is employed when there is no occlusion of the target object, and the particle filter is employed should there be any occlusion.

The results of the experiment that they have performed using the proposed method show that the performance is desirable. However, because their proposed method recognised only colour as the tracking feature, other objects with similar colour distribution may occlude the target object, thereby rendering the system useless.

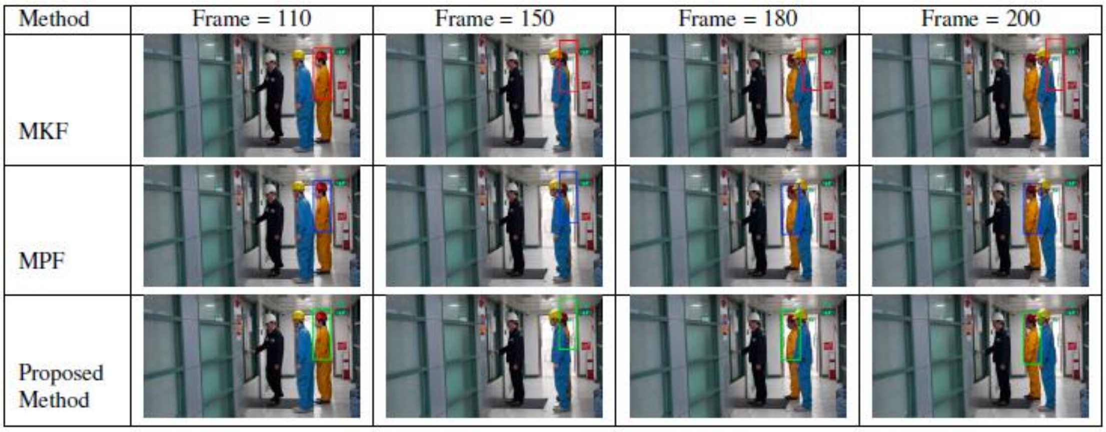

Figure 2.6: Tracking Results for Severe Partial Occlusion (Iswanto & Li, 2017)

Qiao and Yu (2014) proposed tracking techniques by combining particle filter with mean-shift. In their various techniques, mean-shift is utilised to shift each particle towards a local maximum that is close in order to enhance the performance of the particle filter.

This is to ensure the obtainability of a condensed particle set with fewer particles, thereby decreasing the divergence problem. Although the number of particles is reduced, the speed of tracking is not as high as the speed of the target object due to the iteration requirement of mean-shift for each particle.

Moreover, due to the application of the mean-shift algorithm to each particle, the particle set is commonly too concentrated; this condition may affect tracking performance especially when the system needs to track the object re-appearance after severe or full occurrence of occlusion.

2.2.6 Particle-Kalman Filter Algorithms

Particle-Kalman filters are visual object tracking algorithms that combine two conventional algorithms. The general idea of the particle-Kalman filters is based on the phenomenon of most instances in video tracking as shown in Figure 2.7. One of the instances is where the target object motion in global view is in Gaussian motion or linear motion whereas non-linear motion is commonly limited in the local view.

Therefore, the Kalman filter is utilised to perform estimation of global motion in global view and particle filter is utilised to perform local estimation to deal with non-linear motion in local view. In the particle-Kalman filter, the linearity of the Kalman filter is enforced for every particle to obtain more condensed particles to deal with the non-linearity problem.

Through this, the number of particles can be significantly reduced without any additional iteration and this will significantly increase the speed of the system. However, it takes immense time to implement the linearity of the Kalman filter for every particle to obtain a more condensed particle to deal with the non-linearity problem (Iswanto *et al.*, 2019; Iswanto & Li, 2017).

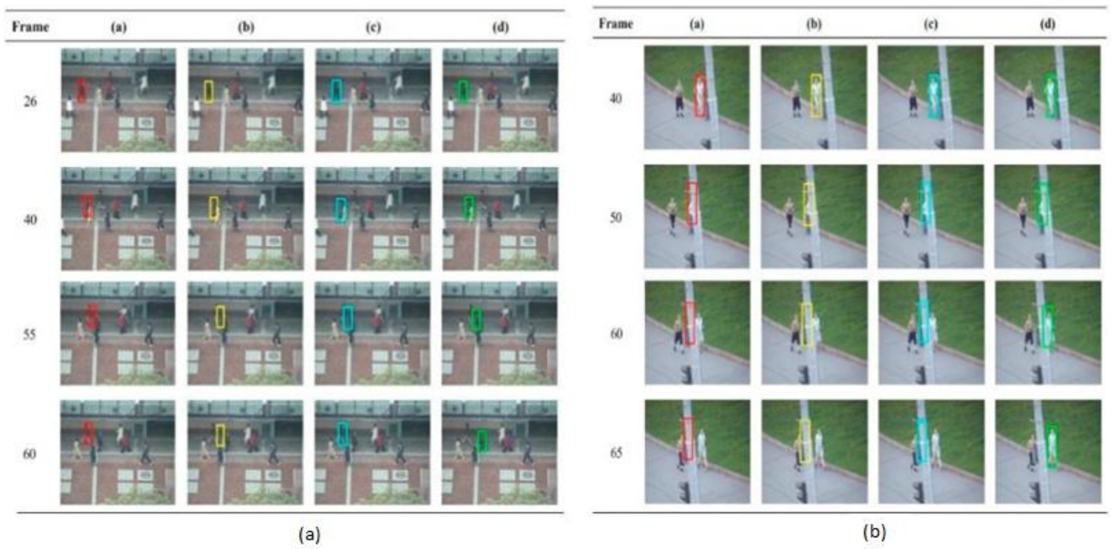

Figure 2.7: Tracking Results for Objects under (a) Similar Colour Interference for Subway Video Sequence and (b) Similar Colour Interference and Full Occlusion for Jogging Video Sequence (Iswanto *et al.*, 2019)