A POND-SURFACE-BASED BIOFLOC-FARM HEALTH-MONITORING SYSTEM FOR AFRICAN CATFISH USING DEEP LEARNING METHODS

NIZAM BIN NORDIN

FACULTY OF COMPUTING AND INFORMATICS
UNIVERSITI MALAYSIA SABAH

2022

A POND-SURFACE-BASED BIOFLOC-FARM HEALTH-MONITORING SYSTEM FOR AFRICAN CATFISH USING DEEP LEARNING METHODS

NIZAM BIN NORDIN

THESIS SUBMITTED IN PARTIAL FULFILLMENT FOR THE DEGREE OF BACHELOR OF COMPUTER SCIENCE WITH HONOURS (NETWORK ENGINEERING)

**FACULTY OF COMPUTING AND INFORMATICS
UNIVERSITI MALAYSIA SABAH**

2022

NAME : NIZAM BIN NORDIN

MATRIC NUMBER : BI18110097

TITLE : A POND-SURFACE-BASED BIOFLOC-FARM HEALTH-MONITORING SYSTEM FOR AFRICAN CATFISH USING DEEP LEARNING METHODS

DEGREE : BACHELORS OF COMPUTER SCIENCE WITH HONOURS (NETWORK ENGINEERING)

VIVA'S DATE : 17th JANUARY 2022

CERTIFIED BY;

1. **SUPERVISOR** Signature

DR. ERVIN GUBIN MOUNG

DECLARATION

I hereby declare that the materials in this thesis is my own except for quotations, equations, summaries and references, which have been duly acknowledge.

17 JANUARY 2022

NIZAM BIN NORDIN

BI18110097

ACKNOWLEDGEMENT

First and foremost, I would like to thank my God for allowing me to gain inspiration and giving me the strength and motivation to allow me to complete this project after such long trials and tribulations I had to face, I have to admit after countless sleepless nights I am so pleased to finally complete this report.

Secondly, I would like to acknowledge and express my gratitude towards my supervisor, Dr. Ervin Gubin Moung for the advices and ideas which helped me a lot during my journey toward completing this project for without it, I doubt this report would be completed. I would like to acknowledge to my friends and fellow classmates who helped shared information, inspiration and knowledge that guided me in completing this project.

Finally, I owe my absolute gratitude towards my family especially towards my parents who supported me mentally and emotionally throughout the course of doing this project for without it, I would have permanently given up on completing this project. Their life advices helped me realize that this is just temporary and life is a long and difficult road.

Nizam Bin Nordin

17 JANUARY 2022

ABSTRACT

Biofloc fish farming is becoming a popular aquaculture tool for fish farmers to farm for fishes as nutrients required for the fishes to feed on are produced naturally within the self-contained ecosystem especially in locations where naturally occurring body of water is difficult to come by and land is limited and expensive. The objective of this research is; (i) To model a standard training database selection criterion on the fish behaviour in relation to their health, (ii) To develop an algorithm that can determine the health of the fish based on the behaviour of the fish, (iii) To implement and test the fish behaviour detection algorithm in a Biofloc fish pond to validate and evaluate the effectiveness of the system. The research methodology stages are; (i) Acquisition of daily recordings of the surfaces of fish ponds and their respective daily health, (ii)Formulating and developing a fish health monitoring algorithm based on their behaviour from the surface, (iii) Selection of training data based on the recordings and health,(iv) performance metric evaluation, (v) Assessment of experimental results of the Biofloc Fish Health Monitoring System On Pond Surface algorithm, and (vi) final report and project wrap-up. The expected outcome of this is an artificial intelligence system that can determine the well-being of the fishes in a Biofloc pond based on the behaviour of the fish that can be observed from the simulated environment of the pond. The proposed system can potentially reduce the already heavy workload of Biofloc fish farmers.

ABSTRAK

Penternakan ikan biofloc menjadi alat akuakultur yang popular bagi petani ikan untuk bertani kerana nutrien yang diperlukan untuk ikan yang dimakan dihasilkan secara semula jadi dalam ekosistem mandiri terutama di lokasi di mana badan air yang wujud secara semula jadi sukar didapat dan tanah adalah terhad dan mahal. Objektif penyelidikan ini adalah; (i) Untuk memodelkan kriteria pemilihan pangkalan data latihan standard mengenai tingkah laku ikan yang berkaitan dengan kesihatan mereka, (ii) Untuk mengembangkan algoritma yang dapat menentukan kesihatan ikan berdasarkan tingkah laku ikan, (iii) Untuk melaksanakan dan menguji algoritma pengesanan tingkah laku ikan di kolam ikan Biofloc untuk mengesahkan dan menilai keberkesanan sistem. Tahap metodologi penyelidikan adalah; (i) Perolehan rakaman harian permukaan kolam ikan dan kesihatan harian masing-masing, (ii)Merumuskan dan mengembangkan algoritma pemantauan kesihatan ikan berdasarkan tingkah laku mereka dari permukaan, (iii) Pemilihan data latihan berdasarkan rakaman dan kesihatan,iv) penilaian metrik prestasi, (v) Penilaian hasil eksperimen algoritma Sistem Pemantauan Kesihatan Ikan Biofloc Pada Permukaan Kolam, dan (vi) laporan akhir dan penutupan projek. Hasil yang diharapkan dari ini adalah sistem kecerdasan buatan yang dapat menentukan kesejahteraan ikan di kolam Biofloc berdasarkan tingkah laku ikan yang dapat diperhatikan dari lingkungan simulasi kolam. Sistem yang dicadangkan berpotensi mengurangkan beban kerja petani ikan Biofloc yang sudah berat.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

In recent years, aquaculture has contributed to the global fisheries where it could be proven by a 21.1% increase in the year 2016 when compared to 2000 making it 46.8%(D. Li et al., 2020). In recent years, Biofloc technology has been on the rise and shows promising potential in aquaculture as it offers a cost-effective, sustainable and eco-friendly method for producing fishes. Biofloc fish farming relies on the growth of heterotrophic bacteria through additional carbon sources and the regulation of the C/N ratio of the aqueous medium. Under specific conditions, these bacteria will consume the waste materials from the water and converts them to proteins which are also form bioflocs (microbial flocs) these are sources of food for the fishes living in that habitat (J. Li et al., 2021). One main issue is that Biofloc fish farming demands for daily visits to each Biofloc pond daily and evaluate the condition of the fish as well as the pond which is labour intensive (Deb et al., 2020).

The global demand for fish is increasing annually with the capture and production of marine resources estimated to be between 6.5 and 14.4 million tons by Indonesia and China respectively in 2014 alone. In aquaculture, the use of fish meal (FM) will not be sustainable in order to meet the increase in growth for the aquaculture industry(Adeoye et al., 2021). Indonesia is the leaders in Biofloc technologies fisheries with 40.2% of it's 10.7 million tons of its fish supplies comes from farmed fishes(Tran et al., 2017). The workload required to maintain the well-being of the fishes are high and requires daily care and a human is required to manually observe each pond and adjust the water concentration accordingly (El-Sayed, 2020b). This requires intense human care and knowledge of the fishes, data collected from already existing Biofloc farms can be recorded and a computer can be trained to do these things automatically using machine learning. This could potentially reduce the workload of the fish farmers as they don't have to manually visit each pond to observe the condition and well-being of the fishes(Terayama et al., 2019).

1.2 Motivation

Aquaculture is nothing new in today's world, but a new method of aquaculture known as Biofloc fish farming. Rural areas would stand to benefit the most from this as they do not require to be near any naturally occurring body of water. Fishes such as tilapia that are commonly used in biofloc can play a significant role for rural areas or still developing countries(El-Sayed, 2020a).

There is an abundance of resources on animal behaviour detection artificial intelligence, but most of the applications do not utilise this for monitoring the health of the fish. Because of this, there is a high demand for automated fish monitoring system.(Terayama et al., 2019)Biofloc fish farms are labour intensive and require constant care in monitoring and maintaining their health. A fish can express their welfare externally through their behaviour and goes through various metabolic changes(Zion, 2012). If an AI is trained to look for these symptoms and signs, it would significantly reduce the workload for the fish farmers.

1.3 Problem Statement

Currently, there are limited articles on fish behavioural algorithm. However, these articles share similar problems statements. It was found that (i) there is no publicly available database for fish behaviour as ground for truth for fish health monitoring investigations. Other than that, (ii) there is no standard algorithm that could determine health of the fish based on its behaviour such as position of the fish in the water and its health. Finally, (iii) there is no standard method and test-bed published method for evaluating the effectiveness of fish health monitoring algorithm. Thus, a comparison between any fish health in relation to their behaviour monitoring system is difficult to make.

1.4 Hypothesis

Research cost of equipment would be reduced as only a standard camera is required along with an integrated computer system.

1.5 Research Question

1. What is the most efficient approach (in terms of detection and resource utilisation) to use fish behaviour data, specifically their state in the water (upside down, upright, vertical) in relation to their health?
2. Since there are different species of fishes, would the catfish be a suitable fish species to use in this research?
3. How can the behaviour of the fish determine the health of the fish?

1.6 Research Objective

1. To model a standard training database selection criterion based on the fish behaviour data and how it relates to the fish's health. (This objective is mapped with Problem Statement 1)
2. To formulate and develop a fish behaviour detection algorithm that can determine the well-being of the fish based on its behaviour. (This objective is mapped with Problem Statement 2)
3. implement and test the fish behaviour detection algorithm in a Biofloc fish pond to validate and evaluate the effectiveness of the system. (This objective is mapped with Problem Statement 3)

1.7 Project Scope

This project is conducted with the main idea that Biofloc fish farmers could potentially reduce their workload on looking after fishes. The scope covered in this project is:

1. Monitor the condition (position) of the fish such as if the fish is vertical or horizontal, and no other behaviours such as swimming patterns or feeding appetite

1.8 Organization of Project

The following chapters are the flow of organization of project.

- Chapter 2 Literature Review
  - In this chapter, existing systems will be reviewed in terms of similarity in order for a comparison to be made between its purpose and function.
- Chapter 3 Methodology
  - This chapter discusses the steps taken at each stage in the methodology during the process of development of this project. Flow charts, activities and milestones will also be included in this chapter.
- Chapter 4 System Analysis and Design
  - This chapter will give an overview of the system design including the entity relationship diagram and the flowchart of the system as well as the UI design of the system in order to help in visualising the system .
- Chapter 5 Implementation
  - In this chapter, the implementation of the system will be discussed. The software and tools required will be stated and shown and explain how the algorithm is put together.
- Chapter 6 Experiment and Results
  - This chapter will discuss the experiment settings and environment in order to test the tracking algorithm as well as the effectiveness of the fish health classification algorithm.
- Chapter 7 Conclusion
  - This chapter will summarize the development of the system. This chapter will also discuss the achievements and future works that can be done to further improve on this project.

1.9 Summary

This chapter provides the idea and motivations that inspired this project. It details the problem statement as well as further explains the questions that is hoped to be answered through this project and the project scope.

CHAPTER 2

LITERATURE REVIEW

2.1 An Overview of Deep Learning Algorithm

A subcategory of machine learning also known as deep neural networks is a network model with nodes also known as neurons connected together with multiple parameters and layers to filter out from between the input and output. Deep learning (DL) is according to the neural network architecture. The advantage of DL is that it provides an automated learning based on features and representation in a hierarchical way. Because of this, DL is a powerful and robust method of machine learning when compared to more traditional method of machine learning. DL is suitable for feature extraction and alteration process whereby each layer would perform a process more complicated compared to the layer before making it an ideal way in dealing with higher complexity tasks (Sharma et al., 2021).

DL is currently the most widely used form of machine learning, DL is such a widely used form of deep learning that it is frequently regarded as a universal learning method. DL is beneficial in various applications in which machine intelligence would be more preferable compared to human intelligence. For example, navigations on extra-terrestrial locations such as Mars where human intelligence is absent and a complete information of the surrounding is almost non-existent, DL would be more suitable to perform such task(Sharma et al., 2021).

Fish behaviour detection has been used in many different scenarios such as for feeding, detecting anomalies and effects of noise on their behaviour. Computer vision technology has been utilised in aquaculture to monitor the well-being and behaviour and can be divided into two main categories, namely direct and indirect methods. Direct methods utilise what is directly visible on the fish such as the behaviour, swimming pattern and/or shape of fishes whereas indirect methods measure what cannot be seen directly on the fish such as excess feeds and fish appetite(D. Li et al., 2020).

2.1.1 Recurrent Neural Network (RNN)

Artificial neural network is a mathematical model that is essentially is a group of connected nodes also called neurons. The recurrent part of RNN stems from how RNN works, according to (Grachev et al., 2017), "The idea of using the recurrent part is to calculate the next state by the input state and the next character of the current word, that is, in the language of finite state machines, to make a transition to the new state labelled with the next character". This basically means that the current layer will calculate the current input to produce and output which will then be the input for the next layer. RNN contains multiple layers, namely input layer, transition tensor, normalization layer and the output layer (Grachev et al., 2017).

The attraction of RNN stems from the various Long Short-Term Memory (LSTM) especially in applications that is sequential in nature(Y. Wang et al., 2020). The input layer of RNN consists of |S| + |Q| neurons. The following input characterize is determined by the first |S| neurons, the output values at these neurons can be measured by an |S|-dimensional vector, so that its vector dimension will correspond to the size of alphabet. The components of vector c that corresponds to the index of the input would not equal to zero. After passing through the (z-1)-th character at the (z-1)-th expression, the final |Q| neurons will decide the state. If z=1, then the transformation is made using the first character of the expression, and the output values of neurons in the input layer are calculated by N-dimensional vector I, which has zeros at all but the first component.

2.1.2 Convolutional Neural Network (CNN)

CNN determines that a network uses a convolution computation model as its name suggests. CNN has a few layers namely associate input layer and associate output layer as well as multiple hidden layers. It has also been shown that CNN became popular due to its simplicity in its architecture and wide usage in all fields, CNN thus became one of best computer vision technique(Kranthi Kumar et al., 2021).

The core of a CNN is the convolutional layer which consists of 3 different elements namely arrangements of space, sharing of parameters and connectivity of locals. In a convolutional layer, the boundaries consist of various channels/portions of different parameters. When an input is passed through these channels, each channel is convolved across different dimensions and height of data volume processing a certain area of an image each time it passes creating a 2-dimentional enactment guide of said channel. Basically, the process of CNN is taking an input and mapping out its feature and sub feature until the required output is obtained and outputting a prediction(Kranthi Kumar et al., 2021).

2.2 Machine Learning Applications in Fish Health Monitoring System

According to Daoliang Li, "The wide application of computer vision technology provides an effective means to perform real-time, automatic, and contactless research" (D. Li et al., 2020). This provides a goods method in monitoring fish behaviour, feeding patterns of fishes are also part of the behaviours of fishes and using cameras to determine the feeding behaviour. When different feeding intensities are displayed, the conditions of each fishes also differ such as colour, shape, texture and the consistency of the fish feeding at any given time. Utilising a novel infrared reflection (IREF) system allows the three-dimensional tracking of fish which allows for a lower hardware cost and less computation when compared to a general stereo vision system (D. Li et al., 2020).

Another method proposed by (J. H. Wang et al., 2020). involves three different stages in tracking the behaviour of fishes namely state definition, tracking/encoding, and decoding dynamic time warping (DTW). The state definition stage involves the utilisation of any deep learning artificial intelligence models in order to perfect the object detection for multiple parts of the moving fish, key parts of the fish is recorded such as the fins, eyes and main body in order to determine the current state of the fish. The tracking/encoding stage occurs when the state of the fish is determined, the tracking of the fist posture is then used to signal the start of the tracking process. In the decoding stage, the state of the fish is compared to 46 different behaviour templates that are suggested by experts on aquaculture experts in order to make a final decision. Positional relationship between key parts of the fish is used to determine the posture. By using deep learning techniques as compared to image processing approaches, the time taken to identify the body parts of the fish is significantly improved despite the messy background. A directed cycle graph (DCG) is then created.(J. H. Wang et al., 2020). In order to continuously capture images of the moving fish, a Faster-rcnn was used and trained with 2986 images of multiple fishes containing body eight key body parts pre-labelled. An object detector provides information either directly or indirectly, this approach requires only a small amount of labelled data for training the detection of fish body parts compared to training Recurrent Neural Network (RNN), Long-short Term Memory (LSTM) and any other methods would require vastly large amount of training data (J. H. Wang et al., 2020). The system proposed is almost suitable to be applied in Biofloc farms as the states of the fish can be determined and when training data is applied the AI can determine the condition and well-being of the fish (Taheri-Garavand et al., 2019).

Besides using the state of the fish, (Anas et al., 2020) uses motion trajectories in ubiquitous environments in order to detect abnormal fish behaviour. Their method involves a pre-processing phase where Multi-Scale Retinex color enhancement algorithm (MSR) is used to improve the quality of the image in the case of unclear water. Then the images are loaded into a trained YOLO model to detect the fishes. The trajectories of the fishes are then extracted from the tracked fishes. The fish's behaviour are then determined and if it is abnormal, farmers will be sent an alert. In the pre-processing phase, the main purpose is to enhance the quality of the image in order to get better results in the processing phase. The images are initially processed by Single-Scale Retinex (SSR) where the logarithm from the original images are subtracted from the gaussian filter of the same image in equation (i) where F(x, y, a) is the Gaussian filter image and I(x, y) is the original image. After that, the image are then sent to MSR in equation (ii) where R(MSR) is the enhanced images where X is the number of scales.

$$\log\big(I(x,y)\big) - \log\big(F(x,y,a) \times I(x,y)\big) = r(x,y) - \text{(i)}$$

$$\sum_{x=1}^{x} \log\big(I(x,y)\big) - \log\big(F(x,y,a) \times I(x,y)\big) = R(MSR) - \text{(ii)}$$

During the processing phase, YOLO algorithm is used in order to detect the fishes as it can detect fish with accuracy, this works by utilizing a whole image as a single neural network. The network then splits each image into different regions and calculates the probability for each region. This method works better than Fast r-cnn as it constructs a full picture context in view at test time. The fishes are then tracked on each video frames and the trajectories of the fish are extracted which will determine if the fish is behaving normally or abnormally using classification method. 2000 images of goldfish were used to create a model. By utilising MSR, images can be enhanced which provides useful when the conditions of the water are not ideal such as the condition of most Biofloc farms. However, the number of images required to create a model is large the computational power required to calculate the trajectory of larger number of fish in order to determine behaviour would be equally large. The behaviour of the fish could also determine the health of the fish

For fish tracking, one method proposed by (Zhao et al., 2019) composes of three parts which are a charge-coupled device camera, experimental fish tank, and a computer. The camera used was Hikvision MV-CA030-10GC industrial camera and directly placed above the experiment tank such that the swimming area of this fish can be captured through the camera. The computer used was a personal computer (Intel(R) Core (TM)-i5-4210U CPU @ 1.70 GHz 2.39 GHz) running Matlab R2017b software. Their research setup is in a laboratory environment with a fixed illumination and a stable environment allowing for background substitution in order to obtain the moving target which is the fish. (Zhao et al., 2019).

2.1.1 Deep Learning Applications in Fish Health Monitoring System

The work done by (D. Li et al., 2020) utilises computer vision technology in which real-time, automatic and contactless research in which fish behaviour is monitored based on the feeding pattern of the fish. Although the feeding patterns of the fish could indicate their health, the challenges for applying this approach for the current research is that there is no readily available training data and models for fish feeding pattern on catfishes. Another challenge for this approach when applying to a Biofloc farm is that the usual conditions for Biofloc farms are usually dark green and filled with bacteria so an IREF system would not be suitable in this situation.

The method used by (Anas et al., 2020) uses motion trajectories in ubiquitous environments. In their method, pre-processing is involved using Multi-Scale Retinex colour enhancement algorithm in order to improve the quality of the image in the case where the water environment is not clear. This could be applied to this project because Biofloc fish farms are usually unclear. The trained YOLO model used in this method could also be applied to this project. Altough, a similar challenge is presented as there is no readily available database for trajectories of fishes in relation to their health.

The most applicable method to this project would be the method done by (J. H. Wang et al., 2020), which involved three different stages in tracking behaviour of the fish.

Usage of deep learning artificial intelligence models are used to detect and record key parts of the fish, with the key parts of the fish detected and recorded a directed cycle graph can be created thus the position and orientation of the fish can be determined. This method is can be used for this project in order to determine the orientation of the catfish which is an indicator of their health. Like previously mentioned before, the limitations of this approach is that the visibility is key in order to get a clear image to determine the orientation of the fish.

2.3 Depth and Analysis

In this chapter, existing systems are explored in order to get a clear understanding of the challenges that needs to be overcome for this project. The comparison of existing system is also made in order to determine the best aspect of each approach in order to apply it to this project.

Table 2.1 summarizes the different DL approach that other researchers have used for fish behaviour detection. There are various approaches that researches have used to achieve this, and their methods vary in terms of the scope for their research, (D. Li et al., 2020) reviewed Back propagation neural network model used in fish feeding behaviour.(J. H. Wang et al., 2020) used Gaussian mixture model on the fish state to detect anomalous behaviour in fishes.(Anas et al., 2020) used profound learning AI models on fish trajectories in determining abnormal fish behaviour.

Table 2. 1 Summary of DL approach for Fish Behaviour Detection

<table>
<tr><th>Author</th><th>Method</th><th>Name of database</th><th>Database criteria</th><th>Machine learning model</th><th>System performance</th></tr>
<tr><td>Daoliang Li, Zhenhu Wang, Suyuan Wu, Zheng Miao</td><td>Fish feeding behaviour</td><td>Own Database</td><td>Difference in frame/texture of feeding fish</td><td>Back propagation neural network model</td><td>Success rate:97%;
Classification success rate: 98.51%,
Detection accuracy:90.8%</td></tr>
<tr><td>Jung-Hua Wang, Shih-Kai Lee, Yi-Chung Lai, Cheng-Chun Lin, Ting-Yuan Wang, Ying-Ren Lin, Te-Hua Hsu</td><td>Fish state</td><td>Own Database</td><td>2986 images containing multiple fish with pre-labelled key body parts</td><td>Gaussian mixture model</td><td>Accuracy = 92.8%% (Faster-rcnn as Object Detector)
Accuracy = 88.0% (YOLO-v4 as Object Detector)
Accuracy = 81.8% (EfficientDet-D1 as Object Detector)</td></tr>
<tr><td>Omar Anas, Youssef Wageeh, Hussam El-Din Mohammed, Ali Fadl, Noha ElMasry, Ayman Nabil, Ayman Atia</td><td>Motion Trajectories of fish</td><td>Google X (Currently X Development LLC)</td><td>Not Specified</td><td>Profound learning AI models</td><td><table><tr><th>Fish trajectories</th><th>Average Accuracy</th></tr><tr><td>5 second</td><td>83.3%</td></tr><tr><td>10 second</td><td>87.8%</td></tr><tr><td>15 second</td><td>82.8%</td></tr><tr><td>20 second</td><td>83.0%</td></tr></table></td></tr>
</table>

CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter introduces the stages in the development of the project. Each stage will describe a step-by-step process in the development of this project.

3.2 Description of Methodology

There are six stages in this research.

Stage 1: Acquisition of daily recordings of the surfaces of fish ponds and their respective daily health. (To achieve Research Objective 1.)

- The database will include daily recordings of this fishes from the side of the fish tank as training data
- Daily health records of the according video and daily recordings taken by experts are also taken
- One camera will be position on the side of the fish tank

Figure 3. 1 Setup of the camera and fish tank for data collection