Ecological Informatics 78 (2023) 102287

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Pattern detection and prediction using deep learning for intelligent decision support to identify fish behaviour in aquaculture☆

S. Shreesha [a], Manohara M M Pai [a,*], Radhika M. Pai [b,*], Ujjwal Verma [c,*]

[a] Department of Information and Communication Technology, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India
[b] Department of Data Science and Computer Applications, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India
[c] Department of Electronics and Communication Engineering, Manipal Institute of Technology Bengaluru, Manipal Academy of Higher Education, Manipal, India

ARTICLE INFO

Keywords:
Computer vision
Pattern analysis
Auto-encoder
Water quality monitoring
Fish behaviour analysis
Outlier detection
Aquaculture

ABSTRACT

The United Nations' Sustainable Development Goals (SDGs) underscore aquaculture as a sustainable practice. However, many countries rely on ineffective conventional methods to monitor the aquaculture. Fish behaviour monitoring systems lack an integrated strategy, often focusing on underwater footage analysis or water quality data. Additionally, limited research exists on using learning algorithms to identify sensor data anomalies, limiting the effectiveness of such systems. This study proposes a comprehensive and effective aquaculture monitoring system to address these issues. The work introduces a novel method integrating fish motion analysis with water quality metrics. The study also presents a learning system that anticipates potential water quality parameters. The proposed model has the potential to significantly enhance aquaculture management procedures and contribute to the UN's SDGs through its ability to forecast water quality parameters and detect abnormal behaviour without requiring training detection and tracking algorithms.

The study employs a large aquarium to capture water quality parameters and fish movement patterns. The study develops an FPGA-based water quality monitoring system and a prediction model to forecast potential water quality metrics. Among multiple LSTM models, the bi-directional LSTM performs well, achieving an RMSE score of 1.01 for water quality prediction. Subsequently, an outlier detection system is applied to identify unusual water quality parameters, triggering the behaviour analysis model. The behaviour analysis model uses an auto-encoder-based reconstruction approach to detect the frantic patterns in fish. The suggested behaviour analysis model has a f1-score of 0.68, demonstrating its reliability in contrast to conventional methods. The developed system is also tested on aquaculture site videos and includes a novel dataset of *Sillago sihama* with annotations for frantic behaviour patterns. The system works effectively as a decision support system for aquaculture.

1. Introduction

Patterns represent regularities in data and find applications across various domains, including robotics, image processing, signal processing, speech analysis, document analysis, and medical image processing (Koskinopoulou et al., 2021; Girisha et al., 2021; Li et al., 2021; Chan et al., 2021). Recent years have witnessed a resurgence of interest in examining patterns in the context of aquaculture applications (Yang et al., 2021; Liu et al., 2021; Thai et al., 2021; Pache et al., 2022; Bergmann et al., 2019; Butail et al., 2015; Abinaya et al., 2021; Jalal et al., 2020; Liu et al., 2023). Analyzing patterns in aquaculture plays a pivotal role in discerning fish behaviour and identifying trends in water quality parameters. Information about fish behaviour and water quality parameters is imperative for monitoring fish health. However, the dearth of reliable techniques for monitoring fish behaviour and water quality hampered the progress of aquaculture. Also, aquaculture faces serious threats posed by unforeseen variations in water quality parameters. It leads to elevated stress levels among fish and alterations in their normal behaviour. The lack of prompt access to aquaculture data impedes fishermen from conducting in-depth analyses of fish behaviour, which, in turn, hinders their ability to mitigate potential issues. Given these challenges, pattern analysis models may be used to monitor water quality and fish behaviour, eventually improving aquaculture harvest by reducing fish mortality.

☆ Research supported by the Department of Science and Technology, India.
* Corresponding authors.
E-mail addresses: mmm.pai@manipal.edu (M.M.M. Pai), radhika.pai@manipal.edu (R.M. Pai), ujjwal.verma@manipal.edu (U. Verma).

https://doi.org/10.1016/j.ecoinf.2023.102287
Received 29 May 2023; Received in revised form 31 August 2023; Accepted 31 August 2023
Available online 7 September 2023
1574-9541/

As fish rely on water for respiration, the water quality significantly impacts their well-being. Therefore, monitoring the water quality parameters within aquaculture systems plays a pivotal role in enhancing fish yield. The emergence of intelligent systems, driven by recent advancements in sensors, controllers, and wireless communication technologies, has found applications in various fields, such as agriculture, surveillance, and environmental monitoring (Yin et al., 2021; Saini et al., 2021; Alam et al., 2021; Nguyen et al., 2021; Lin et al., 2021; Adiono et al., 2021; Chavan et al., 2018; Gokulanathan et al., 2019; Vasudevan and Baskaran, 2021; Singh et al., 2022). In this regard, fishermen can monitor the water quality using sensors and controllers. Existing literature demonstrates that the most frequently observed water quality parameters include temperature, pH, Dissolved Oxygen (DO), conductivity, total dissolved solids, and salinity (Lin et al., 2021; Adiono et al., 2021; Komarudin et al., 2021; Duangwongsa et al., 2021; Shreesha et al., 2021a; Singh et al., 2022). These studies use settings such as aquariums, ponds, and bio-floc systems. Also, the literature uses connectivity technologies like Wi-Fi (Duangwongsa et al., 2021; Rashid et al., 2021; Gopi and Naik, 2021), LoRa (Adiono et al., 2021) and Bluetooth (Putra et al., 2021) to link these devices to the internet. Authors commonly employ the ThingSpeak IoT analytic platform to visualize and analyze the data (Pasika and Gandla, 2020; Simitha and Raj, 2019). While current research predominantly focuses on real-time monitoring of water quality metrics, this perspective may be inadequate for aquaculture as it necessitates proactive measures. To address this, predictive models can simulate the dynamic behaviour of water quality metrics (Zheng et al., 2023). Learning algorithms can forecast water quality parameters for the upcoming hour and detect outliers, affording fishermen ample time for preventive actions. Thus, predictive monitoring aids in the early detection of anomalies and mitigates risks. It further enhances the sustainability of aquaculture. However, few publications in the literature evaluate sensor data and predict anomalies using learning algorithms. Specifically, deep learning algorithms capture intricate correlations within water quality parameters, enhancing the predictive accuracy. The current study bridges this gap by introducing a Field Programmable Gate Array (FPGA)–based sensor system for monitoring water quality parameters and integrating it with learning algorithms to predict water quality metrics and identify outliers (deviations from the norm) for the next time step. Consequently, fishermen can make informed decisions about the ecosystem and implement safety measures.

In addition to monitoring water quality parameters, extracting motion information from underwater videos can provide valuable insights into accurate behaviour analysis. Fish adapt their behaviour in response to minor changes in their ecosystem. For example, fishes swim rapidly when the temperature rises, and as DO levels drop, they swim towards the surface to access oxygen from the air. Studying these behavioural patterns can be highly beneficial for fish farm management. Several studies in the literature have attempted to analyze fish behaviour (Butail and Paley, 2012; Shreesha et al., 2020; Qian et al., 2016; Lee et al., 2010; Saberioon and Cisar, 2016; Rodriguez et al., 2015; Bhaskaran et al., 2019; Papadakis et al., 2012; Shreesha et al., 2021b; Hulse et al., 2022). The experimental research carried out in the literature commonly uses tilapia and zebrafish species (Xu et al., 2006; Xia et al., 2016; AlZu'bi et al., 2015; Bai et al., 2018; Bhaskaran et al., 2019; Lee et al., 2010; Pedersen et al., 2020). The literature has explored various well-known fish behaviours, including fleeing, surface swimming, changes in swim speed, ambient light effects, and collective behaviour (Butail et al., 2015; Pinkiewicz et al., 2011; Xu et al., 2006; Papadakis et al., 2014; Shreesha et al., 2020; Long et al., 2020; Shreesha et al., 2022). The two broad categories of fish behaviour are normal behaviour and anomalous behaviour. Detecting abnormal behaviour is challenging due to a lack of data, and interpreting fish behaviour is often subjective. Thus, defining patterns for computer vision models presents a demanding task. There prevails literature that explores computer vision-based techniques for fish behaviour analysis (Saberioon and Cisar, 2016; Rodriguez et al., 2015; Bhaskaran et al., 2019; Shreesha et al., 2020; Xia et al., 2016). A key concern in many behaviour analysis techniques is their dependency on training object detection and tracking algorithms, which can constrain the effectiveness of behaviour analysis models. The deformable nature of the target objects (fish) and poor under water imaging compounds the challenge of training object detection algorithms on fishes (Yu et al., 2023; Sun et al., 2022). Also, to track the fish, each detection in the current frame must be associated with its corresponding detection in the previous frame. Regardless, it is demanding since fish frequently alter their swim direction, have a similar colour, and are prone to occlusion (Li et al., 2022; Shreesha et al., 2023). The study addresses this gap by proposing a novel approach to fish behaviour analysis that does not depend on training object detection and tracking algorithms. The proposed methodology has several advantages, including reduced computation cost, faster processing of data and broader applicability. Besides, there is gap in literature that studies the frantic movement patterns in fishes. Frantic patterns are characterised with sudden changes in swim speed and directions. Also, there is a lack of literature that combines water quality metrics with fish behaviour. Integrating water quality data with fish behaviour can provide more reliable and valuable insights. In this regard, the current study attempts to integrate the behaviour analysis of fish with the water quality data.

The present study focuses on identifying the behaviour patterns in *Sillago sihama*. It is a bottom-dwelling fish and can adapt to blend itself with the surrounding environment. *Sillago sihama* is a valuable fish species due to its high nutritional and economic significance (Hou et al., 2011). The study is particularly interested in the frantic behaviours of fishes. This pattern is associated with rapid movements, which are crucial since this is the fish's earliest stress-related behaviour. Therefore, identifying this behaviour pattern will help the fisherman manage the culture systems more effectively. However, the scarcity of datasets for fish behaviour analysis limited the development of computer vision-based techniques in aquaculture. In this line, the paper also suggests a trustworthy dataset, SSVid- Frantic for studying frantic behaviours of *Sillago sihama*. Further, to establish the reliability of the developed model, another dataset is generated from real-world cages with Tilapia fish. The proposed study uses these two datasets to evaluate the performance of the developed frantic behaviour detection model. Thus, the study advances the knowledge of frantic behaviours in *Sillago sihama* by addressing the scarcity of comprehensive datasets and provides a valuable resource for developing more effective computer vision-based techniques in aquaculture management.

In real-time, the system will be deployed in remote locations, including rivers and Re-circulatory Aquaculture Systems (RAS). Consequently, the developed model must be computationally efficient, lightweight and reliable. The present study introduces a novel approach to seamlessly integrate water quality data and fish behaviour, ensuring dependable monitoring of fish farms. Fig. 1 illustrates the overview of the proposed study. The FPGA-based system efficiently tracks water quality metrics within the aquaculture environment. Predictive models are employed to forecast water quality measurements for the upcoming period, with outlier detection algorithms identifying deviations in these predictions. The detection of outliers triggers the behaviour analysis model. This model searches for frantic behaviour patterns in the fish, and users are promptly alerted to any identified trends.

The contributions of the paper are as follows:

- A novel dataset named *"SSVid-Frantic"* for studying the frantic patterns of the fish *Sillago sihama* in a large aquarium.
- A novel approach to monitor aquaculture that integrates the water quality parameters with the behaviour patterns of fishes.
- A deep learning based approach for detecting frantic fish behaviour pattern, that is independent of object detection and tracking.
- The proposed and other methods are qualitatively and quantitatively evaluated on the developed dataset.

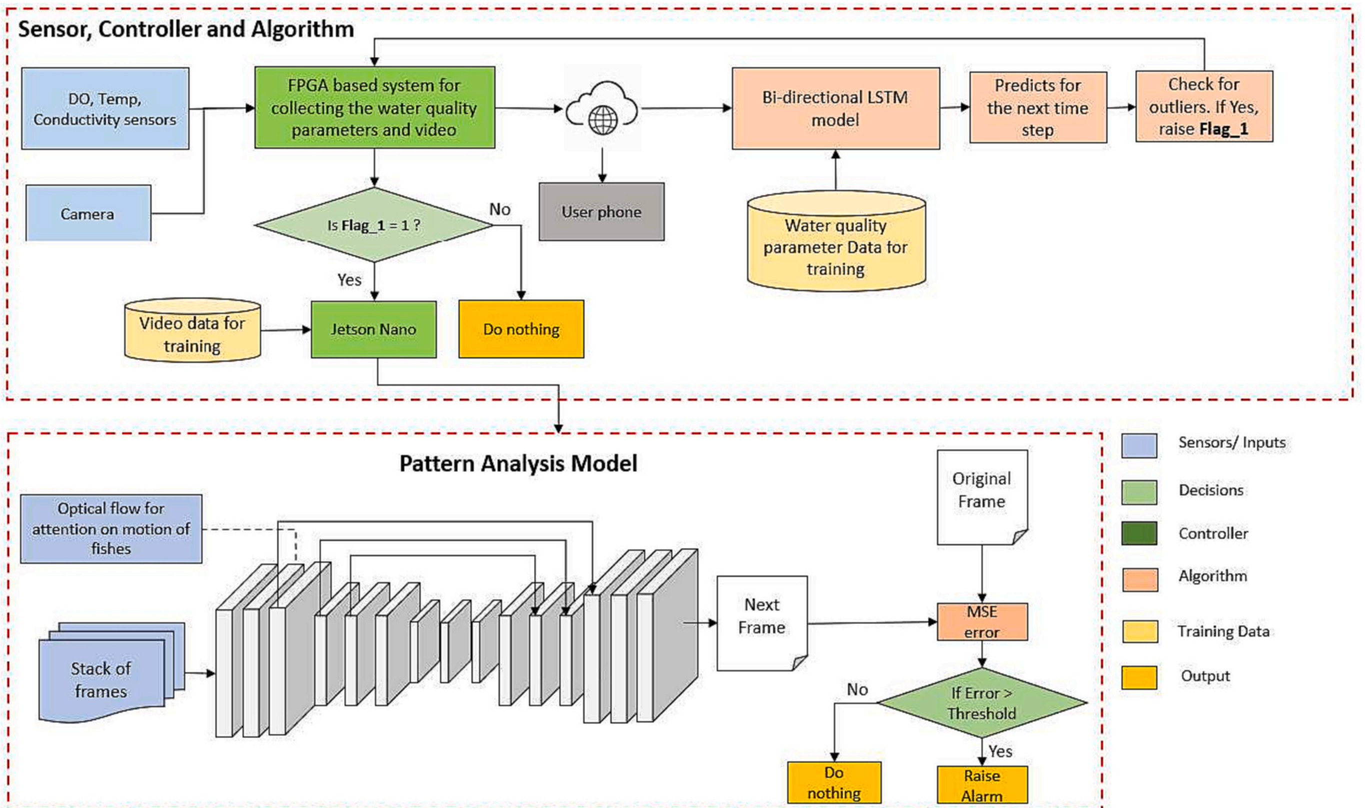


Fig. 1. Overview of the proposed study: Sensors sample the water quality parameters which are processed in the FPGA. It is then sent to the base station where a prediction and outlier detection models are utilized to detect the outliers in the next time stamp. If any outliers are detected, pattern analysis model is triggered which detects the frantic patterns of *Sillago sihama*.

- The performance of the proposed model is validated on the aquaculture cage dataset.

The paper is organised as follows. The description of the dataset and proposed method is presented in Section 2. The Section 3 discusses the performance metrics and the results obtained. The paper is concluded in Section 4.

2. Materials and methods

2.1. Dataset generation

The present study proposes a novel fish dataset for behaviour analysis. The study provides annotations for the frantic movements of the fish. The frantic movement pattern is significant in aquaculture, as fishes exhibit this behaviour when under stress and fleeing. The sources of stress could be changes in the ecosystem, diseases, predation, changes in diet and stocking density. The current section describes the methodology adopted to generate an accurate dataset for frantic pattern analysis of *Sillago sihama*.

The study uses a meticulously constructed aquarium to observe the frantic behaviour of *Sillago sihama's*. The $2.4 \times 0.9 \times 1.22m$ aquarium has a 2 inch sand bed replicating the natural habitat of fish. The bio-filter of the aquarium removes solid waste and lowers ammonia from fish waste to maintain water quality. It is built with acrylic panels for durability and stability. Aeration is another feature embedded in the bio-filter. It provides the necessary DO for the fish, reducing interruptions in video recordings and enabling accurate observations of fish behaviour. High-density polyethene (HDPE) sheets are used along the sides of the aquarium to eliminate external background scenes and focus solely on the fish activities. Wild-caught *Sillago sihama* from the Sita-Swarna river backwaters is used in the dataset to improve representativeness. A fixed camera at one end of the aquarium captures videos of fish at 12 frames per second. The study provides illumination only during the video-capturing process, protecting the natural behavioural patterns and preventing disruption of the natural day-night cycle of the fish. The routine monthly cleaning of the testing environment keeps the fish habitat clean and healthy, ensuring the reliability of the dataset (see Fig. 2).

Sillago Sihama is a long, slender fish with a single dorsal fin. It is a schooling fish commonly found in large groups and is known to be an active swimmers. These coastal marine fish prefer muddy or sandy substrates for scavenging and predator protection. Naturally found in estuaries and shallow waters, these are primary predators of small crustaceans, molluscs, and other invertebrates within the sandy or muddy substrate. They are daytime feeders who engage in both surface- and bottom-feeding behaviours.

The environmental factors significantly influence the behaviour, physiology, and health of the cold-blooded *Sillago Sihama* fish. Variations in water temperature can stress fish and interfere with their metabolism. The undesirable water quality causes stress, respiratory problems, and reduced feeding. Also, habitat changes brought on by coastal development impact their feeding and breeding habits. The predation pressure alters the foraging and movement patterns. Further, the variations in the prey abundance affect the growth and reproduction rates. Managing *Sillago Sihama* populations in aquaculture and natural habitat requires careful consideration of these factors.

The annotations for the frantic patterns of *Sillago sihama* are provided at the frame level. The study makes a significant effort to understand the behaviours of *Sillago sihama* and the accompanying patterns in their movements. Abrupt changes in swim speed, sudden changes in direction, brief bursts of activity, and erratic movements characterise the

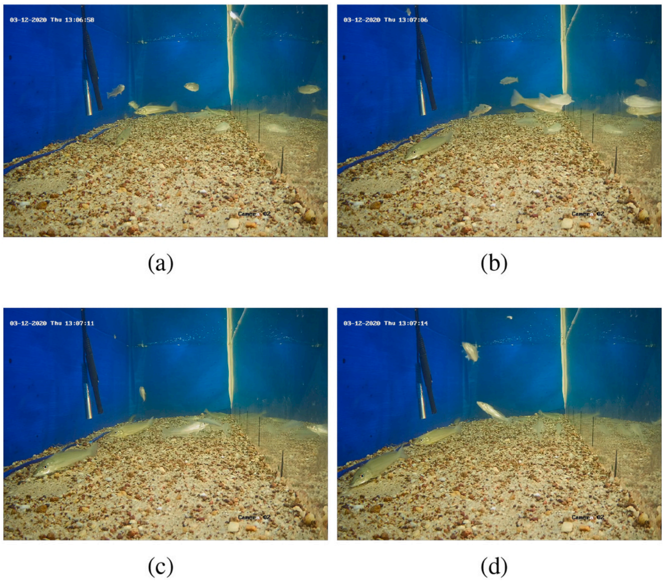


Fig. 2. Sample frames from the captured videos. The blue background is the HDPE sheet used to prevent the capture of external scenes. The fishes seen in the frames are the *Sillago sihama* caught from the wild. The bottom of the aquarium is covered by sand bed.

frantic behaviour of fish. Subsequently, the study provides annotations for the frantic behaviour patterns. In addition, the study conducted a thorough manual review of the dataset to eliminate any labelling irregularities. The frame is regarded as frantic, even if one fish moves quickly. The objective is to detect all the frames with frantic movements. Table 1 provides the details of the dataset. The dataset contains 16 videos split into training and testing sets. The training set consists of 7 videos with usual patterns. Training the pattern analysis model using the videos with normal behaviour patterns allows it to learn the behaviour patterns categorically. The testing set consists of 9 videos with normal and frantic patterns. In addition, the study attempts to explicitly capture the behaviour of *Sillago sihama* by sampling videos at random duration. It eliminates the bias and generalizes the data.

The study evaluated the proposed models by collecting data from an aquaculture site. The study captures the data at 2 feet depth from inside the cage. A strategically placed camera at the edge of the aquaculture cage ensures maximum visual data. Fig. 3 showcases sample images extracted from the collected videos. The study evaluates the model performance using four videos ranging from 30 to 40 s. These videos were recorded at a frame rate of 30 frames per second using an underwater camera. Unlike the aquarium dataset, this particular video set involved camera motion, adding a layer of complexity to the cage dataset. Also, random particles floating in the water posed difficulties for computer vision-based methods. These factors collectively heightened the complexity of analyzing the cage dataset.

Thus, the originality of the fish behaviour dataset and the specific annotations made for the fish's frantic movements make it crucial for

Table 1
Specification of the dataset.

Parameters	Specifications
Dimension of the aquarium	$2.4m \times 0.9m \times 1.22m$
Total number of videos	16
Number of videos for training	7
Number of videos for testing	9
Number of fishes	8–10
Species of fish	Sillago sihama
Frames Per Second	12
Duration of videos	10 s – 30 s

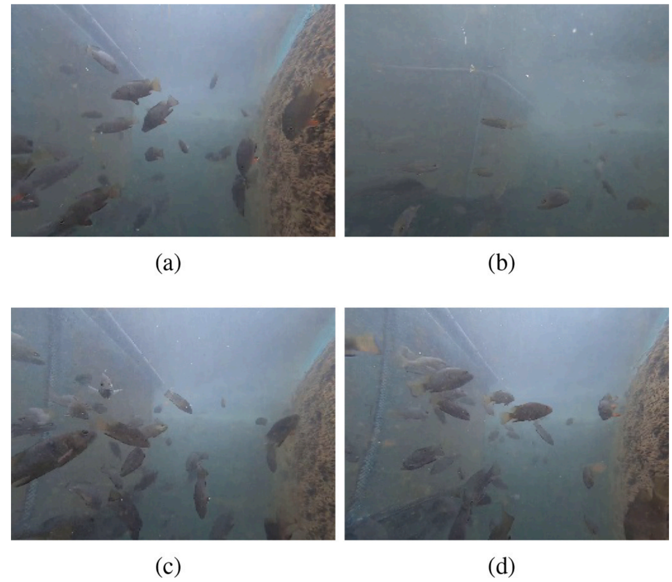


Fig. 3. Sample frames of the captured videos from the cage. The fishes seen in the frames are the *Tilapia* caught from the wild.

this study. In aquaculture settings, frantic fish movement patterns are particularly significant because they are signs of stress and flight behaviour. The dataset serves as a valuable resource for studying stress-related behaviours of fishes. It also serves as a reference for comparison of algorithms developed for fish behaviour analysis.

2.2. Overview of the proposed method

The proposed work presents a novel method to study fish behaviour by analysing underwater videos and water quality parameters. Considering just the water quality might be ineffective since they are changing. In a short time, it might change back to normal. Also, the behavioural pattern analysis model has to analyse every frame, and the ceaseless running of the behavioural analysis model is ineffective. However, integrating the behaviour pattern analysis model with the water quality monitoring system is effective in accuracy and computation. Fig. 1 shows the overview of the proposed model.

The water quality monitoring system consists of sensors and an FPGA-based controller. Sensors sense the water quality parameters in the culture systems. Since the temperature, DO, and salinity affect the growth rate, survival and metabolism of fish, the current study considers these three parameters for monitoring. The FPGA-based control system manages these sensors, and a prediction model predicts the water quality parameters for the next time step. As shown in Fig. 1, DO, temperature and conductivity sensors are connected to the FPGA-based controller (Section 2.3). The prediction model takes in the output of the FPGA, and the outlier detection model detects the outliers in the predicted values. The system raises a flag detecting the outliers and communicates this information to the FPGA controller. The raising of the flag represents the anomalous water quality parameters that may exist in the coming hour. It triggers the camera and the auto-encoder-based behaviour pattern analysis model. The trigger system prevents the necessity of running the camera and pattern analysis model continuously. The behaviour analysis model starts looking for abnormalities in fish movements. It ensures that the pattern analysis model does not miss any critical moments in the aquaculture. The method uses a Jetson-Nano board to run the auto-encoder-based pattern analysis model. Subsequently, the camera and the pattern analysis model run for a specified duration, detecting frantic patterns. The inference model analyzes these patterns and classifies them as normal/abnormal(frantic). These reports are sent back to the users as an alert signal.

2.3. Water quality monitoring system

The water quality monitoring system monitors three parameters, as stated before. The FPGA is programmed to monitor the sensors. The system logs the collected data into the cloud server for immediate access and analysis. A wireless communication technology (Wi-Fi) is opted for logging the data to the server. Subsequently, the study uses a learning algorithm to forecast the water quality parameters and an outlier detection algorithm to detect the anomalous values. Fig. 4 shows the conditioning circuit of the sensors used in the study.

2.3.1. Temperature sensor

The proposed model equips an LM35 sensor to monitor the temperature of the water. It produces an output voltage linear with measured temperature and operates with an accuracy of (+ or -) 0.5 degrees Celcius. The sensor has a range of −55 degrees Celcius to 155 degrees Celcius. Further, it does not require any calibration at room temperature. The sensor output passes through a non-inverting amplifier with a gain of 10 as it is in millivolts. Finally, the amplified signal passes through the Analog Digital Converter (ADC) circuit to convert it to digital format.

2.3.2. Dissolved Oxygen sensor

The study uses a membrane-based Dissolved Oxygen (DO) sensor to monitor the DO in the culture system. This sensor has an output voltage linear with the measured DO with the output voltage varying from 0 to 35–55 millivolts in air. Unlike the temperature sensor, this sensor produces a differential signal output. Accordingly, the study utilizes a differential amplifier with a gain of 10. Subsequently, the analog signal is converted to digital using ADC.

2.3.3. Conductivity sensor

The study monitors the salinity of the culture system using a conductivity sensor. Unlike other sensors in the study, it is a passive device with a cell constant of 0.1/1.0. The conditioning circuit uses the sensor in the voltage divider circuit to get the resistance value of the medium. Subsequently, the resistance value is processed to compute the salinity.

2.3.4. Camera

The system uses a CMOS ov7670 camera to capture the videos of the culture system. It is sensitive to low light operations and supports RGB, YUV and YCbCr image formats. The collected video from the sensor is displayed on the monitor using a Video Graphics Array (VGA) connector.

2.3.5. Controller

The present study uses Spartan-6 XC6SLX9 FPGA as the controller. Spartan-6 is a low-cost, low-power consumption, and logic-optimized FPGA with 9,152 logic cells, 512 KB memory and 16 Discrete Signal Processor (DSP) slices. This FPGA further equips a 50 MHz crystal oscillator, an ADC, a Wi-Fi, a Liquid Crystal Display (LCD), a VGA and a camera (Section 2.3.4). ADC converts the incoming analog signals from the sensors to digital signals, as FPGA can only process digital signals. The study uses an LCD to visualize the sensor values. Also, these processed signals by the FPGA are transmitted to the server using a Wi-Fi module. The attached camera captures the video for pattern analysis. The Jetson Nano board deploys an auto-encoder-based frantic pattern analysis model. The system pushes the collected videos to the Jetson Nano board for pattern analysis.

2.3.6. Water quality parameters prediction

The sampled water quality parameters are further processed to predict the values for the next time stamp. To this end, the study uses Long Short Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997) as the learning model. LSTM and Recurrent Neural Network (RNN) are popular deep learning-based sequence models for predicting time series data. The current study uses LSTM for water quality parameter prediction since RNN suffers from a vanishing gradient problem. LSTM addresses the RNN's drawback of preserving long-term temporal information. The inclusion of memory cells in the LSTM enables it to sustain information for a longer duration (Hochreiter and Schmidhuber, 1997). There are several well-known architectures of LSTM, including stacked LSTM, Convolution LSTM and bi-directional LSTM. The current study uses bi-directional LSTM as the learning model (Schuster and Paliwal, 1997). In the bi-directional LSTM architecture, the sequence is examined both forward and backward by two unidirectional LSTMs. The input token sequence is processed conventionally by one LSTM and reversed by the

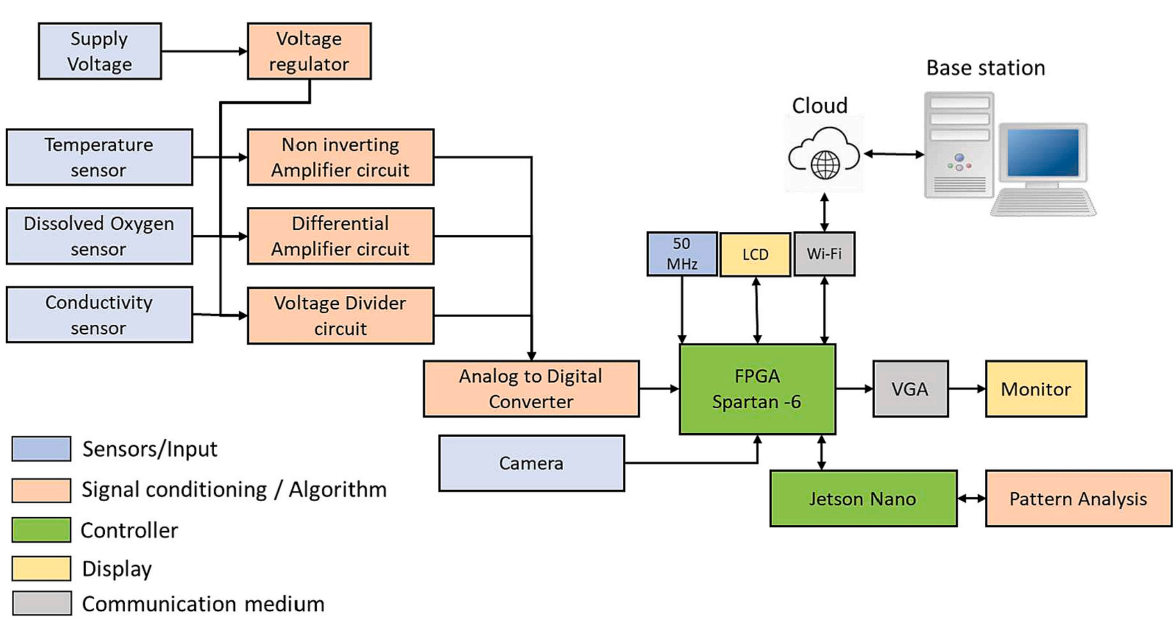


Fig. 4. Overview of the sensor signal conditioning circuit. The study uses Temperature, DO and Conductivity sensors. These sensor outputs are passed through the amplifier and ADC circuits before passing it to the FPGA controller. The sensor signals are converted to a suitable format in the FPGA before being communicated to the base station. Behaviour pattern analysis model is deployed in the Jetson Nano board which is connected to the FPGA board.

other LSTM. The output from each LSTM is a probability vector, and the outputs from the two LSTMs are combined to generate the final output. Fig. 5 illustrates the general architecture of bi-directional LSTM. It ensures the model learns temporal dependencies and complexities in water quality parameters. Let V represent the input to the model. The study considers three water quality parameters namely, temperature (f_t), DO (f_{do}) and Salinity (f_{sal}). Thus, V at a given timestamp (t) is given as follows:

$$V_t = (f_t, f_{do}, f_{sal}) \tag{1}$$

The model considers the current and 4 previous time stamp values of temperature, DO and salinity to make predictions for the next time stamp ($t+1$). Thus, the input to the model is as follows:

$$V = (V_t, V_{t-1}, V_{t-2}, V_{t-3}, V_{t-4}) \tag{2}$$

The model predicts the next time stamp values V_{t+1}. There exist correlations between the water quality parameters. Utilizing an independent bi-directional LSTM model for predicting each parameter is inefficient, as the model cannot learn the correlation between the parameters. The present study addresses this issue by training a single bi-directional LSTM model for predicting different water quality parameters. It ensures that the model learns the correlation between the water quality parameters. Also, it makes the model lightweight and deployable in real time. Subsequently, the predictions are provided to the outlier detection model to detect the abnormal water quality parameters.

2.3.7. Outlier detection model

The study uses an outlier detection system to provide more reasonable decisions on the water quality parameters. Each fish species can survive within a baseline range of water quality parameters. Exposing fishes to water quality other than the baseline range for an extended duration, makes them experience stress and eventually lead to mortality. Thus, outlier detection in aquaculture plays a significant role. To this end, the proposed method uses a box and whisker plot to obtain the data symmetry, skew, variance and outliers. It provides the Inter Quartile Range (IQR) representing the middle portion of the data. If the predicted values are greater than the maximum value identified by the box and whisker plot, then the values are considered outliers. The detection of outliers raises the flag and informs the user via an alert system. It also triggers the video capture and behaviour analysis model to analyze the variations in fish behaviour.

2.4. Behaviour pattern analysis

2.4.1. Frantic pattern detector

The proposed method aims to model the frantic movements of fish characterised by quick movements brought on by excitement or stress.

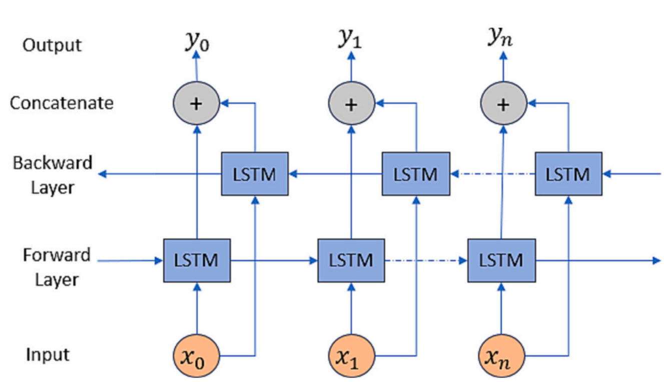


Fig. 5. The architecture of bi-directional architecture. Here, x_i and y_i represents the input and the output of the model. It consists of two layers of LSTM's, one learning the sequence in forward direction and the other learning the sequence in backward direction (Schuster and Paliwal, 1997).

These behaviours are significant in aquaculture because they might occur as a response to anomalous water quality parameters. As a result, identifying and categorising these patterns impacts the aquaculture management. The main goal of the present study is to discern between normal and frantic fish behaviour. Normal behaviour is distinguished by minor displacements of fish from one frame to the next, whereas frantic behaviour entails extensive movement between frames. The study achieves this using an auto-encoder-based reconstruction technique (Hinton and Zemel, 1993). By training the auto-encoder on normal patterns, the model becomes proficient at reconstructing frames with normal patterns. As a result, the Mean Squared Error (MSE) between the original and reconstructed images is small. However, since the model has not seen frantic patterns during its training phase, it has difficulty reconstructing frames with frantic movements, which causes a significant MSE error. The study uses the MSE scores to differentiate between regular frames and frames with frantic patterns. This metric is reliable to categorise behaviours based on their higher MSE values. Additionally, the study uses optical flow in this auto-encoder-based strategy to highlight the region of interest (i.e., moving fish). This straightforward yet efficient technique facilitates the identification of frantic patterns in fish behaviour by doing away with the need to train object detectors and tracking algorithms.

Thus, the proposed methodology recognises and categorises frantic behaviours in real-time fish monitoring, allowing aquaculture operators to respond to possible stress or disturbances in the fish population. Automated detection of frantic behaviours streamlines monitoring processes and reduces the need for constant manual observation. Additionally, this method can help with better fish welfare, early health problem diagnosis, and optimised aquaculture management practises, thus improving the effectiveness and sustainability of aquaculture systems.

The proposed auto-encoder-based model uses encoder-decoder architecture with an attention mechanism. Fig. 6 illustrates the proposed auto-encoder architecture. The design also considers temporal information to model the motion of fish accurately. Let X of dimension $H \times W \times C$ represent the input (set of images) to the encoder, where H is the height, W is the width, and C is the number of channels. Let k represent the current frame. The proposed model takes in the stack of three frames $(k, k-2, k-4)$ to capture the temporal information. Thus, the input to the proposed model X is of size $H \times W \times 9$ and the reconstruction of the proposed model is given as $\widehat{X}$. The encoder consists of 4 convolution blocks. The first three blocks consist of 2 sets of convolution and ReLu activation layers. A batch normalization layer follows it to reduce the convergence period. Introducing the max pooling layer reduces the feature maps' dimension and propagates the most prominent features to the deeper layers. The last block of the encoder consists of a convolution layer with 512 filters. Likewise, the decoder consists of 4 sets of convolution blocks. The first 3 blocks incorporate 2 sets of convolution and ReLu activation layers. However, the upsampling layer precedes it to increase the dimension of the feature maps. The last block of the decoder includes a convolution and the sigmoid activation layer. The proposed model uses skip connections with an attention mechanism from encoder to decoder. It compensates for the loss of spatial resolution due to max pooling and enables feature re-usability.

In addition, the proposed method uses an attention mechanism to focus on the regions of interest (fish), ensuring accurate reconstruction. Each frame consists of large background information. The attention mechanism aids the model in concentrating on fish regions and filtering out irrelevant information. Specifically, the study introduces the Optical Flow information in the auto-encoder model. Significantly, the literature consists of two stream models to reconstruct the original frame and its corresponding optical flow (Cruz-Esquivel and Guzman-Zavaleta, 2022; Ravanbakhsh et al., 2017; Li et al., 2021). It ensures that the model learns temporal and spatial information. Contrarily, the use of two streams increases the computation cost. The proposed method addresses

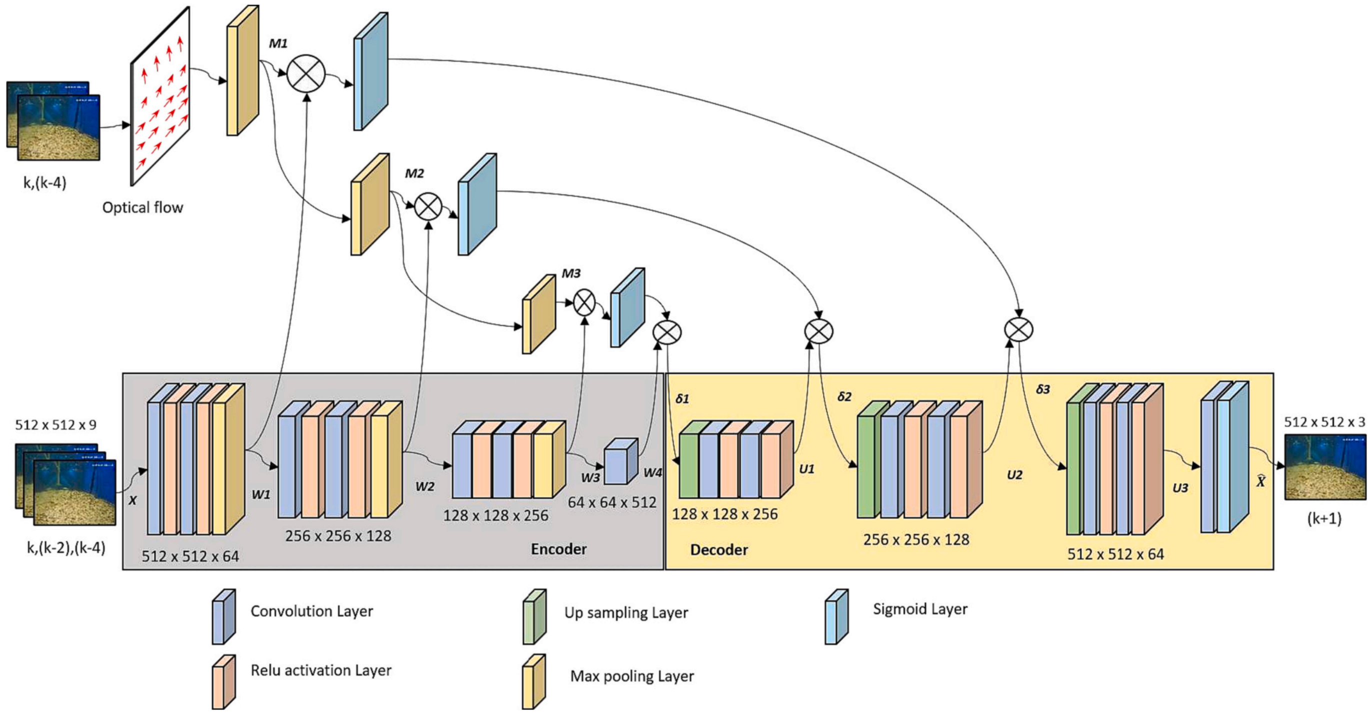


Fig. 6. The auto-encoder based architecture of the proposed pattern analysis model. The proposed encoder-decoder based model integrates motion information using Optical Flow to reconstruct the frame.

this issue by utilizing a single-stream model. The single-stream model integrates the optical flow-based attention. Let $OF = (u, v)$ represent the dense optical flow vector for the entire image computed between k^{th} and $(k-4)^{th}$ frame. Subsequently, the proposed method computes the displacement magnitude (M) of the optical flow vector as,

$$M = \sqrt{u^2 + v^2} \tag{3}$$

where, u and v represents the gradients in x and y directions respectively. The region in M corresponding to fishes has a better response than the other regions. Hence, M is used as an attention map to focus on the region of interest. Let W_i represent the output of the convolution block in the encoder section (encoder Fig. 6) where $i = 1, 2, 3, 4$. Likewise, U_i represents the output of the convolution block in the decoder section (decoder Fig. 6). Further, the methodology multiplies W_i with the M_i to highlight the features concerning the fish. The resultant feature map of this operation is multiplied by the output of the previous convolution block U_i. Thus, input to each convolution block in the decoder δ is as follows,:

$$\begin{aligned} \delta_{i=1} &= (W_{4-i} \times M_{4-i}) \times W_4 \\ \delta_{i=2,3} &= (W_{4-i} \times M_{4-i}) \times U_{i-1} \\ \delta_{i=4} &= U_{i-1} \end{aligned} \tag{4}$$

Significantly, the only moving objects in the frame are the fishes, and optical flow provides the motion information of the fishes. The model is trained to predict the next frame $(k+1)$.

2.4.2. Inference of behaviour patterns

The above-noted behaviour pattern analysis model is used to reconstruct the next frame. The inference algorithm attempts to identify the patterns based on the capacity of the auto-encoder to reconstruct the next frame. The model computes the Mean Squared Error ∊ between the original and the reconstructed frame.

$$\epsilon = \frac{1}{n^*m} \sum_{i=1}^{n} \sum_{j=1}^{m} \left(X_{i,j} - \widehat{X}_{i,j}\right)^2 \tag{5}$$

where X represents the original frame, $\widehat{X}$ represents the reconstructed frame, n represents the height and m represents the image width. If the error is greater than the set threshold τ, the frame is considered anomalous. The value of τ is set experimentally on the dataset. A high mean square error signifies a huge difference between the original and reconstructed frame. Since the reconstructed frame constitutes a normal behaviour, a high MSE indicates an anomalous event.

2.5. Implementation details

The study involves two components for monitoring the aquaculture system: the hardware component (sensors and control units) and the software component (LSTM and auto-encoder-based prediction model). The study uses temperature, DO and salinity sensors to monitor the parameters of the ecosystem. The output of these sensors is distinct from each other. Accordingly, necessary signal conditioning circuitry is adopted to convert it to an FPGA readable format. The voltage regulator 7805 IC provides a 5 V voltage supply to the signal conditioning circuits. It ensures a steady voltage supply and offers protection from overvoltage. The LM324 IC is an Operational Amplifier (Op-Amp) utilized in the circuit design to generate the necessary amplification for the sensor output. The amplified sensor signals are then passed through the ADC to convert it to a digital signal. The present study utilizes MCP3208 as the ADC for the amplified sensor signals. It adopts Successive Approximation Registers(SAR) architecture to provide low power consumption, high resolution and accuracy. The study uses FPGA as the control unit. It offers advantages such as cost efficiency and high parallelism in program execution. The FPGA is programmed using Xilinx Integrated Synthesis Environment (ISE). Finally, the system integrates the ESP8266 Wi-Fi module to push the data to the server. Fig. 7 shows the implementation of the proposed water quality monitoring system using FPGA.

The water quality prediction model (Sections 2.3.6) used in the study is trained using an Intel(R) Core (TM) i7-8550U CPU and an NVIDIA 1650 GPU. Section 3.2 provides a performance evaluation of various LSTM models using different optimizers and loss functions. The study trains the water quality prediction model using data collected from the

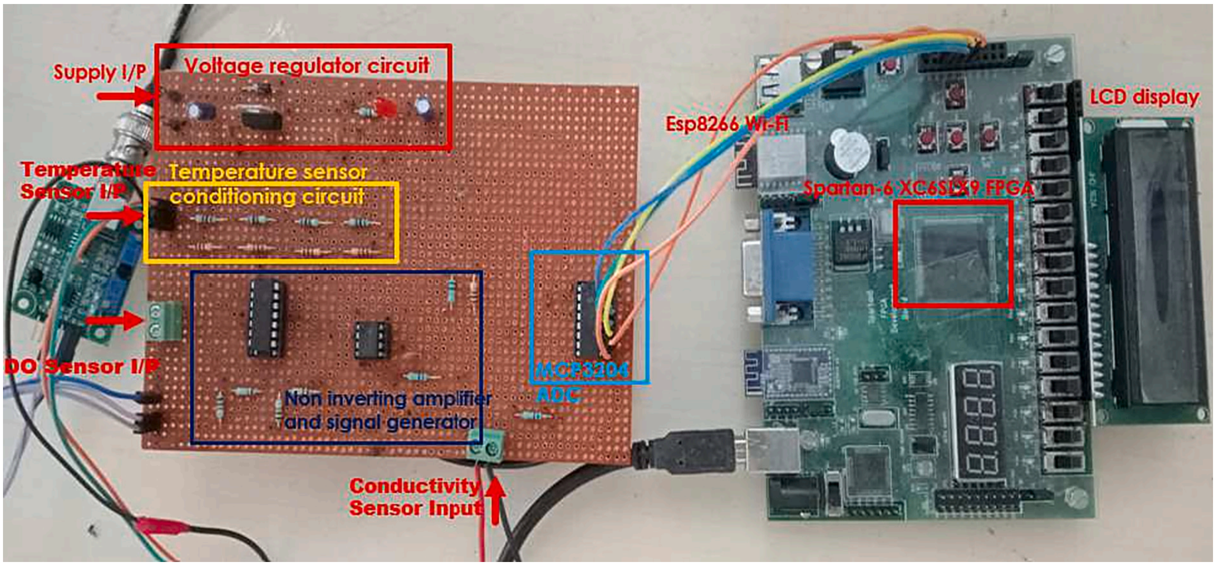


Fig. 7. FPGA based water quality monitoring system. The developed system monitors temperature, Dissolved Oxygen and Salinity.

aquarium. The generated dataset includes 7325 sample values for temperature, DO, and salinity taken once every 90 min. It is divided in the proportion 70:30 into train and test data. With a learning rate of 0.00005 and a batch size of 8, the LSTM models are trained over 200 epochs.

The present study uses an auto-encoder-based reconstruction approach to detect a frantic pattern of the fishes (Section 2.4.1). Again, the study trains the auto-encoder using an Intel(R) Core (TM) i7-8550U CPU and an NVIDIA 1650 GPU. The model is trained over 300 epochs with a batch size of 1 and a 0.001 learning rate. Videos gathered from the aquarium are used in the study to train the model. Fig. 8 illustrates the input to the auto-encoder as a stack of three 512×512 frames, denoted by $(k), (k-2)$ and $(k-4)$. The Mean Squared Error (MSE) loss function and Adam optimizer guide the training process. The study experimentally sets the threshold τ to 1200.

3. Results and discussion

3.1. Performance metrics

The study uses two learning models, namely bi-directional LSTM and Auto-encoder. It calculates the Root Mean Square Error (RMSE) between the ground truth and the predicted values to evaluate the performance of the bi-directional LSTM model. The study estimates the accuracy, precision, recall, and f1-score to assess the performance of the auto-encoder

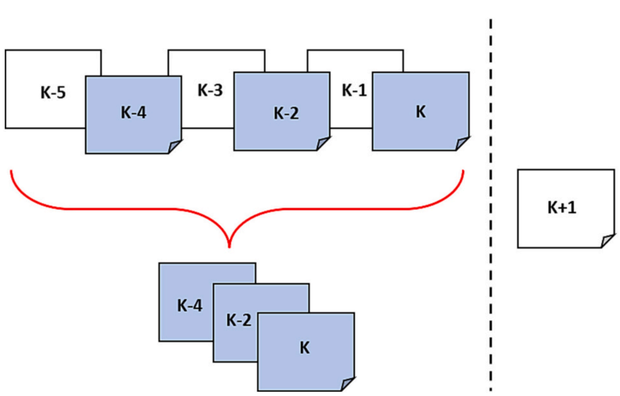


Fig. 8. $k, k-2, k-4$ are the inputs to the auto-encoder model and $k+1$ (next frame) is the output of the model.

model. To this end, it compares the frame-level anomaly scores predicted by the model with the ground-truth frame-level anomaly scores.

3.2. Performance evaluation of water quality parameter prediction and outlier detection system

The study trains the model using data sampled for 90 min. It chooses a 90-min window to enable the model to comprehend the underlying trends in water quality measurements that develop over this duration. Consequently, the model simulates the expected values for the following 90^{th} minute while making predictions. A shorter time limit, such as 15 or 30 min, would force the model to learn and base its predictions primarily on the patterns occurring within those shorter time frames. However, the fisherman might be unable to take precautions due to the shorter time window. Instead, the longer window benefits the fisherman since it gives them enough time to take safety precautions. There exists literature that uses machine learning algorithms such as Artificial Neural Networks (ANN), Regression Trees and Support Vector Machine (SVM) to forecast water quality parameters (Chou et al., 2018; Ahmed et al., 2019). However, these machine-learning models cannot capture long-range dependencies due to a lack of memory cells. LSTM can capture the long-term temporal dependencies using the memory cells, allowing them to analyze time series data more accurately. Further, as LSTM models require a large dataset to train, the authors in Zheng et al. (2023); Haq and Harigovindan, 2022 explored the hybrid methods. However, the current study uses considerable data with 7325 sample values. As a result, the present study does not consider hybrid models for predicting water quality parameters as it increases the complexity. The current study examines and compares the performance of several LSTM models, including LSTM, Stacked LSTM, and bi-directional LSTM, in forecasting water quality parameters. Table 2 presents the results

Table 2
RMSE scores for various learning models. The best model is highlighted in bold.

Learning models	RMSE Scores Train
LSTM (RMS prop, MSE Loss)	1.20
Stacked LSTM (RMS prop, MSE Loss)	1.36
Bi-directional LSTM (RMS prop, MSE Loss)	1.09
Bi-directional LSTM (SGD, MSE Loss)	15.08
Bi-directional LSTM (Adam, MAE Loss)	1.10
Bi-directional LSTM (Adam, Huber Loss)	1.07
Bi-directional LSTM (Adam, MSE Loss)	**1.01**

demonstrating the bi-directional LSTM's superior efficiency over the stacked LSTM and LSTM models, as shown by lower Root Mean Squared Error values. The bi-directional LSTM has the advantage of learning the patterns in both forward and backwards directions, which makes it better able to identify data trends than the other models. The bi-directional LSTM model is then tuned using hyperparameters. Consequently, the study investigated the model's performance with the Adam and Stochastic Gradient Descent (SGD) optimizers. In addition, an

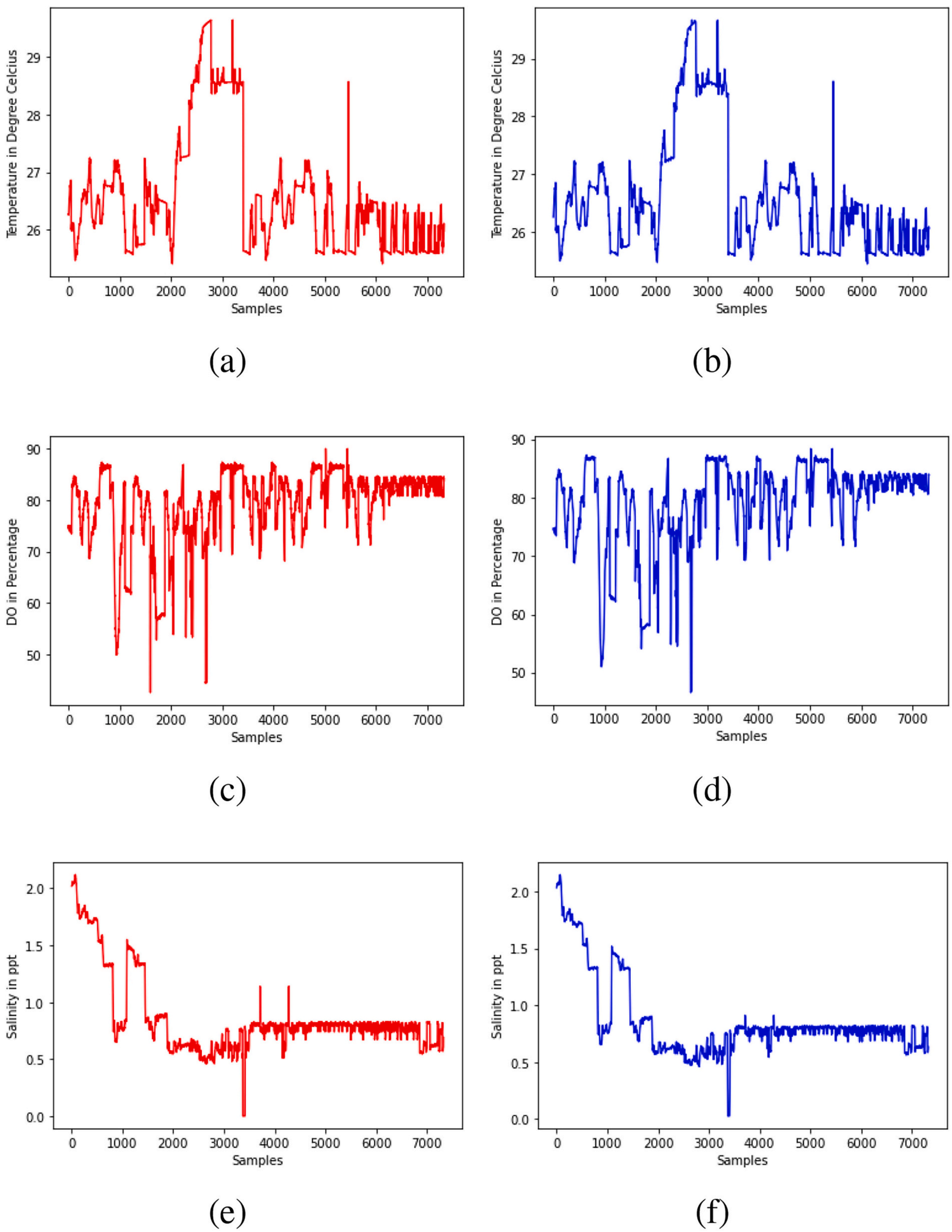


Fig. 9. Sample results of bi-directional LSTM model trained with Adam optimizer and MSE loss function in predicting the water quality parameters.The first column in red shows the ground truth values from the sensors. The second column in blue shows the predicted values from the model. The first row (a), (b) represents the temperature, second row (c), (d) represents the DO and the third row (e), (f) represents the salinity parameter.

experiment with different loss functions such as MSE, MAE, and Huber is conducted. The study dropped the SGD optimizer since the model did not perform well. The bi-directional LSTM model performed favourably for the Adam and RMS prop optimizers. However, the bi-directional LSTM model outperforms the alternative architectures when using Adam optimizers and MSE loss function, as indicated with a lower train RMSE error of 0.01. The results highlight the importance of considering temporal dependencies and complexities while forecasting water quality parameters.

Fig. 9 shows the ground truth and predicted water quality parameters by the bi-directional LSTM models. There are sudden peaks and bottoms in the ground truth of water quality parameters illustrated in the figure. However, the model predicted these sudden changes accurately. Thus, the results highlight the effectiveness of bi-directional LSTM models in modelling the dynamic nature of the water quality parameters. Fig. 10 shows the box and whisker plot for temperature, DO, and salinity. The maximum temperature and salinity values are 28.5 $^\circ C$ and 1.1*ppt*, respectively. Thus, if any new predictions are over the identified threshold (i.e., 28.5 $^\circ C$ and 1.1*ppt*), then it is considered outliers. However, DO values need to be as high as possible for the growth of the fish. Hence, the DO parameter considers the minimum value. The study identifies the minimum DO as 63% using the Fig. 10. If predictions for DO fall below the threshold, it is considered an outlier.

3.3. Ablation study

The study uses an auto-encoder-based reconstruction approach to identify the frantic movement patterns of the fishes. In this regard, the analysis consists of two experiments to identify the optimum model. Firstly, the experiment tests different configurations of auto-encoder to determine the optimum model. Section 3.3.1 gives the details of the same. Secondly, it examines the performance of the model for different time windows. Section 3.3.2 details the time window experiments.

3.3.1. Effectiveness with different architectures

Initially, the study configures an auto-encoder without skipping connection or attention modules. The experiment calculates the accuracy, precision, recall, and F1 score to evaluate the performance of different architectures. In an ideal case, the best-performing model will have the values of these parameters close to 1. Table 3 shows the results obtained for various auto-encoder designs. All the experimented configurations included batch normalization layers to make the network more stable during training. As anticipated, the model performed poorly without skip connections and attention modules. Without an attention module, an autoencoder may process the entire image without discrimination, making it harder to identify the relevant features. The model struggles to preserve the contextual and spatial features from the preceding network layers due to the absence of skip connections. This configuration had the poorest performance, with an accuracy of 0.45. It is reasonable as the model could not locate the object of interest and focused on the entire frame, resulting in poor frame reconstruction. Fig. 11 illustrates this scenario. Subsequently, skip connections were introduced to propagate the features from the encoder to the decoder. It helps retain contextual and spatial information. The encoder and decoder, however, might not be learning the characteristics that align exactly in spatial location due to a lack of attention mechanism. This misalignment may cause errors in locating the target objects, reducing the performance. The model performed with an accuracy of 0.48. There is a minor improvement in the performance compared to the previous model, and yet not satisfactory. Subsequently, the study introduces only the attention module to the auto-encoder configuration. It allows the model to identify the region of interest. However, the quality of the features it attends to determines the efficacy of the attention mechanism. The model might not have access to the most informative features without skip connections, resulting in ineffective attention allocation and subpar performance. The model performance increased to 0.53 with the inclusion of attention to the auto-encoder configuration. Finally, The study designs an auto-encoder with skip connections and an attention module. Multiplying the encoded feature map with the optical flow-based attention map allowed the model to focus on the region of interest. Skip connections propagate and store spatial and contextual

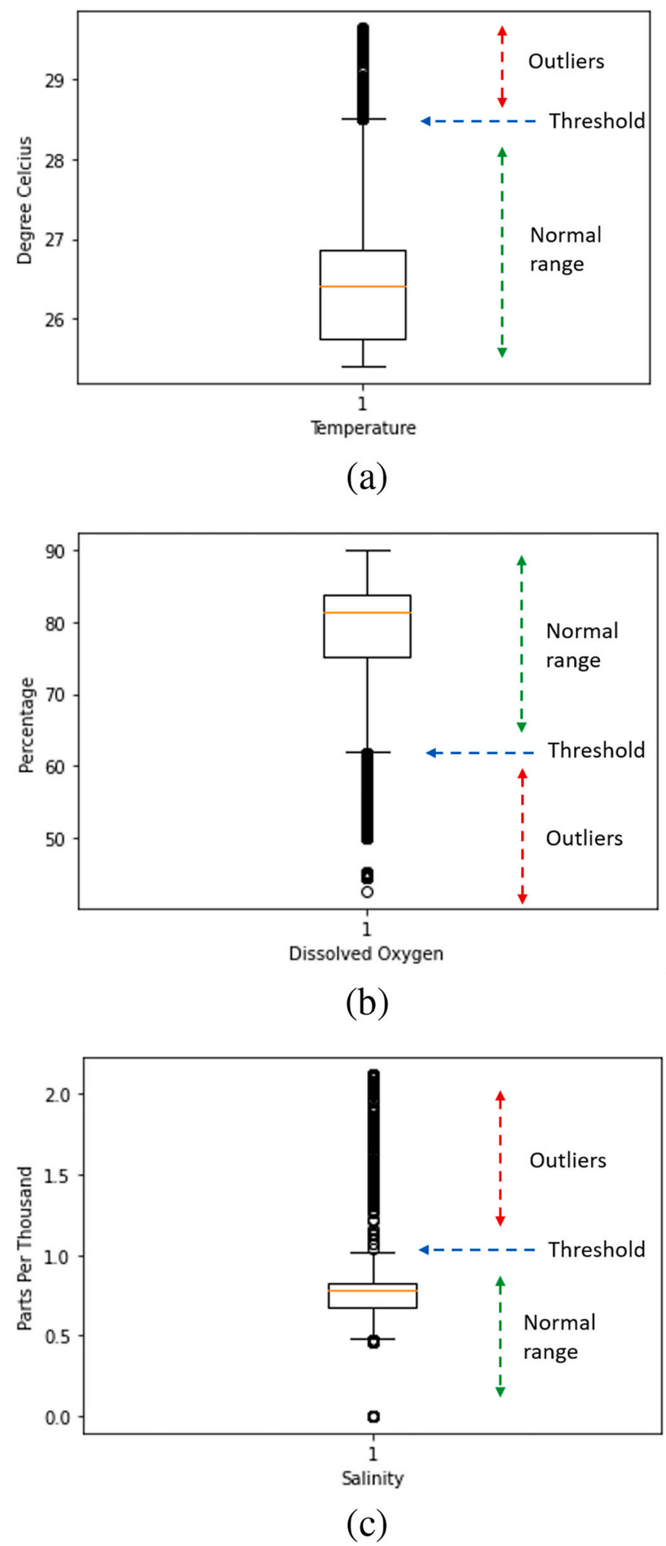


Fig. 10. Outlier detection using box plot for temperature, Dissolved Oxygen and Salinity. The maximum for the temperature is at 28.5 degree Celcius and salinity is at 1.1ppt. The minimum for DO is at 63%.

Table 3
Accuracy, precision, recall and f1-scores for different configuration of auto-encoder. The best model is highlighted in bold.

Models	Accuracy	Precision	Recall	F1-score
Auto-encoder	0.45	0.50	0.45	0.46
Auto-encoder (With Skip connection)	0.48	0.55	0.48	0.48
Auto-encoder (With optical flow-based attention)	0.53	0.52	0.53	0.52
Auto-encoder (With Skip connection and optical flow based attention)	**0.68**	**0.70**	**0.68**	**0.60**

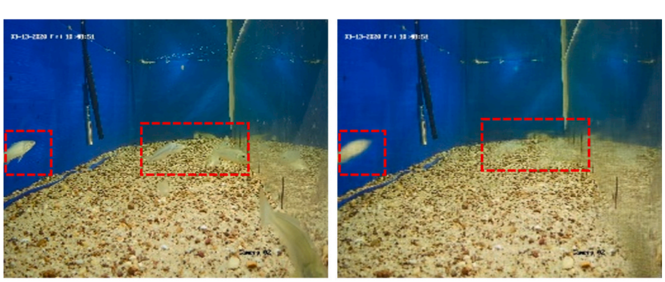
(a) (b)

Fig. 11. The inability of the model to localize the fishes (auto-encoder without skip connections and optical flow-based attention). (a) shows the ground truth frame and the corresponding reconstructed frame is shown in (b).

information. The attention mechanism focuses the model's attention on the areas of interest, improving the capacity to identify the target object. Fig. 12 illustrates this, and it performed considerably better than the other configurations, with an accuracy of 0.68. The study revealed that the model could identify almost all anomalous frames accurately. Fig. 13 shows the ground truth, the corresponding optical flow and the reconstructed frames.

3.3.2. Effectiveness with different temporal window

The input to the model is the stack of frames, which allows the model to extract the temporal information accurately. The temporal information plays a significant role as the model has to detect frantic patterns. It facilitates the model to capture the displacement information of the fish accurately. The study examines the effect of different temporal windows on the model performance. Since the videos are recorded at 12 fps, processing every consecutive frame for temporal information is redundant and expensive for video processing. Thus, the current study considers alternative frames for capturing temporal information. Let k represent the current frame. The goal of the model is to predict the next frame $(k + 1)$. Table 4 details the model's performance for different temporal windows. Initially, the study conducts performance evaluation with a single current frame as the input, i.e. (k). The temporal window in this scenario is small, and as a result, there is a loss of relevant temporal information. The model could not capture the motion information of the fish, resulting in poor performance. Subsequently, the study examined the performance of the model when the input is the stack of 6 frames, i.e. $(k), (k-2), (k-4), (k-6), (k-8), (k-10)$. It is interesting to note that the recorded videos are at 12 fps, and the considered window is approximately equal to 12 fps. As a result, this window captures the motion of fish for one second. The data were generalised by this wide window, making it difficult for the model to identify specific patterns and make accurate predictions. Then, to get the model's optimal input stack of frames, the time window is shortened gradually. As a result, the study reduces the window size to a stack of 4 frames, i.e. $(k), (k-2), (k-4), (k-6)$. The model performance improved when compared with previous approaches since there is less generalisation in the input data to the model. Likewise, the study tested the model's performance with a stack of 2 and 3 frames. The model's performance increased with a stack of 2 frames, i.e. $(k), (k-2)$. The rationale is that there is better temporal information to collect and less data generalisation. This temporal window produced an accuracy of 0.67. Compared with the previous window size of 4 frames, the accuracy increased by 0.03. With the 3 frames, i.e. (k), $(k-2), (k-4)$, the model's performance improved considerably, with an accuracy of 0.68. Here, the model arbitrates between capturing temporal information and reducing data generalisation. Also, increasing the temporal window beyond the stack of 3 frames decreases the model's performance due to data generalisation. Hence, in the current study, the temporal window is set as a stack of 3 frames (i.e. $(k), (k-2), (k-4)$).



(a) (b)

Fig. 12. Effectiveness of multiplying the encoded feature map with the optical flow based attention map. (a) shows the ground truth frame and the corresponding reconstructed frame is shown in (b). The model capacity to focus on the region of interest is highlighted in the red bounding box.

3.4. Comparative analysis

The analysis involves comparing the proposed auto-encoder model against the conventional methods on the proposed dataset. Several pieces of literature study the different behaviours of fishes (Bertolini et al., 2022; de Vargas Guterres et al., 2020; Bhaskaran et al., 2019; Papadakis et al., 2012; Hulse et al., 2022). However, none of these studies focuses on modelling the frantic movement patterns of fishes. In this regard, the study selects relevant literature that addresses similar challenges in other research domains (anomalous patterns in humans). These models are trained on the proposed dataset for performance evaluation. Table 5 shows the accuracy, precision, recall, and f1-scores of the models on the dataset. Fig. 14 shows the sample results. The methods described in Ravanbakhsh et al. (2017); Li et al., 2021; Cruz-Esquivel and Guzman-Zavaleta, 2022 performed poorly in reconstructing the regions with fishes as they struggled to identify the region of interest. Significantly, the fish in the proposed dataset have the same colour as the sand, making target object detection difficult. The rapid motions of fish further enhance the challenge. Also, when the background and the moving objects have similar visual properties, it can be difficult for the model to capture the motion information merely based on optical flow. Further, these models cannot capture the movement information accurately as these models consider a small temporal window as the input. As a result, these models fail to capture intricate and rapid fish motions during frantic patterns. However, a wider temporal window enables the model to represent long-term dependencies. Skip connections, attention mechanisms, and a wider temporal window assisted the suggested model in successfully addressing the difficulties of recognising fish areas with movement and capturing precise temporal information. The study's feature propagation technique enables retaining the contextual and spatial data. In addition, it uses a stack of three images (i.e. $(k), (k-2), (k-4)$) which enables the model to capture the temporal information more accurately. The model may prioritise the regions crucial to the fish's motion and behaviour by attending to the moving regions of the input frames using an attention mechanism. Finally, the proposed approach multiplies the encoded features with the

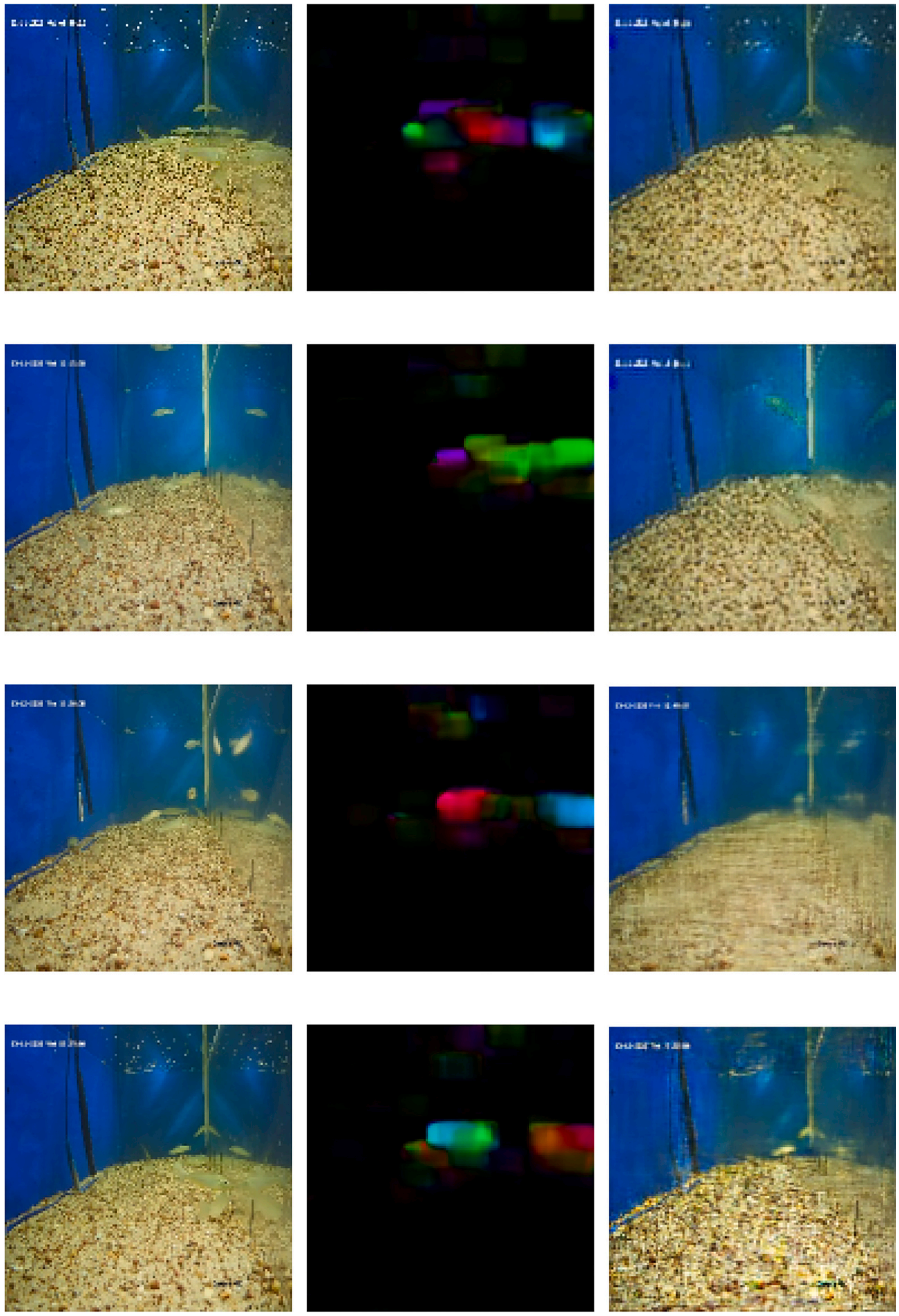


Fig. 13. Sample results of the proposed model in reconstructing the images. First column represents the ground truth, the second column represents the corresponding optical flow and the third column represents the reconstructed image.

attention map. This multiplication acts as a type of attention mask by highlighting important motion information and attenuating irrelevant features, allowing the model to focus more effectively on the motion of the fish. Consequently, the proposed model outperformed the traditional approaches. Thus, the model focuses on the fishes and captures their frantic patterns. The proposed model outperforms the other methods with an accuracy of 0.68, precision of 0.70, recall of 0.68 and f1-score of 0.68. Fig. 15 compares the ground truth anomaly score with the model prediction. The red shade presents the ground truth scores, and the black line represents the model predictions. Also, the figure displays the frames corresponding to ground truth and reconstructed output. It is seen that the model performs considerably well in reconstructing the normal frames and poorly in reconstructing frames with frantic patterns.

3.5. Performance on aquaculture cage dataset

The study assesses the performance of the proposed model using the cage dataset. Table 6 summarizes the results obtained from this evaluation. The model demonstrated satisfactory performance in detecting frantic patterns on the cage dataset, achieving an accuracy of 0.58.

Table 4
Comparison of accuracy, precision, recall and f1-scores of the proposed model against various temporal window.The best time window is highlighted in bold.

Temporal Window	Accuracy	Precision	Recall	F1-score
(k)	0.61	0.63	0.58	0.60
(k), (k-2)	0.67	0.67	0.64	0.65
(k), (k-2), (k-4)	**0.68**	**0.70**	**0.68**	**0.68**
(k), (k-2), (k-4), (k-6)	0.64	0.66	0.64	0.64
(k), (k-2), (k-4), (k-6), (k-8), (k-10)	0.55	0.52	0.57	0.54

Table 5
Comparison of accuracy, precision, recall and f1-scores of the proposed model against related work.The best model is highlighted in bold.

Methods	Accuracy	Precision	Recall	F1-score
Two-Stream DSTAE (Li et al., 2021)	0.44	0.41	0.44	0.37
Two-Stream GAN (Ravanbakhsh et al., 2017)	0.62	0.63	0.62	0.62
Top Heavy Auto-encoder (Cruz-Esquivel and Guzman-Zavaleta, 2022)	0.60	0.56	0.60	0.57
Proposed Method	**0.68**	**0.70**	**0.68**	**0.68**

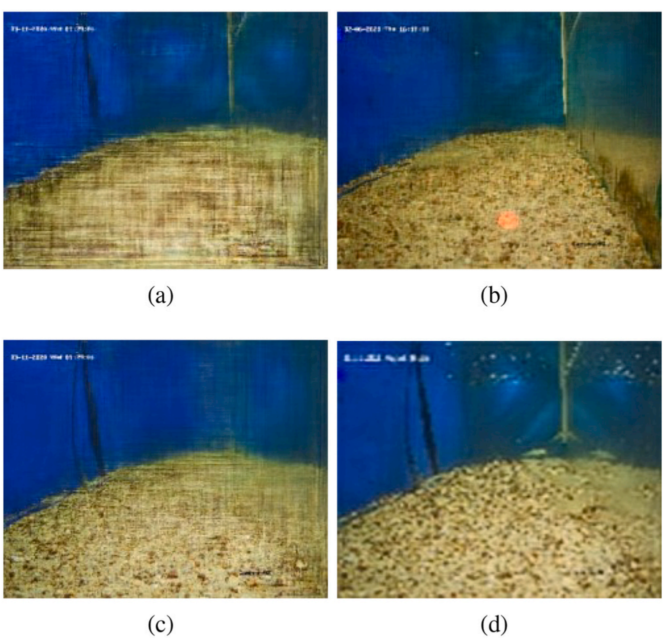


Fig. 14. Reconstructed frames of the proposed and other compared conventional methods. (a), (b) and (c) shows the results of Li et al. (2021); Ravanbakhsh et al. (2017) and Cruz-Esquivel and Guzman-Zavaleta (2022) respectively. The result of proposed model is shown in (d).

Fig. 16 displays sample ground truth and reconstructed frames. However, it is worth noting that the model's performance was comparatively better on the aquarium dataset than on the cage dataset. This disparity can be attributed to the absence of camera motion in the aquarium dataset, whereas the cage dataset contained camera motion. The presence of camera motion, combined with the natural movements of fish, sometimes led to misclassification of fish motion as frantic patterns. Improving the data collection process addresses this limitation and enhances the model's performance. Another noteworthy aspect is that the aquarium and cage datasets involved different fish species. It demonstrates that the proposed approach is not dependent on the species under consideration, further emphasizing the model's reliability in detecting frantic patterns in various cage scenarios. Thus, the results obtained validate the proposed model's reliability in detecting frantic patterns of fish in cage scenarios. However, there is room for improvement through refining data collection techniques and addressing the challenges introduced by camera motion in the cage dataset.

3.6. Aquaculture viewpoint

Monitoring water quality parameters and fish behaviour plays a significant role in aquaculture. The water quality monitoring system senses the minute changes in the ecosystem, and connecting these systems to the internet makes the system remotely deployable. In this regard, FPGA-based systems have several advantages, including high accuracy when handling big datasets and complex calculations in real-time. The calibration of the proposed system at regular intervals provides accurate results. Their capacity for parallel processing enables quick detection of and response to changes in water quality, lowering risks to aquatic life. These systems are easily adaptable to different scales since adding sensors and features does not degrade the performance. Also, the suggested FPGA-based system is more affordable than existing alternatives. Another significant benefit is customization, which enables operators to adapt monitoring algorithms to their particular requirements. Thus, real-time monitoring, high accuracy, and scalability of FPGA-based systems make them ideal for considerable aquaculture facilities with varied prerequisites. However, the existing systems are expensive and are suitable for small-scale aquaculture operations as these systems come with a fixed number of in/out ports for sensors. Thus, these systems lack customization and scalability for large-scale aquaculture facilities.

A prediction system for monitoring the water quality makes it more beneficial to the fishermen. It allows the fishermen to be aware of their culture system in advance (at least 90 min) and take preventive measures. The outlier detection system determines the outliers in the predictions, allowing the fishermen to make more reasonable decisions regarding aquaculture. The proposed methodology integrates the water quality monitoring system with the fish behaviour detection system. The fish behaviour system detects frantic movement patterns usually observed under stressful conditions. Fishermen have limited access to their culture system and cannot monitor the system 24/7. Alternatively, the fish behaviour detection system can reliably monitor the fish for the fishermen. Thus, the model provides better decisions on the culture system by monitoring the water quality and fish behaviour 24/7. Consequently, the developed Decision Support System (DSS) benefits the fishermen by better managing the culture systems and improving the economic gains.

3.7. Contributions to Sustainable Development Goals

The current study proposes a novel method for detecting abnormal behaviour in aquaculture, combining Water quality parameters with fish motion to support the UN's Sustainable Development Goals (SDGs). By emphasising aquaculture as a sustainable method of fishing, the research is in line with SDG 14 (Life Below Water). The study contributes to the goal of effectively regulating harvesting by detecting abnormal fish behaviour using the proposed method. Furthermore, the development of an FPGA-based water quality monitoring system and prediction model directly links the study to SDG 9 (Industry, Innovation, and Infrastructure). The study advances research and promotes resource efficiency in aquaculture, utilising technology and learning algorithms. Including these techniques in a decision support system for aquaculture monitoring has the potential to support both sustainable fishing methods and economic gains.

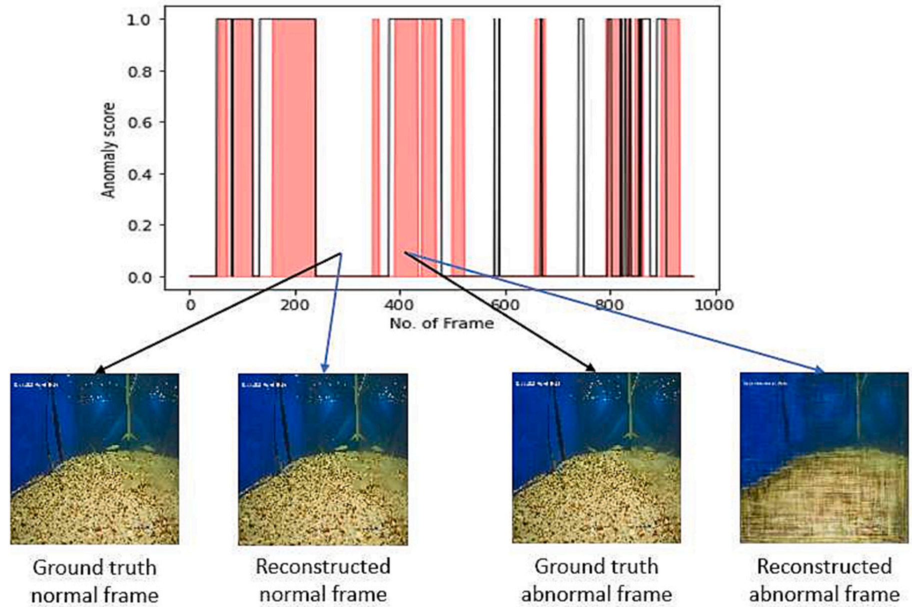


Fig. 15. Comparison of ground truth anomaly score with the model prediction. The red region shows the ground truth anomalous frames. The black line shows the model predictions. Corresponding ground truth and reconstructed frames are shown in the second row.

Table 6
Comparison of accuracy, precision, recall and f1-scores of the proposed model on aquarium and cage dataset.

Dataset	Accuracy	Precision	Recall	F1-score
Aquarium	0.68	0.70	0.68	0.68
Cage	0.59	0.76	0.60	0.67

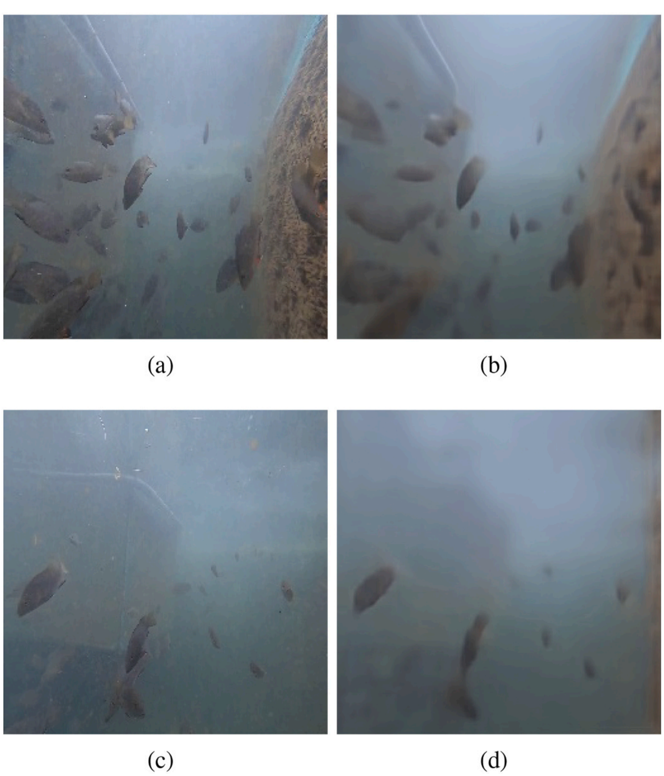


Fig. 16. Sample results of the proposed model on cage dataset. The left column shows the groundtruth frames and the right column shows the corresponding reconstructed frames.

4. Conclusion

The present study develops an efficient system for monitoring water quality parameters and fish in aquaculture. In this regard, the study utilizes a large aquarium with 8 to 10 *Sillago sihama* fish. It further designs an FPGA-based water quality monitoring system with a wireless connection module to transmit data to the cloud. The proposed method uses a learning algorithm to forecast water quality parameters. To this end, the study uses a bi-directional LSTM module trained on the water quality data collected from the aquarium. Subsequently, the approach employs a box and whisker plot to establish thresholds for the water quality parameters. The detection of outliers triggers the behaviour analysis model, which detects anomalous (frantic) behaviours over a predetermined time frame. The behaviour analysis model uses auto-encoder-based reconstruction methodology to identify frantic patterns of fishes. Consequently, the study integrates water quality parameters with the behaviour analysis model to enhance the monitoring of aquaculture systems.

The study proposes a novel fish dataset with annotations for frantic patterns of *Sillago sihama*. The results highlight the dependability of the proposed dataset for developing pattern analysis algorithms. The FPGA-based water quality monitoring system is seen as reliable for sampling water quality parameters. The study demonstrated the effectiveness of bi-directional LSTM over other LSTM models for predicting water quality parameters. Also, the results indicate that the bi-directional LSTM model with Adam optimizer and MSE loss function has better RMSE scores of 1.01. The study highlights the effectiveness of the box and whisker plot for identifying the threshold values of the water quality parameters in aquaculture.

The auto-encoder-based reconstruction approach outperforms the conventional methods for detecting the pattern with an accuracy of 0.68, precision of 0.70, recall of 0.68 and f1-score of 0.68. It is significant to note that the proposed pattern analysis methodology is computationally efficient as it does not require training in object detection and tracking to capture the behaviour of fish. In addition, the optical flow-based attention positively influenced the results, highlighting the necessity of localizing the region of interest for the model during training. Further, it achieved an accuracy of 0.59 on the cage dataset, thereby showcasing the reliability of the proposed model in real-time scenarios. The fisherman can act quickly since the use of learning algorithms allows them to know the circumstances of the culture system well in advance. Finally, the proposed methodology to monitor the water quality parameters and the behaviour of fish is reliable and can aid in preventing economic losses for the fishermen.

Compliance with ethical standards

All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Abinaya, N., Susan, D., Kumar, R., 2021. Naive bayesian fusion based deep learning networks for multisegmented classification of fishes in aquaculture industries. Ecol. Inform. 61, 101248.

Adiono, T., Toha, A.M., Pamungkas, S., Sutisna, N., Sumiarsih, E., 2021. Internet of things for marine aquaculture. In: 2021 International Symposium on Electronics and Smart Devices (ISESD). IEEE, pp. 1–5.

Ahmed, A.N., Othman, F.B., Afan, H.A., Ibrahim, R.K., Fai, C.M., Hossain, M.S., Ehteram, M., Elshafie, A., 2019. Machine learning methods for better water quality prediction. J. Hydrol. 578, 124084.

Alam, A.U., Clyne, D., Deen, M.J., 2021. A low-cost multi-parameter water quality monitoring system. Sensors 21 (11), 3775.

AlZu'bi, H., Al-Nuaimy, W., Buckley, J., Sneddon, L., Young, I., 2015. Real-time 3d fish tracking and behaviour analysis. In: 2015 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT). IEEE, pp. 1–5.

Bai, Y.-X., Zhang, S.-H., Fan, Z., Liu, X.-Y., Zhao, X., Feng, X.-Z., Sun, M.-Z., 2018. Automatic multiple zebrafish tracking based on improved hog features. Scient. Rep. 8 (1), 1–14.

Bergmann, P., Meinhardt, T., Leal-Taixe, L., 2019. Tracking without bells and whistles. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 941–951.

Bertolini, C., Capelle, J., Royer, E., Milan, M., Witbaard, R., Bouma, T., Pastres, R., 2022. Using a clustering algorithm to identify patterns of valve-gaping behaviour in mussels reared under different environmental conditions. Ecol. Inform. 69, 101659.

Bhaskaran, R., Baskaran, R.K., Vijayalakshmi, C., 2019. Machine learning technique for analyzing the behavior of fish in an aquarium. In: International Conference on Modeling, Machine Learning and Astronomy. Springer, pp. 55–65.

Butail, S., Paley, D.A., 2012. Three-dimensional reconstruction of the fast-start swimming kinematics of densely schooling fish. J. Roy. Soc. Interface 9 (66), 77–88.

Butail, S., Salerno, P., Bollt, E.M., Porfiri, M., 2015. Classification of collective behavior: a comparison of tracking and machine learning methods to study the effect of ambient light on fish shoaling. Behav. Res. Methods 47 (4), 1020–1031.

Chan, W., Park, D., Lee, C., Zhang, Y., Le, Q., Norouzi, M., 2021. Speechstew: Simply mix all available speech recognition data to train one large neural network. arXiv preprint arXiv:2104.02133.

Chavan, M., Patil, M.V.P., Chavan, S., Sana, S., Shinde, C., 2018. Design and implementation of iot based real time monitoring system for aquaculture using raspberry pi. Int. J. Recent and Innovat. Trends Comput. Commun. 6 (3), 159–161.

Chou, J.-S., Ho, C.-C., Hoang, H.-S., 2018. Determining quality of water in reservoir using machine learning. Ecol. Inform. 44, 57–75.

Cruz-Esquivel, E., Guzman-Zavaleta, Z.J., 2022. An examination on autoencoder designs for anomaly detection in video surveillance. IEEE Access 10, 6208–6217.

de Vargas Guterres, B., da Silveira Guerreiro, A., Sandrini, J.Z., da Costa Botelho, S.S., 2020. Feasibility of visual signals on the construction of biosensors based on behavioral analysis of perna mussels. Ecol. Inform. 59, 101118.

Duangwongsa, J., Ungsethaphand, T., Akaboot, P., Khamjai, S., Unankard, S., 2021. Real-time water quality monitoring and notification system for aquaculture. In: 2021 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering. IEEE, pp. 9–13.

Girisha, S., Verma, U., Pai, M.M., Pai, R.M., 2021. Uvid-net: Enhanced semantic segmentation of uav aerial videos by embedding temporal information. IEEE J. Select. Top. Appl. Earth Observ. Remote Sens. 14, 4115–4127.

Gokulanathan, S., Manivasagam, P., Prabu, N., Venkatesh, T., 2019. A gsm based water quality monitoring system using arduino. Shanlax Int. J. Arts, Sci. Human. 6 (4), 22–26.

Gopi, A.P., Naik, K.J., 2021. A model for analysis of iot based aquarium water quality data using cnn model. In: 2021 International Conference on Decision Aid Sciences and Application (DASA). IEEE, pp. 976–980.

Haq, K.R.A., Harigovindan, V., 2022. Water quality prediction for smart aquaculture using hybrid deep learning models. IEEE Access 10, 60078–60098.

Hinton, G.E., Zemel, R., 1993. Autoencoders, minimum description length and helmholtz free energy. Adv. Neural Inform. Process. Syst. 6.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780.

Hou, H., Li, B., Zhao, X., Shao, X., Liu, Z., Yan, Y., Wang, Z., Tao, Y., Yang, X., 2011. Proteins characteristics and lipid profiles of silver sillago (sillago sihama). In: 2011 International Conference on Remote Sensing, Environment and Transportation Engineering. IEEE, pp. 7738–7741.

Hulse, S.V., Renoult, J.P., Mendelson, T.C., 2022. Using deep neural networks to model similarity between visual patterns: Application to fish sexual signals. Ecol. Inform. 67, 101486.

Jalal, A., Salman, A., Mian, A., Shortis, M., Shafait, F., 2020. Fish detection and species classification in underwater environments using deep learning with temporal information. Ecol. Inform. 57, 101088.

Komarudin, M., Septama, H., Yulianti, T., Yudamson, A., Hendri, J., Arafat, M., 2021. Multi node sensors for water quality monitoring towards precision aquaculture. In: IOP Conference Series: Earth and Environmental Science, volume 739. IOP Publishing, p. 012026.

Koskinopoulou, M., Raptopoulos, F., Papadopoulos, G., Mavrakis, N., Maniadakis, M., 2021. Robotic waste sorting technology: Toward a vision-based categorization system for the industrial robotic separation of recyclable waste. IEEE Robot. Autom. Magaz. 28 (2), 50–60.

Lee, J.-H., Wu, M.-Y., Guo, Z.-C., 2010. A tank fish recognition and tracking system using computer vision techniques. In: 2010 3rd International Conference on Computer Science and Information Technology, volume 4. IEEE, pp. 528–532.

Li, T., Bo, W., Hu, C., Kang, H., Liu, H., Wang, K., Fu, H., 2021. Applications of deep learning in fundus images: A review. Med. Image Anal. 69, 101971.

Li, T., Chen, X., Zhu, F., Zhang, Z., Yan, H., 2021. Two-stream deep spatial-temporal auto-encoder for surveillance video abnormal event detection. Neurocomputing 439, 256–270.

Li, W., Li, F., Li, Z., 2022. Cmftnet: Multiple fish tracking based on counterpoised jointnet. Comput. Electron. Agric. 198, 107018.

Lin, J.-Y., Tsai, H.-L., Lyu, W.-H., 2021. An integrated wireless multi-sensor system for monitoring the water quality of aquaculture. Sensors 21 (24), 8179.

Liu, H., Liu, T., Gu, Y., Li, P., Zhai, F., Huang, H., He, S., 2021. A high-density fish school segmentation framework for biomass statistics in a deep-sea cage. Ecol. Inform. 64, 101367.

Liu, J., Zhang, L., Li, Y., Liu, H., 2023. Deep residual convolutional neural network based on hybrid attention mechanism for ecological monitoring of marine fishery. Ecol. Inform. 77, 102204.

Long, L., Johnson, Z.V., Li, J., Lancaster, T.J., Aljapur, V., Streelman, J.T., McGrath, P.T., 2020. Automatic classification of cichlid behaviors using 3d convolutional residual networks. Iscience 23 (10), 101591.

Nguyen, N.H., Nguyen, H.X., Le, T.T., Vu, C.D., et al., 2021. Evaluating low-cost commercially available sensors for air quality monitoring and application of sensor calibration methods for improving accuracy. Open Journal of Air Pollution 10 (01), 1.

Pache, M.C.B., Sant'Ana, D.A., Rozales, J.V.A., de Moraes Weber, V.A., Junior, A.d.S.O., Garcia, V., Pistori, H., Naka, M.H., 2022. Prediction of fingerling biomass with deep learning. Ecol. Inform. 71, 101785.

Papadakis, V.M., Glaropoulos, A., Kentouri, M., 2014. Sub-second analysis of fish behavior using a novel computer-vision system. Aquacultural engineering 62, 36–41.

Papadakis, V.M., Papadakis, I.E., Lamprianidou, F., Glaropoulos, A., Kentouri, M., 2012. A computer-vision system and methodology for the analysis of fish behavior. Aquacultural engineering 46, 53–59.

Pasika, S., Gandla, S.T., 2020. Smart water quality monitoring system with cost-effective using iot. Heliyon 6 (7), e04096.

Pedersen, M., Haurum, J.B., Bengtson, S.H., and Moeslund, T.B. (2020). 3d-zef: A 3d zebrafish tracking benchmark dataset. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 2426–2436.

Pinkiewicz, T., Purser, G., Williams, R., 2011. A computer vision system to analyse the swimming behaviour of farmed fish in commercial aquaculture facilities: A case study using cage-held atlantic salmon. Aquacult. Eng. 45 (1), 20–27.

Putra, T., Surtono, A., Puazi, G., Suciyati, S., Firdaus, I., Rumiyanti, L., et al. (2021). Measurement of physical parameters of water quality in real-time based on arduino. In Journal of Physics: Conference Series, volume 1751, page 012067. IOP Publishing.

Qian, Z.-M., Wang, S.H., Cheng, X.E., Chen, Y.Q., 2016. An effective and robust method for tracking multiple fish in video image based on fish head detection. BMC bioinformatics 17 (1), 1–11.

Rashid, M.M., Nayan, A.-A., Simi, S., Saha, J., Rahman, M.O., Kibria, M.G., 2021. Iot based smart water quality prediction for biofloc aquaculture. Int. J. Adv. Comput. Sci. Appl. 12 (6).

Ravanbakhsh, M., Nabi, M., Sangineto, E., Marcenaro, L., Regazzoni, C., and Sebe, N. (2017). Abnormal event detection in videos using generative adversarial nets. In 2017 IEEE International Conference on Image Processing (ICIP), pages 1577–1581.

Rodriguez, A., Rico-Diaz, A.J., Rabunal, J.R., Puertas, J., Pena, L., 2015. Fish monitoring and sizing using computer vision. In: International work-conference on the interplay between natural and artificial computation. Springer, pp. 419–428.

Saberioon, M., Cisar, P., 2016. Automated multiple fish tracking in three-dimension using a structured light sensor. Computers and Electronics in Agriculture 121, 215–221.
Saini, J., Dutta, M., Marques, G., 2021. Sensors for indoor air quality monitoring and assessment through internet of things: a systematic review. Environ. Monit. Assess. 193 (2), 1–32.
Schuster, M., Paliwal, K., 1997. Bidirectional recurrent neural networks. IEEE Trans. Signal Process. 45 (11), 2673–2681.
Shreesha, S., MM, M.P., Pai, R.M., and Verma, U. (2021a). Lstm-based prediction of water quality parameters system in backwaters. In 2021 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), pages 1–6. IEEE.
Shreesha, S., MM, M.P., Verma, U., and Pai, R.M. (2020). Computer vision based fish tracking and behaviour detection system. In 2020 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), pages 252–257. IEEE.
Shreesha, S., Pai, M.M., Verma, U., Pai, R.M., 2023. Fish tracking and continual behavioral pattern clustering using novel sillago sihama vid (ssvid). IEEE Access 11, 29400–29416.
Shreesha, S., Ujjwal, V., MM, M.P., and Pai, R.M. (2022). Fish tracking and unsupervised fish behaviour analysis of sillago sihama. In CV4Animals: Computer Vision for Animal Behavior, Tracking and Modeling, page Track III:Journal Track. CVPR.
Shreesha, S., Verma, U., Pai, R.M., Girisha, S., et al., 2021b. Behavioural pattern analysis of fishes for smart aquaculture: An object centric approach. In: TENCON 2021–2021 IEEE Region 10 Conference (TENCON). IEEE, pp. 917–922.
Simitha, K., Raj, S., 2019. Iot and wsn based water quality monitoring system. In: 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA). IEEE, pp. 205–210.
Singh, S., Rai, S., Singh, P., Mishra, V.K., 2022. Real-time water quality monitoring of river ganga (india) using internet of things. Ecol. Inform. 71, 101770.
Sun, H., Yue, J., Li, H., 2022. An image enhancement approach for coral reef fish detection in underwater videos. Ecol. Inform. 72, 101862.
Thai, T.T.N., Nguyen, T.S., Pham, V.C., 2021. Computer vision based estimation of shrimp population density and size. In: 2021 International symposium on electrical and electronics engineering (ISEE). IEEE, pp. 145–148.
Vasudevan, S.K., Baskaran, B., 2021. An improved real-time water quality monitoring embedded system with iot on unmanned surface vehicle. Ecol. Inform. 65, 101421.
Xia, C., Chon, T.-S., Liu, Y., Chi, J., Lee, J., 2016. Posture tracking of multiple individual fish for behavioral monitoring with visual sensors. Ecol. Inform. 36, 190–198.
Xu, J., Liu, Y., Cui, S., Miao, X., 2006. Behavioral responses of tilapia (oreochromis niloticus) to acute fluctuations in dissolved oxygen levels as monitored by computer vision. Aquacultural engineering 35 (3), 207–217.
Yang, L., Liu, Y., Yu, H., Fang, X., Song, L., Li, D., Chen, Y., 2021. Computer vision models in intelligent aquaculture with emphasis on fish detection and behavior analysis: a review. Archives of Computational Methods in Engineering 28 (4), 2785–2816.
Yin, H., Cao, Y., Marelli, B., Zeng, X., Mason, A.J., Cao, C., 2021. Soil sensors and plant wearables for smart and precision agriculture. Adv. Mater. 33 (20), 2007764.
Yu, G., Cai, R., Su, J., Hou, M., Deng, R., 2023. U-yolov7: A network for underwater organism detection. Ecol. Inform. 75, 102108.
Zheng, Z., Ding, H., Weng, Z., Wang, L., 2023. Research on a multiparameter water quality prediction method based on a hybrid model. Ecol. Inform. 76, 102125.