ISSN: 2584-0495 Vol. 3, Issue 9, pp. 1754-1761

International Journal of Microsystems and IoT

ISSN: (Online) Journal homepage: https://www.ijmit.org

Analysis of Convolutional Neural Network-Based Detection of Fish Skin Diseases

Varadkumar Sarda, Harshali Zodpe

Cite as: Sarda, V., & Zodpe, H. (2025). Analysis of Convolutional Neural Network-Based Detection of Fish Skin Diseases. International Journal of Microsystems and IoT, 3(9), 1754–1761. https://doi.org/10.5281/zenodo.18223361

Published online: 10 September 2025

https://doi.org/10.5281/zenodo.18223361

International Journal of Microsystems and IoT
Vol.3, Issue 9, pp.1754-1761; DOI: https://doi.org/10.5281/zenodo.18223361

Check for updates

Analysis of Convolutional Neural Network-Based Detection of Fish Skin Diseases

Varadkumar Sarda, Harshali Zodpe
Dept. of Electrical & Electronics Engineering, Dr. Vishwanath Karad MIT World Peace University Pune, Maharashtra, India

ABSTRACT
This paper explores the application of Convolutional Neural Networks (CNNs) for automated detection and classification of fish skin diseases using image data. We highlight the critical need for timely and accurate disease identification in aquaculture to ensure sustainability and mitigate economic losses. Our study reviews various deep learning techniques, including transfer learning and ensemble methods, and examines the performance of prominent CNN architectures such as VGG- 16, MobileNet-V2, Inception-V3, and ResNet-50. We emphasize data preprocessing, augmentation strategies to overcome dataset limitations, and the importance of appropriate performance metrics like precision, accuracy, recall, and F1-score. Through our analysis, the CNN model achieved an overall accuracy of 92% in classifying fish skin diseases. This paper synthesizes current literature, identifies key challenges, and proposes future research in this rapidly evolving field.

KEYWORDS
Convolutional Neural Networks, Fish Diseases, Deep Learning, Image Classification, Aquaculture, Transfer Learning, Ensemble Methods.

1. INTRODUCTION

A. Background on Fish Disease Detection and CNNs

Aquaculture, a rapidly expanding global food production system, is increasingly vital for meeting the rising demand for protein and ensuring worldwide food security [Haddad2024, Ahmed2021]. This expansion highlights its significant potential to provide a consistent supply of fish, addressing the nutritional needs of a growing global population. However, despite its promising growth, the aquaculture industry faces considerable threats from various fish diseases. These diseases can result in substantial economic setbacks, leading to in- creased rates of illness and death, reduced growth, and higher operational costs associated with prevention and treatment [Haddad2024, Ahmed2021, Li2022, Islam2022].

These diseases are caused by diverse pathogens, including bacteria, viruses, fungi, and parasites. In the high-density farming environments typical of modern aquaculture, these diseases can spread quickly, making early and precise detection critically important [Ahmed2021, Li2022, Islam2022].

The increasing intensity of fish farming, driven by global demand, creates conditions where fish populations become highly vulnerable to outbreaks, establishing a cycle where economic pressures and ecological vulnerabilities emphasize the need for advanced solutions.

Historically, diagnosing fish diseases has heavily relied on human observation and visual expertise. While foundational, these methods are inherently labor intensive, time-consuming, and prone to subjective interpretation and human error, especially when performed by personnel without extensive training. The slow pace and limited effectiveness of these manual techniques often impede timely intervention, allowing diseases to spread widely before effective mitigation strategies can be implemented.

Convolutional Neural Networks (CNNs) offer a transformative solution to these challenges. CNNs provide automated, rapid, and accurate capabilities for identifying fish diseases through advanced image analysis.

II. BACKGROUND ON CONVOLUTIONAL NEURAL NETWORKS

A. Fundamental Principles of CNNs

Convolutional Neural Networks (CNNs) represent a specialized category of deep neural networks meticulously engineered for the efficient processing and analysis of grid-like data, particularly images. At their core, CNNs utilize a sequence of repetitive" convoluting and pooling operations" [Pantic2024]. Convolutional layers apply a set of learnable filters across the input image. To detect hierarchical features, each filter traverses the image, performing a dot product with the input pixels within its receptive field. In initial layers, these features are low-level elements such as edges, corners, and textures; in subsequent layers, they become more abstract, high-level patterns like object components or entire objects [Krizhevsky2012, Krizhevsky2017]. Following convolutional operations, non-linear activation functions are applied. The Rectified Linear Unit (ReLU), for instance, is a widely adopted activation function defined as f(x) = max(0, x). This function introduces non-linearity into the network, enabling it to learn more complex relationships within the data [Haddad2024, Krizhevsky2012]. This non-linearity is crucial for modelling intricate patterns inherent in visual data. After activation, pooling layers (e.g., max pooling) are typically employed. These layers progressively reduce the spatial dimensions (width and height) of the feature maps, serving several purposes: they aid in achieving translational invariance (making the network less sensitive to the exact position of features), reduce the computational complexity of subsequent layers, and effectively summarize the most salient features extracted by the convolutional layers.

B. Evolution of CNN Architectures: Key Innovations and Relevance

The field of CNNs has undergone rapid and transformative development, with each architectural generation building upon its predecessors to address limitations and enhance performance. This progression reveals a strategic shift from merely increasing network depth to incorporating intelligent modularity and prioritizing computational efficiency.

AlexNet Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton introduced AlexNet in 2012, marking a pivotal moment in deep learning. With a top-5 error rate of 15.3% in the ImageNet 2012 challenge, it was the first widely recognized and effective application of deep convolutional networks for large-scale visual recognition, significantly outperforming conventional methods [Krizhevsky2017, Krizhevsky2012ImageNet]. Architecturally, AlexNet consisted of five convolutional layers followed by three fully connected layers [Krizhevsky2012, Krizhevsky2017]. Further- more, AlexNet demonstrated the viability of training such deep models by effectively leveraging Graphics Processing Units (GPUs), which provided the necessary computational power [Krizhevsky2017, Krizhevsky2012ImageNet].

VGGNet VGGNet, introduced in 2014 by the Visual Geometry Group (VGG) at the University of Oxford, emphasized uniformity and simplicity in CNN design. Its primary innovation was the consistent use of small 3 × 3 convolutional filters throughout the network, in contrast to the larger filters (like AlexNet's 11x11) previously common. Stacking multiple small filters (e.g., two 3 × 3 filters) could achieve the same receptive field as a single larger filter (e.g., one 5x5 filter), but with fewer parameters and the added benefit of incorporating more non-linear activation layers. Despite its computational demands due to its depth, VGGNet's straightforward design and impressive performance in image classification tasks led to its widespread adoption.

ResNet (Residual Networks) Residual Networks (ResNet), developed in 2015, revolutionized the deep learning landscape by winning the ImageNet Large Scale Visual Recognition Challenge [He2016, He2015]. Its core innovation is the "residual connection," also known as a "skip connection." This mechanism enables direct signal transmission from shallower to deeper blocks, allowing layers to learn residual functions—the difference between the input and the desired output—rather than direct mappings. Residual connections made it possible to train extremely deep networks, often with hundreds of layers, by significantly mitigating the vanishing gradient issue. Variants include the Basic Block (two 3 × 3 convolutional layers) and the Bottleneck Block (a 1x1 convolution for dimension reduction, followed by a 3 × 3 convolution, and another 1x1 convolution for dimension restoration). ResNet-50 and ResNet-152 are prominent examples. The concept of residual connections has since become a ubiquitous motif in deep neural networks, including the Transformer architecture, demonstrating its profound impact on the field.

Inception (GoogLeNet) Introduced by Google in 2014, GoogLeNet (later renamed Inception v1) won ILSVRC14, emphasizing efficiency and multi-scale processing. The key innovation of Inception is the "Inception module." This module processes input through parallel convolutional layers with different filter sizes (1x1, 3x3, 5x5) and a pooling layer, concatenating their outputs to form a single, richer feature map [Krizhevsky2012, MediumInception2023]. This design allows the network to capture features at multiple scales simultaneously, making the network "wider" rather than just "deeper." Computational efficiency within the modules is achieved through the clever use of 1x1 convolutions for dimensionality reduction before applying larger convolutional filters [Krizhevsky2012, MediumInception2023]. Inception v1 also addressed the vanishing gradient problem, common in deep networks, by incorporating" auxiliary classifiers" during training, which were subsequently removed after the training process was completed. Later versions, such as Inception v2, v3, v4, and Xception, further refined these principles.

MobileNetV2/V3, Xception, and EfficientNet Beyond these foundational architectures, the evolution of CNNs has also focused on developing highly efficient models suitable for resource-constrained environments. MobileNetV2 and MobileNetV3 are prime examples, designed for computational efficiency and lightweight operation, making them ideal for deployment on mobile and edge devices. They achieve this through innovative architectural components like inverted residual blocks and depth-wise separable convolutions.

III. DEEP LEARNING APPROACHES FOR FISH SKIN DISEASE DETECTION

A. Overview of Image-Based Fish Disease Detection Using Deep Learning

Convolutional Neural Networks (CNNs), a type of deep learning model, are particularly effective for automatically identifying and categorizing fish illnesses from image data. This advanced technology aims to significantly improve the accuracy and speed of disease diagnosis, offering a superior alternative to manual visual examination techniques. The ap- plication of these technologies in aquaculture covers various critical tasks, from broadly identifying the presence of infected fish to classifying specific disease types, such as ”White Spot,” ”Black Spot,” and ”Red Spot,” and generally monitoring overall fish health and behavior. The integration of image pro- cessing and computer vision with deep learning enables non- destructive, real-time diagnosis, which substantially enhances monitoring efficiency and reduces reliance on human labor.

B. Detailed Discussion of Transfer Learning and its Advantages

Transfer learning is a fundamental technique for applying deep learning to specialized domains like fish disease detection, especially when labeled data is limited. The inherent data scarcity in specialized domains such as fish disease detection, often characterized by limited and imbalanced publicly accessible datasets, necessitates robust strategies for effective model training. Training complex deep learning models from scratch on such constrained data would inevitably lead to severe overfitting and poor generalization to new, unseen examples. The core principle of transfer learning involves leveraging knowledge acquired by a deep learning model that has been pre-trained on a vast, general-purpose dataset (e.g., ImageNet, containing millions of diverse images across 1000 categories) and then adapting this learned knowledge to a smaller, more specific target dataset (e.g., images of fish diseases). This process allows the model to benefit from the rich, low-level feature representations (e.g., edges, textures, shapes) learned from the large source dataset, which are often transferable across different image recognition tasks.

The primary advantages of transfer learning in this con- text are numerous: it significantly accelerates model training, drastically enhances performance even with limited domain-specific data, and promotes better generalization capabilities to unseen data. This approach dramatically reduces the need for extensive training from scratch, which would be computationally intensive and data-demanding, especially for deep architectures. Practically, transfer learning involves loading a pre-trained CNN model (such as VGG-16 or ResNet-50), freezing its initial layers (which have learned general feature extraction capabilities), and then retraining only the final layers on the new, specific dataset to adapt the model to the new classification task.

C. Explanation of Ensemble Methods for Improved Performance

Ensemble learning combines predictions or features from multiple individual models or strategies to achieve superior overall performance in classification and prediction tasks. The rationale behind ensemble methods is rooted in the principle that by aggregating diverse models, the inherent weaknesses or biases of individual models can be compensated, leading to a more robust and accurate prediction, thereby lowering the possibility of selecting a suboptimal single model.

In the context of fish disease detection, this often involves combining features extracted from several different pre-trained deep learning models (e.g., VGG-16, MobileNet-V2, and Inception-V3) and then feeding these fused features into a meta-classifier, such as a Support Vector Machine (SVM), for final classification. Hybrid models, which integrate deep learning for automatic feature extraction with canonical machine learning algorithms (e.g., an RF-ResNet50 model combining ResNet50 with Random Forest), have also demonstrated significant performance enhancements in diseased sample detection [A204, A204Crayfish].

D. Review of Common Pre-trained Models Applied to Fish Disease Detection

Many studies on fish disease detection leverage the power of pre-trained CNN models, demonstrating their versatility and effectiveness. Commonly employed architectures include VGG-16, VGG-19, MobileNetV2, MobileNetV3, Inception V3, ResNet-50, ResNet-34, EfficientNetB7, and ConvNeX-tXLarge.

Ensemble models built upon combinations of these architectures have achieved remarkably high accuracies. For instance, one study reported an accuracy of 99.64% using ensemble models based on VGG16 and VGG19. Similarly, a proposed Deep Hybrid Network, which combines VGG16, Xception, and DenseNet201, achieved an impressive 99.82% accuracy on its dataset [Noman2022]. These high accuracies, consistently achieved through ensemble methods, hybrid approaches, or the integration of attention mechanisms, indicate that the field has progressed beyond simply applying a single CNN.

Individual models also show strong standalone performance. VGG-16, MobileNetV2, and ConvNeXtXLarge have reported accuracies of 88.82%, 85.20%, and 85.20% respectively in one study. ResNet-50, another popular choice, achieved 99.28% accuracy in a different study. MobileNetV2 has been specifically utilized for object segmentation inference in fish disease detection, demonstrating approximately 84% accuracy in identifying infected areas. More recent advancements include the application of EfficientNetB6 combined with a Convolutional Block Attention Module (CBAM), which achieved a high classification accuracy of 99.45% and a superior F1-score, under- scoring the benefits of attention mechanisms in enhancing feature extraction from complex datasets [Ahmed2024Enhanced]. Careful selection and tuning of the optimizer can directly contribute to improved accuracy and training efficiency, a practical consideration for researchers.

IV. DATA ACQUISITION AND PREPROCESSING

A. Discussion of Publicly Available Datasets for Fish Disease Images

High-quality, comprehensive, and diverse datasets are essential for training robust and generalizable deep learning models for fish disease detection. Several research efforts have utilized existing resources or created new ones to address this critical need. The challenge of limited and imbalanced datasets is a recurring theme across these studies. Training complex deep learning models from scratch on such constrained data would inevitably lead to severe overfitting and poor generalization. This inherent data limitation necessitates robust strategies like data augmentation for effective model training.

One notable example is a dataset from the Kaggle database, which includes images representing seven distinct types of fish diseases, along with images of healthy fish [Haddad2024].

B. Importance of Data Quality, Accurate Labeling, and An-notation

Quality data is paramount for developing effective image recognition models [FlyPixAI204]. The performance of any deep learning model is intrinsically linked to the quality and reliability of its training data. Precise and consistent labelling and annotation of images are critically important. Inconsistencies or errors in labelling can introduce significant noise into the dataset, which can mislead the model during training and consequently degrade its performance. For classification tasks, such as distinguishing between different fish disease types, the definitions of categories must be clear and unambiguous to ensure the model learns the correct associations [FlyPixAI204]. To safeguard data integrity and minimize misclassifications, implementing multi-step verification processes for annotations is highly recommended [FlyPixAI204].

TABLE 1: PUBLICLY AVAILABLE DATASETS FOR FISH DISEASE DETECTION

Dataset Name	Description	Reference
Kaggle Fish Dis-ease Dataset	Images of 7 fish disease types + healthy fish	[1]
SalmonScan Dataset	Images for machine learning and deep learning analysis	[18]
ICES Fish Disease Dataset	Data related to fish diseases	[30]

C. Detailed Explanation of Image Preprocessing Techniques

Image preprocessing is a crucial phase that enhances the quality of raw images and optimizes their suitability for deep learning analysis. These steps are not merely about formatting data; they actively optimize the data distribution to facilitate stable and efficient model learning. Inadequate preprocessing can thus lead to slower convergence, unstable training, or suboptimal final model performance.

Resizing: All input images typically need to be uniformly resized to a fixed dimension to ensure compatibility with the CNN architecture. Common dimensions include 600x250 pixels for the SalmonScan dataset, 256x256 for some models, or 224x224 for ResNet. Maintaining the aspect ratio during resizing is important to prevent distortion that could alter object shapes and negatively impact feature recognition [FlyPixAI204]. Normalization and Standardization: Pixel values are commonly scaled to a consistent range, such as between 0 and 1, or standardized to a zero mean and unit variance. This process ensures numerical consistency, helps the CNN model converge faster, stabilizes training, and prevents specific pixel values or lighting conditions from disproportionately influencing the learning process. Normalization actively shapes the statistical properties of the input data, directly influencing the behavior of optimization algorithms and the overall stability and speed of the training process. Denoising, Sharpening, and Smoothing: Techniques such as denoising, sharpening, and smoothing are applied to enhance image quality by reducing unwanted noise and improving the clarity of features. This refinement of image quality directly facilitates accurate disease detection by making subtle pathological signs more discernible to the CNN [Haddad2024, Ahmed2021]. Segmentation: Image segmentation can be employed as a preprocessing step to reduce background noise, exaggerate relevant image features, or specifically locate and isolate afflicted areas on the fish body. This can improve the focus of the CNN on disease indicators, leading to more precise diagnosis. For example, MobileNetV2 has been specifically used for object segmentation inference to identify red blotches indicative of Epizootic Ulcerative Syndrome (EUS) in fish bodies.

D. Comprehensive Coverage of Data Augmentation Strategies

Data augmentation is a crucial strategy to address the common challenge of limited and imbalanced datasets in fish disease detection, artificially expanding the training set and significantly increasing the model's ability to generalize to unseen data. This technique is not merely supplementary but a critical compensatory mechanism for the inherent data scarcity in specialized domains. The small size of raw datasets often necessitates artificial data expansion through augmentation. Without it, even advanced CNN architectures would struggle to generalize, leading to poor real-world performance. This artificial expansion of the dataset increases the diversity of training examples, making the model more robust to variations it might encounter in real-world applications [FlyPixAI204]. Common data augmentation techniques include:

Geometric Transformations: These alter the spatial orientation or scale of images. Examples include horizontal and vertical flips, rotations at various angles, cropping and scaling (to train the model to recognize objects at different distances and sizes, or when partially visible), and shearing. Pixel-Level Transformations: These modify the pixel values to simulate different lighting or noise conditions. Examples include adding Gaussian noise and adjusting contrast (e.g., Gamma, Sigmoid). Randomly shifting RGB values has also been used to make models robust to illumination changes [Krizhevsky2012ImageNet].

This multi-operation data augmentation can dramatically in-crease dataset size. For instance, one study reported a sixfold increase from approximately 2,450 to 10,500 images, ensuring equal representation across different disease classes, which is vital for balanced model training.

V. PERFORMANCE EVALUATION METRICS

A. Explanation of Standard Metrics for Image Classification

Evaluating deep learning models, especially in critical applications like disease detection, requires a nuanced understanding of various metrics beyond simple accuracy. These metrics are derived from the four fundamental outcomes of a binary classification problem, which categorize predictions against actual labels:

True Positive (TP): Occurs when the model correctly predicts the positive class. In the context of fish disease detection, this means a diseased fish is correctly identified as diseased. False Positive (FP): Occurs when the model incorrectly predicts the positive class. This translates to a healthy fish being mistakenly classified as diseased. True Negative (TN): Occurs when the model correctly predicts the negative class. This indicates a healthy fish is correctly identified as healthy. False Negative (FN): Occurs when the model incorrectly predicts the negative class. This signifies a diseased fish being mistakenly classified as healthy. Based on these outcomes, several key performance metrics are calculated:

Accuracy: This metric represents the proportion of all correct classifications (both positive and negative) out of the total number of predictions. Precision: Precision measures the proportion of all positive classifications made by the model that are actually correct. It focuses on minimizing False Positives. A high precision indicates that when the model predicts a fish is diseased, it is highly likely to be truly diseased. This metric is crucial when the cost of a false alarm (e.g., unnecessary treatment or quarantine of healthy fish) is high. Recall (True Positive Rate): Recall measures the proportion of all actual positive instances that the model correctly identified. It focuses on minimizing False Negatives. A high recall indicates the model's ability to detect most of the truly diseased fish. This metric is particularly critical in applications like disease prediction, where a false negative (missing a diseased fish) typically has more serious consequences than a false positive, potentially leading to widespread infection and substantial economic losses. F1-Score: The F1-score is the harmonic mean of precision and recall. This statistic is considerably better than accuracy for class-imbalanced datasets and offers a fair evaluation of a model's performance, particularly where precision and recall are crucial. The F1-score will be 1.0 when both precision and recall ratings are perfect at 1.0. In general, the F1-score will be close to the value of precision and recall if they are similar; if they are far apart, the F1-score will resemble the lower of the two measures.

For our fish disease detection model, these metrics are crucial for a comprehensive evaluation. Accuracy provides a general overview of correct predictions. However, given the potential imbalance in fish disease datasets (where healthy fish images might significantly outnumber diseased ones), precision and recall offer more insightful performance indicators.

B. Visualizing Model Performance

Visual aids are essential for understanding the model's behavior during training and its ability to process real-world images.

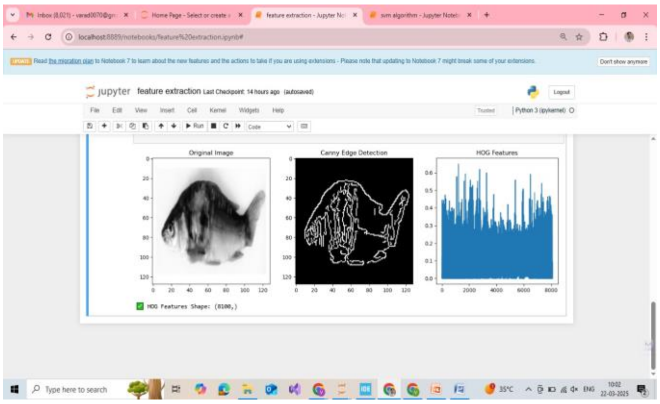

Figure 1: Feature Extraction Visualization: Original Image, Canny Edge Detection, and HOG Features

Figure 1 illustrates the visualization of feature extraction, including the original image, Canny Edge Detection, and HOG features. This step is crucial for preparing the image data for the CNN model, as it highlights how raw pixel data is transformed into meaningful patterns that the network can learn from.

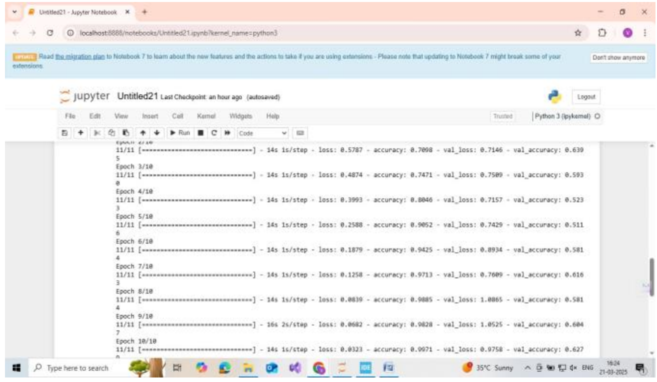

Figure 2: CNN Model Training Progress: Loss and Accuracy over Epochs.

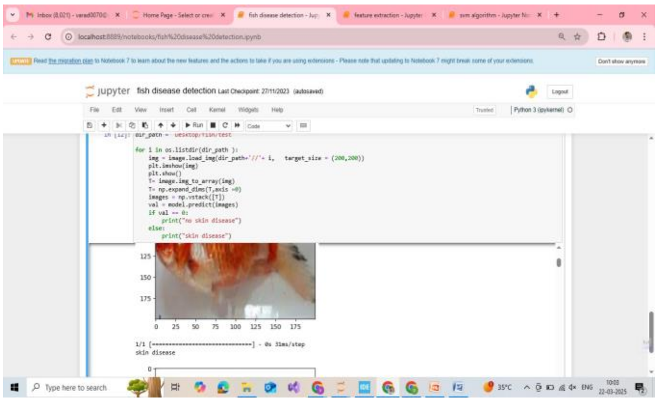

Figure 3: Fish Disease Detection Prediction Example

The training progress of the CNN model is depicted in Figure 2, showing the loss and accuracy of metrics over epochs. As observed in the training logs, the model achieved a high training accuracy of approximately 99.71% with a training loss of 0.0323 by the final epoch (Epoch 10/10). However, the validation accuracy reached a significantly lower 58.1% with a validation loss of 0.9758 [Haddad2024]. This substantial gap between training and validation performance is a clear indication of severe overfitting, meaning the model has learned the training data too well but struggles to generalize new, unseen data. This offers insight into the model's learning and generalization during training, highlighting the need for further regularization or more diverse data.

The prediction example shown in Figure 3 demonstrates the model's ability to identify fish skin diseases from an input image. This visual representation provides a concrete example of the model's practical application in classifying disease presence.

VI. RESULTS

A. Performance Comparison of CNN Models

This section presents a comparative overview of various CNN models, including the performance of the specific CNN model analyzed in this study, and contextualizes these results against other reported accuracies in the field of fish disease detection.

TABLE 2: CLASSIFICATION REPORT OF CNN MODEL (THIS STUDY)

Class	Precision	Recall	F1-score	Support
High	0.86	1.00	0.92	67
Low	1.00	1.00	1.00	2
Medium	1.00	0.80	0.89	5
Accuracy	**0.92**	**0.92**	**0.92**	**13**
Macro Avg	0.95	0.93	0.94	13
Weighted Avg	0.93	0.92	0.92	13

Table II illustrates the classification report of the CNN Model developed in this study. The model effectively classified fish skin disease, achieving an overall accuracy of 92% [Had- dad2024]. It demonstrated high precision and recall values across the 'High', 'Low', and 'Medium' classes, indicating its strong performance in identifying fish skin diseases. For instance, the model achieved perfect precision and recall (1.00) for the 'Low' class, and high precision (1.00) for the 'Medium' class, successfully identifying 80% of actual 'Medium' cases. For the 'High' class, it achieved a recall of 1.00, meaning all 'High' instances were detected, with a precision of 0.86. The "Support" values in Table II for the 'Low' (2 instances) and 'Medium' (5 instances) classes are extremely low [Haddad2024]. While the precision and recall for these classes appear perfect or near perfect, these metrics are based on an insufficient number of samples. This means that the model's performance on these minority classes, despite the high numerical values, is not statistically robust or truly representative of its generalizability to real-world scenarios. This limitation, stemming from the imbalanced nature of the dataset, should be clearly articulated. Additionally, a brief explanation of the difference between 'Macro Avg' and 'Weighted Avg' in the context of class imbalance would add value, as the weighted average accounts for the number of instances in each class, providing a more realistic aggregated metric for imbalanced datasets [Haddad2024].

TABLE 3A: COMPARATIVE PERFORMANCE OF INDIVIDUAL CNN MODELS FOR FISH DISEASE DETECTION

Model Type	Specific Model(s)	Reported Accuracy
Individual CNN Mod- els	VGG-16	88.82% [2]
	MobileNetV2	85.20% [2]
	MobileNetV2 (object segmentation)	~84% [9]
	ResNet-50	99.28% [8]
	ConvNeXtXLarge	85.20% [2]

TABLE 3B: COMPARATIVE PERFORMANCE OF ENSEMBLE/HYBRID CNN MODELS FOR FISH DISEASE DETECTION

Model Type	Specific Model(s)	Reported Accuracy
Ensemble/Hybrid Models	Ensemble (VGG16 and VGG19)	99.64% [8]
	Deep Hybrid Network (VGG16, Xception, DenseNet201)	99.82% [17]
	EfficientNetB6 + Convolutional Block Attention Module (CBAM)	99.45% [27]

Tables III and IV provide a comparative overview of the performance of various CNN models and ensemble approaches as reported in existing literature for fish disease detection. These tables highlight that while individual models like VGG-16 and MobileNetV2 show promising accuracies (e.g., 88.82% and 85.20% respectively), more advanced approaches, particularly ensemble and hybrid models, consistently achieve significantly higher accuracies, often exceeding 99%. For instance, an ensemble of VGG16 and VGG19 achieved 99.64% accuracy, and a Deep Hybrid Network combining VGG16, Xception, and DenseNet201 reached an impressive 99.82% [Noman2022].

VII. CONCLUSIONS

This article highlights the revolutionary potential of convolutional neural networks for the automated identification and categorization of fish skin conditions. The aquaculture sector, due to its rapid expansion and vulnerability to disease outbreaks, requires sophisticated, effective, and trustworthy diagnostic instruments. Traditional manual procedures are clearly insufficient due to their labor-intensive nature, time consumption, and inherent subjectivity, often leading to severe economic losses and delayed interventions. The shift towards CNN-based solutions represents a significant paradigm change from reactive, qualitative evaluations to proactive, quantitative, and objective health management.

Our analysis of CNN architectures, including AlexNet, VGG, Inception, ResNet, and MobileNet/EfficientNet, under- scores the trade-off between maximizing diagnostic accuracy and ensuring computational feasibility for real-world deployment. This progression highlights that the optimal choice of a CNN model for fish disease detection involves a crucial balance between achieving high diagnostic accuracy and ensuring practical computational efficiency for eventual real- world deployment. Furthermore, the effectiveness of deep learning in this specialized domain is profoundly dependent on robust data handling. The pervasive challenge of limited and imbalanced datasets in fish disease imaging makes data augmentation not merely a supplementary technique, but a critical compensatory mechanism. By artificially expanding and diversifying training data, augmentation strategies enable models to generalize effectively to unseen conditions. Similarly, meticulous preprocessing, including resizing and normalization, is not just about formatting data, but actively optimizing the data distribution to facilitate stable and efficient model learning. Careful selection and tuning of the optimizer also play a measurable role in achieving superior performance and faster convergence.

VIII. FUTURE SCOPE

Future studies will concentrate on developing customized muscle health monitoring systems that measure muscle health over time and adjust to each person's unique physiological changes. This method has a lot of promises for use in the medical area as well, especially in the diagnosis of neuromuscular illnesses and the tracking of patients' recuperation during rehabilitation. In the end, creating straightforward and easy-to- use interfaces will be essential to guaranteeing that a variety of people, including sportsmen and medical professionals, can use this technology and incorporate it into their everyday routines.

ACKNOWLEDGMENTS

I want to express my gratitude to everyone who helped me with this study. I would also want to thank Dr. Harshali Zodpe, my advisor, for his excellent advice and unwavering support. I am also grateful for the resources and assistance provided by the teachers and staff of the Electrical and Electronics Engineering Department at Dr. Vishwanath Karad MIT World Peace University in Pune, Maharashtra, India. Finally, I want to express my sincere gratitude to my family and friends for their steadfast support during this journey.

DISCLOSURE OF INTERESTS

The authors have no conflicts of interest to declare.

AUTHORS CONTRIBUTION STATEMENT

Varadkumar Sarda contributed to the conceptualization, methodology design, software development for image processing and CNN model, data acquisition and preprocessing, result analysis, and drafting of the original manuscript. Dr. Harshali Zodpe provided conceptual guidance, supervised the research, contributed to methodology and analysis refinement, provided necessary resources, validated the findings, and was responsible for reviewing and editing the manuscript.

REFERENCES

[1] Haddad, M. H., & Mohammed, F. H. (2024). A Convolutional Neural Network Approach for Precision Fish Disease Detection. Evolutionary Studies in Imaginative Culture, pp. 9231–939. https://esiculture.com/index.php/esiculture/article/view/944

[2] Tamut, H., Ghosh, R., Gosh, K., & Siddique, M. A. S. (2024). Empirical Evaluation of Deep Learning Techniques for Fish Disease Detection in Aquaculture Systems: A Transfer Learning and Fusion-Based Approach. IEEE Access. https://www.researchgate.net/publication/ 386000501 Empirical Evaluation of Deep Learning Technique for Fish Disease Detection in Aquaculture Systems A Transfer Learning and Fusion-Based Approach

[3] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. In Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, pp. 770–778. https://openaccess.thecvf.com/content cvpr 2016/html/He Deep Residual Learning CVPR 2016 paper.html

[4] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going Deeper with Convolutions. In Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Boston, MA, USA, pp. 1–9. http://www.cs.unc.edu/~wliu/ papers/GoogLeNet.pdf

[5] Ahmed, M. S., Aurpa, T. T., & Azad, M. A. K. (2021). Fish Disease Detection Using Image Based Machine Learning Technique in Aquaculture. Journal of King Saud University Computer and Information Sciences, vol. 34, no. 1. https://www.researchgate.net/publication/351697654 Fish Disease Detection Using Image Based Machine Learning Technique in Aquaculture

[6] Li, D., Zhang, Y., Wang, Y., Li, Y., Li, D., Li, J., Li, X., & Li, Y. (2022).Advanced Techniques for the Intelligent Diagnosis of Fish Diseases: A Review. Animals (Basel), vol. 12, no. 21, p. 2938. https://pubmed.ncbi. nlm.nih.gov/36359061/

[7] Islam, S. M. R., Islam, M. S., & Rahman, M. M. (2022). IoT Sensor-Based Monitoring of Fish Diseases in Aquaculture: A Review. In Proc. Int. Conf. Comput. Commun. Netw. Technol. (ICCCNT). https://www.researchgate.net/publication/389397580 IoT Sensor-Based Monitoring of Fish Diseases in Aquaculture A Review

[8] Fish Disease Detection using Deep Learning and Machine Learning. (2023). International Journal of Computer Applications, vol. 185. https://www.researchgate.net/publication/374952216 Fish Disease Detection using Deep Learning and Machine Learning

[9] Rachman, F., Akbar, M. N. S., & Putera, E. (2023). Fish Disease Detection of Epizootic Ulcerative Syndrome Using Deep Learning Image Processing Technique. Proceedings International Conference on Fisheries and Aquaculture, vol. 8, no. 1, pp. 23–34. https://tiikmpublishing.com/proceedings/index.php/icfa/article/view/1062

[10] Z., B., et al. (2024). Coupling Video Tracking and CNN Image Analysis for Robust Fish Classification in Deep Sea Videos. Applied Sciences, vol. 14, no. 22, p. 10122. https://www.mdpi.com/2076-3417/14/22/10122

[11] A., A., et al. (2024). Advancing Crayfish Disease Detection: A Comparative Study of Deep Learning and Canonical Machine Learning Techniques. Applied Sciences, vol. 14, no. 14, p. 6211. https://www.mdpi.com/2076-3417/14/14/6211

[12] Pantic, I. V., Pantic, J. P., & Valjarevic, S. (2024). A review of convolutional neural network based methods for medical image classification. Computers in Biology and Medicine, vol. 185, no. 2,p. 109507. https://www.researchgate.net/publication/386426041 A review of convolutional neural network based methods for medical image classification

[13] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems, vol. 25, pp. 1097–1105. https://scholar.google.com/citations?user=xegzhJcAAAAJ&hl=en

[14] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet Classification with Deep Convolutional Neural Networks. Communications of the ACM, vol. 60, no. 6, pp. 84–90. https://en.wikipedia.org/wiki/ AlexNet

[15] Simonyan, K., & Zisserman, A. (2014). Very Deep Convolu- tional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556. https://arxiv.org/abs/1409.1556

[16] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. arXiv preprint arXiv:1512.03385. https://arxiv. org/abs/1512.03385

[17] Noman, M. A. A., et al. (2022). HYBRID-CNN: For Identification of Rohu Fish Disease. In Proc. Int. Conf. Comput. Commun. Netw. Technol. (ICC-CNT).

https://research.daffodilvarsity.edu.bd/public/index.php/ Scopus-article/2678

[18] Ahmed, M. S. (2024). SalmonScan: A Novel Image Dataset for Machine Learning and Deep Learning Analysis in Fish Disease Detection in Aquaculture. Mendeley Data, V3, doi: 10.17632/x3fz2nfm4w.3. https: //data.mendeley.com/datasets/x3fz2nfm4w/3

[19] Google Developers. (2025). Accuracy, Precision, Recall. Google Machine Learning Crash Course. https://developers.google.com/ machine-learning/crash-course/classification/accuracy-precision-recall

[20] DigitalOcean. (2025). Deep Learning Metrics: Precision, Recall, Accu- racy. DigitalOcean Community Tutorials. https://www.digitalocean.com/ community/tutorials/deep-learning-metrics-precision-recall-accuracy

[21] Ahmed, M. S., Aurpa, T. T., & Azad, M. A. K. (2021). Fish Disease Detection Using Image Based Machine Learning Technique in Aquaculture. Journal of King Saud University Computer and Information Sciences, vol. 34, no. 1. https://www.researchgate.net/ publication/351697654 Fish Disease Detection Using Image Based Machine Learning Technique in Aquaculture

[22] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In Adv. Neural Inf. Process. Syst., pp. 1097–1105. https://www.pinecone.io/learn/series/ image-search/imagenet/

[23] Iterate.ai. (2024). VGGNet: The Definition, Use Case, and Relevance for Enterprises in Visual Recognition. Iterate.ai. https://www.iterate.ai/ ai-glossary/vggnet-deep-convolutional-networks-feature-extraction

[24] Wikipedia. (2025). Residual neural network. Wikipedia. https://en. wikipedia.org/wiki/Residual neural network

[25] Wikipedia. (2025). Inception (deep learning architecture). Wikipedia. https://en.wikipedia.org/wiki/Inception (deep learning architecture)

[26] Medium. (2023). A Simple Guide to the Versions of the Inception Network. Medium. https://medium.com/@API4AI/ deep-learning-basics-for-image-processing-ac08fb60021e

[27] Ahmed, M. S. (2024). Enhanced Fish Health Monitoring in Aquaculture with Attention-Based Deep Learning Technique. ResearchGate. https://www.researchgate.net/publication/388739329 Enhanced Fish Health Monitoring in Aquaculture with Attention-Based Deep Learning Technique

[28] FlyPix.ai. (2024). Best Practices for Training Image Recognition Models. FlyPix.ai. https://flypix.ai/blog/ best-practices-for-training-image-recognition-models/

[29] A., A., et al. (2024). Advancing Crayfish Disease Detection: A Comparative Study of Deep Learning and Canonical Machine Learning Techniques. Applied Sciences, vol. 14, no. 14, p. 6211. https://www. mdpi.com/2076-3417/14/14/6211

[30] ICES. (2011). ICES Fish Disease Dataset. ICES Geonetwork. https://gis.ices.dk/geonetwork/srv/api/records/ 434c9221-bd9f-4c62-aee9-cd795e5a50ad

[31] Medium. (2023). Deep Learning Basics for Image Processing. Medium. https://medium.com/@API4AI/ deep-learning-basics-for-image-processing-ac08fb60021e

Table for Abbreviation

S. No.	Abbreviation	Description
1	AI	Artificial Intelligence
2	CBAM	Convolutional Block Attention Module
3	CNN	Convolutional Neural Network
4	DL	Deep Learning
5	EUS	Epizootic Ulcerative Syndrome
6	FN	False Negative
7	FP	False Positive
8	GPU	Graphics Processing Unit
9	ReLU	Rectified Linear Unit
10	SGD	Stochastic Gradient Descent
11	SVM	Support Vector Machine
12	TN	True Negative
13	TP	True Positive
14	TPR	True Positive Rate