



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: I Month of publication: January 2026

DOI: <https://doi.org/10.22214/ijraset.2026.76812>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

CNN and Transformer-Based Deep Learning Models for Coral Health Classification

Hajara Sabnam Kareem Navaz¹

¹Department of Information Technology, Amity University Dubai, Dubai, United Arab Emirates

Abstract: Coral reefs are among the most important marine ecosystems, providing habitat for nearly 32% of marine species and supporting the livelihoods and food security of more than 500 million people worldwide. In addition to sustaining marine biodiversity, coral reefs protect coastlines and contribute to economic and cultural activities in many regions. Despite their significance, coral reefs are rapidly declining due to climate change, ocean warming, and, in particular, coral bleaching. The most severe global coral bleaching event to date began in 2023 and is still ongoing, affecting approximately 84% of the world's coral reefs, highlighting the urgent need for reliable methods to detect bleaching at an early stage. In response to this challenge, this paper proposes deep learning-based models for coral health classification using underwater images. The approach compares the performance of three convolutional neural network architectures, ResNet50, EfficientNetB3, and ConvNeXt-Tiny, with a transformer-based model, Swin Transformer-Tiny, for classifying corals as healthy or bleached. Transfer learning is applied to all models, and their performance is evaluated using a publicly available dataset containing 923 labeled coral images. The results show that all models achieve effective classification performance, with ConvNeXt-Tiny and Swin Transformer-Tiny attaining the highest accuracy of 86.33%, outperforming ResNet50 and EfficientNetB3. These findings provide insight into the advantages of newer CNN and transformer-based architectures for learning complex visual patterns in underwater coral images. The results further demonstrate their suitability for practical coral reef monitoring systems, supporting reliable and early detection of bleaching in real-world conservation and reef management applications.

Keywords: Coral Health Classification, Convolutional Neural Networks (CNNs), Transformer Models, Transfer Learning, Deep Learning, Coral Bleaching Detection

I. INTRODUCTION

Coral reefs, often called “the rainforests of the sea,” are colonies of corals or small animals called polyps that resemble tiny sea anemones. These polyps secrete limestone, which builds up over time to form complex reef structures. These coral reefs, made up of thin layers of calcium carbonate in the seas, serve as refuges, food reserves, and nurseries for many habitats in the ocean, ranging from smallest algae to huge fishes and invertebrates [1, 2]. They are one of the most valuable marine organisms as they create habitats for about 32% of the entire marine species [3]. Globally, they support the livelihoods and food security of over 500 million people, particularly in developing countries and small island states. They also act as natural breakwaters, reducing coastal erosion and protecting shorelines from storms, with estimated flood protection benefits valued at over \$130 billion globally [4]. Not only are they essential for the survival of many marine species, but they also provide several other benefits, including serving as a source of novel pharmaceuticals and holding high cultural, spiritual, and educational value for many communities.

However, despite holding so much significance in different aspects, coral reefs are declining rapidly due to ocean acidification, climate change, habitat destruction, and most commonly due to coral bleaching and therefore are at the verge of endangerment. Coral bleaching, generally defined as the loss of the coral's symbiotic algae (Symbiodinium) or their pigments when the coral is under stress typically happens in three stages. At first, the coral's microbiota try to maintain balance, then they shift to support heat-tolerant symbionts, and finally, if the stress continues, these protective responses fail, which can lead to the coral's death [5]. This bleaching leads to substantial coral mortality, hence, compromising the structure and function of their ecosystems and acting as one of the major causes of coral reef endangerment [6]. Over the last two decades, coral bleaching has occurred in several distinct episodes, each differing in severity and duration. The frequency, scale, and intensity of mass bleaching events have increased dramatically since the 1980s, with global-scale events in 1998, 2010, and 2014–2017 causing unprecedented coral loss [7]. However, the most intense and severe coral bleaching event on record started from the beginning of 2023 and is still ongoing according to scientists and has affected about 84% of the world's coral reefs [8].

With an increasing trend in the bleaching of corals, it becomes necessary and important to develop certain methods for early identification of bleaching in corals, as it allows for rapid intervention in conservation efforts before the corals die, therefore increasing the chances of recovery and preventing permanent loss of reef structure and biodiversity.

Hence, recognizing this is a severe issue, this paper proposes computer vision-based deep learning models using CNNs and transformer-based approaches, respectively, ResNet50, EfficientNetB3, ConvNeXt-Tiny, and Swin Transformer, to detect and classify corals as healthy or bleached based on their images in order to facilitate timely intervention in coral reef conservation, improve monitoring accuracy and aid in coral reef management, so as to prevent permanent loss of corals due to bleaching. ResNet50 is employed as a traditional CNN architecture leveraging residual learning, EfficientNetB3 represents a modern CNN optimized for parameter efficiency through compound scaling, and ConvNeXt-Tiny reflects recent advancements in CNN design inspired by transformer principles. Additionally, the Swin Transformer-Tiny is included as a representative transformer-based architecture to assess the applicability of attention-based models for coral classification tasks under limited data conditions.

II. LITERATURE REVIEW

In recent years, a growing number of studies have explored and examined how machine learning and deep learning techniques can enhance coral reef monitoring, species classification, and overall ecosystem assessment.

Raphael et al. in their work have reviewed and discussed the developments in applying deep learning methods to coral reef research between 2016 and 2018, highlighting how these models improve coral identification, species classification, and monitoring accuracy compared to other manual methods. They also identified key limitations in current approaches and outlined future needs, such as fine-scale species detection and automated tracking of coral health and diversity [9]. Chowdhury et al. in their paper present a multidisciplinary review of how GIS, remote sensing, ML and DL techniques are used to monitor coral reef changed under climate stress. By evaluating publicly available datasets from organizations like NOAA, they evaluate existing ML/DL models and mention the requirement for better models to monitor their health [10]. Shahid et al. in their research work developed a Bag-of-Features-based method to detect and localize bleached corals, evaluating multiple handcrafted descriptors (such as LBP, HOG, and GLCM) alongside deep CNN models including AlexNet, VGG-19, ResNet-50, and Inception v3. Their approach achieved 99.08% accuracy, outperforming current state-of-the-art techniques in both classification and localization [11]. Ahmed et al. in their paper introduced a hybrid CNN–Vision Transformer model for automated coral reef health assessment using 923 labeled underwater images. By combining local feature extraction from CNNs with global context modeling from ViTs, their approach achieved 83.78% accuracy and demonstrated potential for scalable, real-time coral monitoring [12]. Kaur et al. proposed a transfer-learning approach using the VGG19 convolutional neural network to automate coral reef classification on a dataset of 923 labeled images. Their model, trained and evaluated through standard metrics such as accuracy, confusion matrix, and classification report, achieved 74% accuracy. The study highlights the benefits of CNNs in enabling efficient, scalable, and real-time monitoring of coral health, offering a valuable tool for conservation efforts [13]. Thamarai et al. applied CNN-based deep learning methods to classify coral reefs into healthy and stressed categories using ResNet50, InceptionV3, and a custom CNN model. After hyperparameter tuning, the pretrained models reached 70% and 55% accuracy, while their proposed model achieved a higher accuracy [14]. Xin et al. introduced FCOS_EfficientNET, an improved EfficientNet-based object detection model for coral bleaching, incorporating BiFPN and optimized attention mechanisms. On a custom coral dataset, FCOS_EfficientNETb3 achieved 81.5% accuracy and 59.3% recall, while other variants balanced speed and precision for real-time monitoring, demonstrating the model's effectiveness for reef conservation and mobile/edge deployment [15].

The authors mostly use CNN for coral classification or traditional ML algorithms like SVM and Random Forest for classifying coral bleaching status, or use a hybrid approach by combining multiple ML techniques such as CNNs with colour normalization and texture analysis for higher accuracy. These studies clearly demonstrate the potential of machine learning and deep learning techniques in developing models for coral reef monitoring and coral bleaching detection. While hybrid models and transfer learning approaches also demonstrate good performance and scalability, the challenges regarding generalization, and interpretation still remain highlighting the need for better models with higher accuracy especially in diverse environmental conditions.

III. MATERIALS AND METHODS

This paper presents a comparative analysis of deep convolutional neural networks, ResNet50, EfficientNetB3, ConvNeXt-Tiny and a transformer-based model, Swin Transformer-Tiny, in effectively classifying corals as healthy or bleached based on image features extracted from underwater photographs. By using transfer learning to fine-tune pre-trained models, this paper additionally focuses on improved classification accuracy and generalizability.

The methodology followed to develop these models follows a systematic approach from dataset acquisition and preprocessing, through model selection and fine-tuning, to performance evaluation using standard metrics and is depicted in Fig.1.

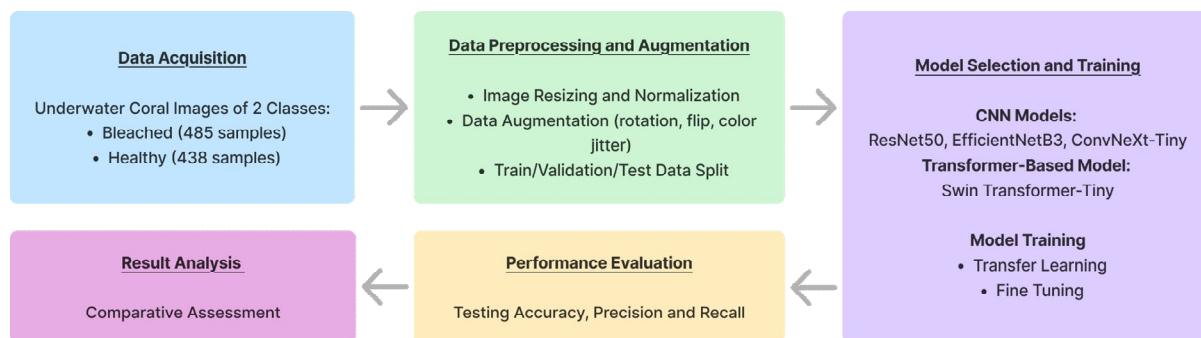


Fig. 1 Methodology Workflow for Coral Health Classification

A. Dataset Description and Preprocessing

The dataset, obtained from publicly available resources originally consisted of underwater coral images categorized into two classes, healthy corals and bleached corals. It consisted of a total of 923 images, including 485 images of bleached corals and 438 images of healthy corals. Table 1 shows a sample image from both the classes.

TABLE I
SAMPLE IMAGES FROM THE DATASET CLASSES

Class Label	Sample Image
Healthy coral	
Bleached coral	

The dataset was split into 3 subsets for training, validation and testing in a ratio of 70:15:15 respectively. The dataset images underwent a preprocessing and augmentation stage to enhance model robustness before model training. For data preprocessing, two separate transformation pipelines (train_transform, and test_transform objects) were defined for training and testing phases of the models. To improve model generalization and to reduce overfitting, extensive data augmentation was applied to the training images. Each of the training images was first resized to 256×256 pixels. Then, to introduce a spatial variation and to enable the models to learn important features anywhere and not just fixed position, a 224×224 region was randomly cropped from the 256×256 image. After this, random horizontal flip with a probability of 0.5 and random vertical flip with a probability of 0.3 was applied in order to simulate different viewing orientations commonly encountered in underwater images. Additionally, images were randomly rotated within a range of ± 20 degrees to account for camera angle variations during image acquisition. Then, colour-based augmentations were applied using colour jittering to reflect the diverse lighting conditions of underwater images. Following this, random affine transformations with translation up to 10% of the image dimensions were used to introduce minor spatial shifts. After these augmentations, the images were converted into tensors and normalized using ImageNet mean and standard deviation values (mean = [0.485, 0.456, 0.406], standard deviation = [0.229, 0.224, 0.225]) to ensure compatibility with the pre-trained deep learning models.

B. Model Development

After data preprocessing, deep learning models were developed using ResNet50, EfficientNetB3, ConvNeXt-Tiny and Swin Transformer-Tiny architectures and their performances were evaluated and compared. These four architectures were selected to represent both conventional CNN-based models and modern transformer-based models, enabling a comprehensive comparison between residual learning, compound scaling, and attention-based mechanisms for coral health classification. The first model developed was a ResNet50 CNN model. It is a deep convolutional neural network (CNN) architecture widely used for image classification and related computer vision tasks. It is a part of Residual Network family, which basically introduced the concept of residual learning to enable the training of very deep networks without the vanishing gradient problem [16]. The vanishing gradient problem occurs when gradients become progressively smaller during backpropagation in deep neural networks, therefore preventing early layers from learning effectively. To overcome this problem, residual networks introduce the use of residual connections or skip connections, which is a shortcut that basically bypasses one or more layers, allowing the network to learn residual functions, i.e., the difference between input and output rather than direct mappings [17]. ResNet50 is a widely adopted residual network, was used as a strong baseline model in this paper due to its stable optimization and performance in visual image recognition tasks. The second model was developed using EfficientNetB3, a deep CNN model from the EfficientNet family, which was designed for high accuracy, compound scaling and computational efficiency in image classification tasks [18]. Instead of making the model deeper, this architecture focuses on scaling the depth, width and resolution in a balanced way. B3 is a mid-sized model of the versions B0-B7. The balanced and optimized scaling strategy in these models gives better accuracy per parameter [19]. EfficientNetB3 is a much more recent and parameter-efficient architecture that jointly scales network depth, width and resolution. The third model developed was a ConvNeXt-Tiny model. It is a modernized convolutional network that adopts transformer-inspired design choices while preserving the inductive biases of CNNs, offering improved feature representation for visual patterns. The fourth model was developed using Swin Transformer-Tiny architecture. In contrast to other three models, the Swin Transformer-Tiny is an attention-based model that captures long-range dependencies through hierarchical self-attention, making it particularly effective for learning global contextual information. All four models were developed using the same training methodology and evaluation pipeline. Fig. 2 describes the development setup of the models.

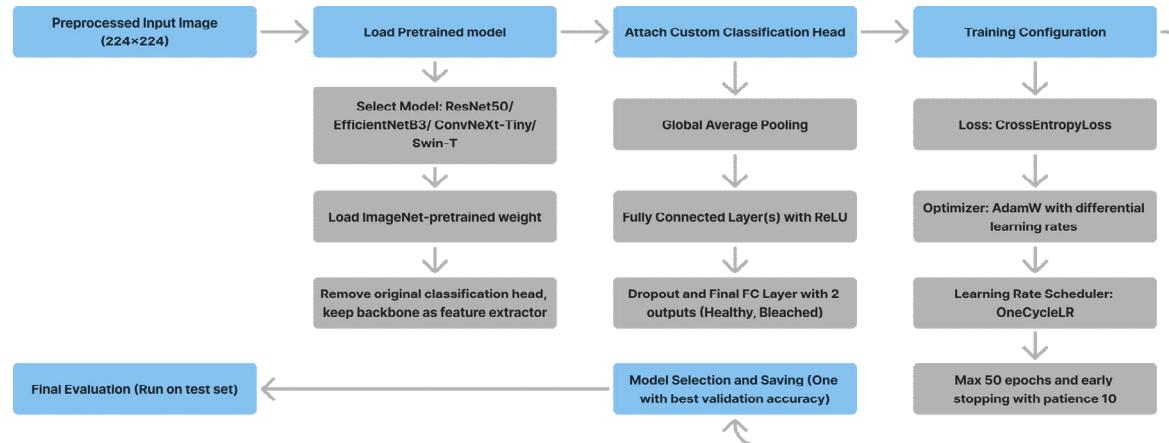


Fig. 2 Model Training and Development Pipeline

For each architecture, the pre-trained model was initialized with weights learned from the ImageNet dataset. The original ImageNet-specific classification head was removed, and the remaining layers were retained as a feature extractor to learn high-level visual representations from underwater coral images. A custom classification head was added to each model, consisting of fully connected layers with ReLU activation and dropout to reduce overfitting. The final layer produced two output scores corresponding to healthy and bleached coral classes. The models were trained for a maximum of 50 epochs with early stopping applied using a patience of 10 epochs. CrossEntropyLoss was used as the loss function. The AdamW optimizer was employed with differential learning rates, where a lower learning rate was assigned to the pre-trained backbone and a higher learning rate to the newly added classification layers to enable stable fine-tuning. A OneCycle learning rate scheduler was used to dynamically adjust the learning rate during training, improving convergence and generalization. Model performance was monitored using training and validation loss, accuracy, and precision. The model achieving the highest validation accuracy was saved as the best-performing model. Final evaluation was conducted on the test set using model accuracy, precision, recall, F1-score and confusion matrix analysis.

IV. RESULTS AND DISCUSSION

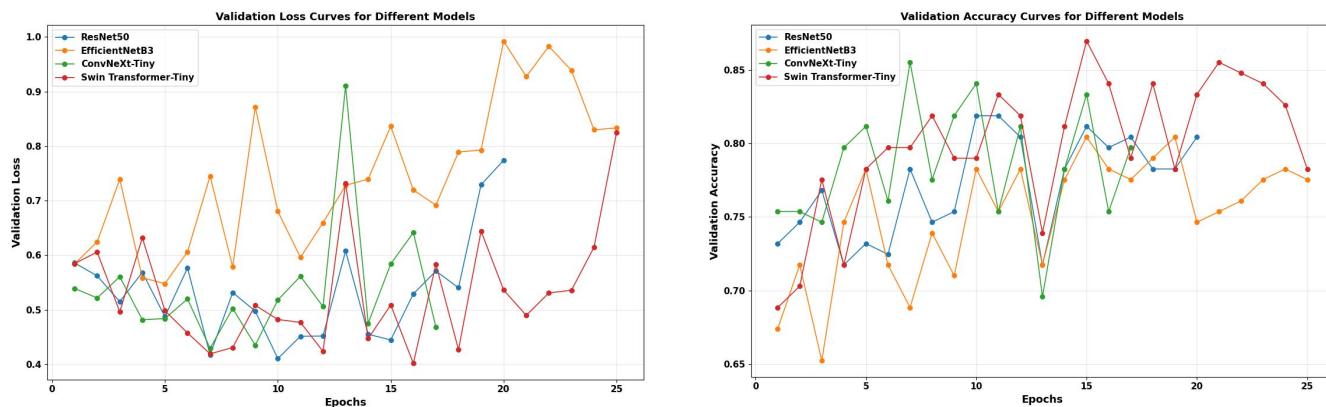
In this paper, three CNN models, ResNet50, EfficientNetB3, ConvNeXt-Tiny and a Transformer-based model, Swin Transformer-Tiny were compared as backbone models for coral health classification. Table 2 displays a comparison of the performance of both the models based on certain evaluation metrics.

TABLE II
PERFORMANCE COMPARISON OF THE DEVELOPED MODELS

Accuracy Metrics	ResNet50		EfficientNetB3		ConvNeXt-Tiny		Swin Transformer-Tiny	
	Bleached Corals	Healthy Corals	Bleached Corals	Healthy Corals	Bleached Corals	Healthy Corals	Bleached Corals	Healthy Corals
Precision	0.82	0.87	0.86	0.81	0.86	0.87	0.88	0.85
Recall	0.89	0.79	0.82	0.85	0.89	0.83	0.86	0.86
F1-score	0.86	0.83	0.84	0.83	0.87	0.85	0.87	0.86
Test Accuracy	84.17%		83.45%		86.33%		86.33%	

All the developed models demonstrated a good performance in distinguishing healthy and bleached corals, therefore indicating the effectiveness of deep convolutional neural networks for this task. Fig. 3 (a) displays the validation loss curves and (b) displays the validation accuracy curves of the developed models. The confusion matrices of all the models is presented in Fig. 4.

ResNet50 model stopped training at 20 epochs due to early stopping and gave an overall test accuracy of 84.17%. Out of 73 bleached corals, the model correctly classified 65 as bleached and incorrectly classified 8 as healthy. Out of 66 healthy corals, it classified 52 correctly and 14 incorrectly. As for the EfficientNetB3 model, early stopping triggered at 25 epochs and the model gave an overall test accuracy of 83.45%. The model correctly classified 60 bleached coral images and misclassified 13, while correctly classifying 56 healthy coral images and misclassifying 10. The ConvNeXt-Tiny and Swin Transformer-Tiny models both gave the highest test accuracy of 86.33%. ConvNeXt-Tiny model stopped training at 17 epochs due to early stopping. It also correctly also classified 65 bleached corals and misclassified 8, while correctly classifying 55 healthy corals and misclassifying 11. The Swin Transformer-Tiny model stopped training at 25 epochs. It further correctly classified 63 bleached corals and misclassified 10, while correctly classifying 57 healthy corals and misclassifying 9.



a) Validation Loss Curves

b) Validation Accuracy Curves

Fig. 3 Validation Loss and Accuracy Curves of the Developed Models

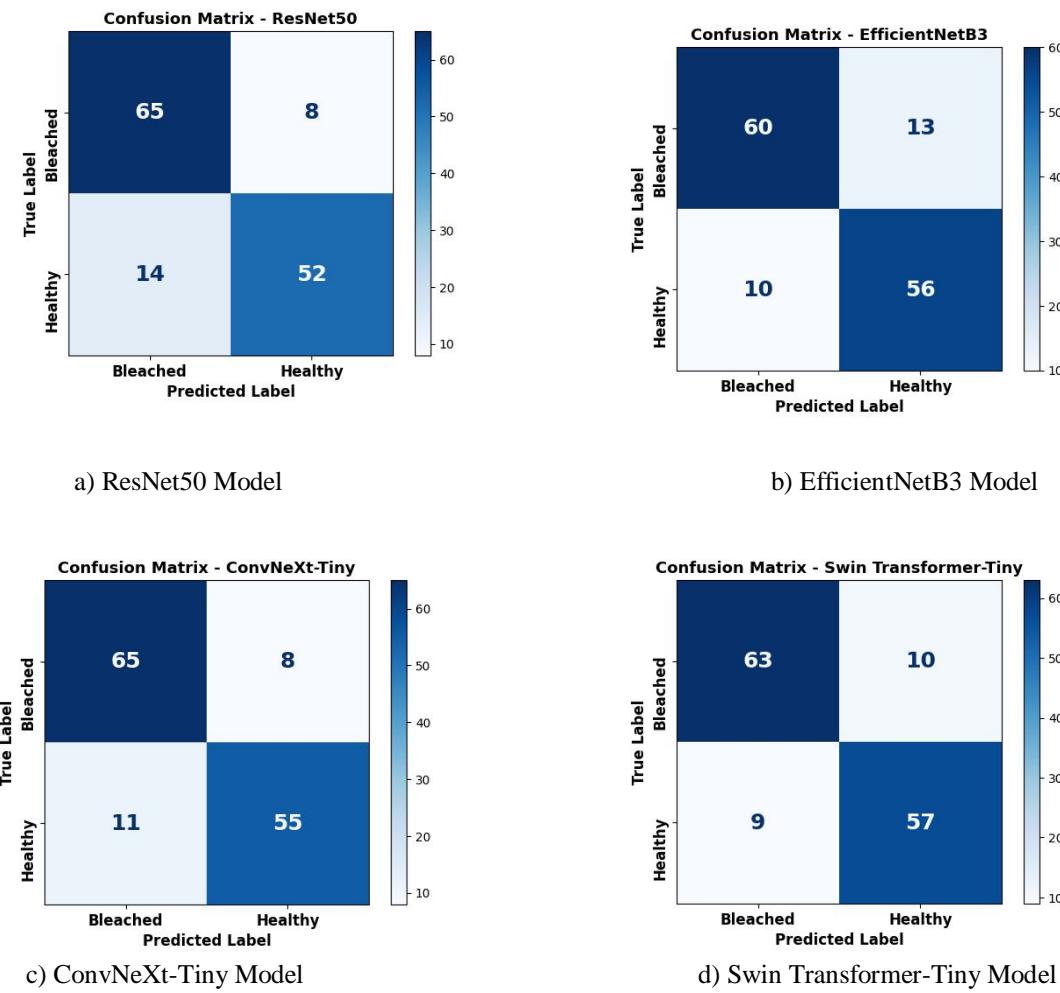


Fig. 4 Confusion Matrices of the Developed Models

Overall, ConvNeXt-Tiny and Swin Transformer-Tiny achieved the highest test accuracy, suggesting that recent CNN and transformer-based architectures can improve classification performance on coral images compared to ResNet50 and EfficientNetB3. The slightly lower performance of ResNet50 can be attributed to its architecture, which primarily increases depth through residual blocks, making it less efficient on relatively small datasets. EfficientNetB3, despite its parameter efficiency and compound scaling of depth, width, and resolution, achieved slightly lower accuracy than the recent architectures but still demonstrated balanced feature extraction and generalization capabilities. ConvNeXt-Tiny (CNN) and Swin Transformer-Tiny (Transformer-based), being more recent architectures, effectively captured richer spatial and contextual features, improving classification performance on underwater coral images.

V. CONCLUSION AND FUTURE WORK

This paper presented a comparative analysis of CNN and Transformer-based models, ResNet50, EfficientNetB3, ConvNeXt-Tiny and Swin Transformer-Tiny for effective coral health classification. Among the models, ConvNeXt and Swin Transformer achieved the highest test accuracy of 86.33%, outperforming ResNet50 (84.17%) and EfficientNetB3 (83.45%). By comparing these four models, we gain validated insights into their capabilities in coral health classification. Specifically, we observe that while all models are effective at distinguishing healthy and bleached corals, the more recent architectures, ConvNeXt and Swin Transformer, better capture complex spatial and contextual features, leading to higher classification performance. ResNet50 and EfficientNetB3, though effective, are comparatively less efficient on this dataset, highlighting the significance of architecture choice for tasks involving limited and complex image data. Overall, the results demonstrate the potential of deep learning models for automated coral reef monitoring, even when trained on relatively small datasets.

The findings of this study suggest that Transformer-based and modern CNN architectures can be effectively used for real-world coral health assessment and early bleaching detection. As future work, the dataset can be expanded to include more diverse coral species and environmental conditions to further improve model robustness and generalization. Additionally, integrating temporal data, attention-based ensemble approaches, or deploying the model in real-time underwater monitoring systems could further enhance the practical applicability of this work in coral reef conservation efforts.

VI. ACKNOWLEDGMENT

I sincerely thank my mentors, family, and friends for their guidance and support throughout this research. I extend my gratitude to the researchers whose work provided valuable insights that helped shape and develop this research.

REFERENCES

- [1] National Ocean Service, "What is a coral reef made of?," *Noaa.gov*, 2019. <https://oceanservice.noaa.gov/facts/coralmadeof.html>
- [2] "The coral - Monaco Oceanographic Institute, Albert I Foundation," Institut océanographique, Jul. 19, 2021. <https://www.oceano.org/en/thematic-pages/the-coral/> (accessed Jan. 02, 2026).
- [3] Y. K. Wong, Z. Zheng, M. Zhang, D. J. Suggett, and S.-K. Yeung, "CoralSCOP-LAT: Labeling and analyzing tool for coral reef images with dense semantic mask," *Ecological Informatics*, vol. 91, p. 103402, Nov. 2025, doi: <https://doi.org/10.1016/j.ecoinf.2025.103402>.
- [4] D. O. Obura et al., "Coral Reef Monitoring, Reef Assessment Technologies, and Ecosystem-Based Management," *Frontiers in Marine Science*, vol. 6, no. 580, Sep. 2019, doi: <https://doi.org/10.3389/fmars.2019.00580>.
- [5] D. N. Handiani, N. S. Ningsih, and E. Belyiana, "Coral bleaching occurrence and its relation to marine heatwave events in the Southwestern waters of South Sulawesi, Indonesia, as part of the Coral Triangle region," *Journal of Marine Systems*, vol. 252, p. 104136, Sep. 2025, doi: <https://doi.org/10.1016/j.jmarsys.2025.104136>.
- [6] T. N. F. Roach, J. Dilworth, C. M. H. A. D. Jones, R. A. Quinn, and C. Drury, "Metabolomic signatures of coral bleaching history," *Nature Ecology & Evolution*, pp. 1–9, Feb. 2021, doi: <https://doi.org/10.1038/s41559-020-01388-7>.
- [7] C. M. Eakin, H. P. A. Sweatman, and R. E. Brainard, "The 2014–2017 global-scale Coral Bleaching event: Insights and Impacts," *Coral Reefs*, vol. 38, no. 4, pp. 539–545, Jul. 2019, doi: <https://doi.org/10.1007/s00338-019-01844-2>.
- [8] Margaux Monfared, "84% of the world's coral reefs impacted in the most intense global coral bleaching event ever | ICRI," ICRI, Apr. 23, 2025. <https://icriforum.org/4gbe-2025/>
- [9] A. Raphael, Z. Dubinsky, D. Iluz, and N. S. N., "Neural Network Recognition of Marine Benthos and Corals," *Diversity*, vol. 12, no. 1, p. 29, Jan. 2020, doi: <https://doi.org/10.3390/d12010029>.
- [10] Chowdhury, M. Jahan, Shahriar Kaisar, M. E. Khoda, A. Karim, and R. Naha, "Coral Reef Surveillance with Machine Learning: A Review of Datasets, Techniques, and Challenges," *Electronics*, vol. 13, no. 24, pp. 5027–5027, Dec. 2024, doi: <https://doi.org/10.3390/electronics13245027>.
- [11] S. Jamil, M. Rahman, and A. Haider, "Bag of Features (BoF) Based Deep Learning Framework for Bleached Corals Detection," *Big Data and Cognitive Computing*, vol. 5, no. 4, p. 53, Oct. 2021, doi: <https://doi.org/10.3390/bdcc5040053>.
- [12] Ajay, Akanksh M, and Mamatha Balipa, "Enhancing Coral Health Monitoring with a Hybrid CNN-ViT Model for Bleaching Prediction," pp. 74–79, May 2025, doi: <https://doi.org/10.1109/icaiss61471.2025.11041783>.
- [13] A. Kaur, K. S. Gill, M. Kumar, and R. Rawat, "VGG19's Role in Safeguarding Coral Reefs using a CNN-Based Monitoring Model," 2024 4th Asian Conference on Innovation in Technology (ASIANCON), pp. 1–5, Aug. 2024, doi: <https://doi.org/10.1109/asialcon62057.2024.10838114>.
- [14] M. Thamarai and S. P. Aruna, "Stressed Coral Reef Identification Using Deep Learning CNN Techniques," *Journal of Electronic & Information Systems*, vol. 5, no. 2, pp. 1–9, Sep. 2023, doi: <https://doi.org/10.30564/jeis.v5i2.5808>.
- [15] G. Xin, H. Xie, S. Kang, Y. Chen, and Y. Jiang, "Improved research on coral bleaching detection model based on FCOS model," *Marine Environmental Research*, vol. 200, p. 106644, Sep. 2024, doi: <https://doi.org/10.1016/j.marenvres.2024.106644>.
- [16] Y. Shuang et al., "Classification of pain expression images in elderly with hip fractures based on improved ResNet50 network," *Frontiers in Medicine*, vol. 11, Jul. 2024, doi: <https://doi.org/10.3389/fmed.2024.1421800>.
- [17] L. Mahin, A. Shaikh, Aniket Sandupatla, Rushikesh Pudale, A. Bakare, and Prof. Mallesh Chavan, "Classification of Simple CNN Model and ResNet50," *International journal for research in applied science and engineering technology*, vol. 12, no. 4, pp. 4606–4610, Apr. 2024, doi: <https://doi.org/10.22214/ijraset.2024.60677>.
- [18] Deep Ajabani, Z. A. Shaikh, A. Yousef, K. Ali, and M. A. Albahar, "Enhancing skin lesion classification: a CNN approach with human baseline comparison," *PeerJ Computer Science*, vol. 11, pp. e2795–e2795, Apr. 2025, doi: <https://doi.org/10.7717/peerj-cs.2795>.
- [19] G V Kartheek and S. Rani, "Robust Multi-Class Classification for Real-Time Agricultural Applications Using Efficient and Adaptive Deep Learning," *International Journal of Scientific Research in Science and Technology*, vol. 11, no. 6, pp. 90–99, Nov. 2024, doi: <https://doi.org/10.32628/ijsrst2411490>.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)