International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

BSI-EUNet: Birds Species Identification using Enhanced UNET Architecture

[1]K. Annalakshmi, [2]Dr. R. Rajeswari

Submitted: 02/10/2023 **Revised**: 21/11/2023 **Accepted**: 01/12/2023

Abstract: The precise and automated identification of bird species is becoming increasingly important as biodiversity assessment and conservation activities become more prominent. This paper introduces a new strategy, BSI-EUNet, for the identification of bird species in image processing and classification, in response to the increasing demand for more effective and efficient methods. The dataset containing bird species has been obtained from the Kaggle repository. The image dataset has undergone denoising using non-adaptive thresholding to improve the quality of the input images. The segmentation was performed utilizing the upgraded UNET architecture, often known as E-UNET. This form of UNET is specifically designed to enhance the accuracy and precision of segmenting bird species in photographs, resulting in a meticulous and precise depiction of the borders of each species. The training and classification process involves utilising a Hybrid Neural Network (HNN) that combines the Visual Geometry Group (VGG)-16 and Convolutional Neural Network (CNN) with a modified Densenet. This approach aims to develop a strong and adaptable model for identifying different species. The integration of these structures improves the network's capacity to detect complex patterns and hierarchical characteristics, facilitating precise categorization of various bird species. The model's higher performance is demonstrated by experimental assessments on the datasets, surpassing existing approaches.

***Keywords**: Bird species, CNN, Deep Learning, Hybrid Neural Network, E-UNET architecture*

1. Introduction

Various groups of enthusiasts and experts are interested in bird species identification for various reasons, including the aesthetic appeal of birds and their songs and their ecological value [1]. Bird identification is a task that ornithologists are acquainted with, and it has been recognized as a scientific endeavour since antiquity [2]. Ornithologists study various elements of bird life, such as how birds interact with their environment, their anatomy, the songs they produce, and their distribution [3]. Observing, researching, and monitoring birds have practical purposes [4-8].

VGG-16, known for its robust feature extraction capabilities, collaborates with a CNN-modified Densenet, forming a synergistic architecture. VGG-16 captures high-level features, while the modified Densenet enhances hierarchical feature learning, collectively enabling precise and efficient bird species identification [9-12].

This study introduces BSI-EUNet, a comprehensive framework designed for Birds Species Identification. Central to its efficacy are three key components: non-adaptive thresholding for image denoising, segmentation using an enhanced UNET [15] architecture, and a HNN for training and classification, fusing the strengths of VGG-16 [17] and a CNN [17]-modified Densenet [18].

The primary contributions and objectives of this manuscript may be summarized as follows.

- Non-adaptive thresholding [19] for Image denoising
- Bird Species Segmentation has been done with enhanced UNET architecture
- Training and Bird Species Classification using a HNN, VGG-16, with CNN-modified Densenet

This paper's remaining structure as follows. In Section 2, different authors address different types of bird's species classification. In Section 3, presents the BSI-EUNet model. Section 4 discussed with experiments results and discussion. Finally the conclusion has represented at Section 5.

2. Background Study

Islam, S. et al. [2] in this work, machine learning techniques was utilized to classify images of birds from Bangladesh into their respective species. These authors have achieved an accuracy of 89% using SVM and kernel approach. Jian, L. et al. [4] this study developed a web service for identifying bird species based on maximum identification similarity. They used a single feature value, which made it difficult to distinguish between certain bird species that were relatively similar, requiring further comparison of morphological attributes. Marini, A. et al. [7] this study used aural and visual

[1]Research Scholar, Dept. of Computer Applications, Dr M.G.R. Educational and Research Institute, Chennai, Tamil Nādu, India.
annalakshmi.krishnan@gmail.com
[2]Associate Professor, Dept. of Computer Applications, Dr M.G.R. Educational and Research Institute, Chennai, Tamil Nādu, India.
rajeswari.cse@drmgrdu.ac.in

features to improve bird species identification based on bird images. They found that combining auditory and visual cues provided a correct classification rate of 35%, rejecting 30% of the samples. Nanni, L. et al. [8] this study developed a new method for classifying bird calls by utilizing visual and auditory characteristics of sound and texture descriptors derived from a spectrogram of the sound. They analyzed and compared these descriptors and their combinations on a large database of bird songs from 46 different species. Pang, C. et al. [9] this study presented a new fine-grained bird image retrieval task that looks for comparable samples from various species of birds. Authors suggested a two-step approach to tackle this difficulty, which was empirically demonstrated to be successful using the CUB200 dataset. In this study, Qiao, B. et al. [10] employed a Support Vector Machine (SVM) decision tree to propose a straightforward technique for bird classification. These researchers discovered that the accuracy rate fluctuated based on the specific beak characteristic. Qiu, Z. et al. [11] this study developed a multi-class SVM model to identify five common bird species that may disrupt transmission lines. These authors used image processing techniques such as Segmentation, greying, and filtering, along with colour, texture, and form characteristics to distinguish between bird species. Ragib, K. et al. [12] this study focused on identifying different species of birds using deep learning (DL) techniques. They created a dataset of birds based on their surroundings and compared two different models, with their best model achieving a 97.98% accuracy rate.

2.1 Problem Definition

The problem being addressed in the various studies mentioned is bird species identification. The goal is to accurately classify and identify bird species based on visual features. This is an important problem in fields such as ornithology, conservation biology, and ecology, where accurate identification of bird species is necessary for research and conservation efforts. This research overarching goal is to create model and method for machine learning that can reliably distinguish between different bird's species of using datasets.

3. Materials and Methods

The bird species classification has an important one for natural life. The BSI-EUNet model has utilized denoising, segmentation and classification with different algorithms. The dataset has been collected from benchmark datasets and trained using hybrid neural networks.

3.1 Dataset

The dataset has been collected from https://www.kaggle.com/datasets/gpiosenka/100-bird-species. The dataset contains train, test and valid folders. There are 400 different classes of bird species, and the dataset has an image category.

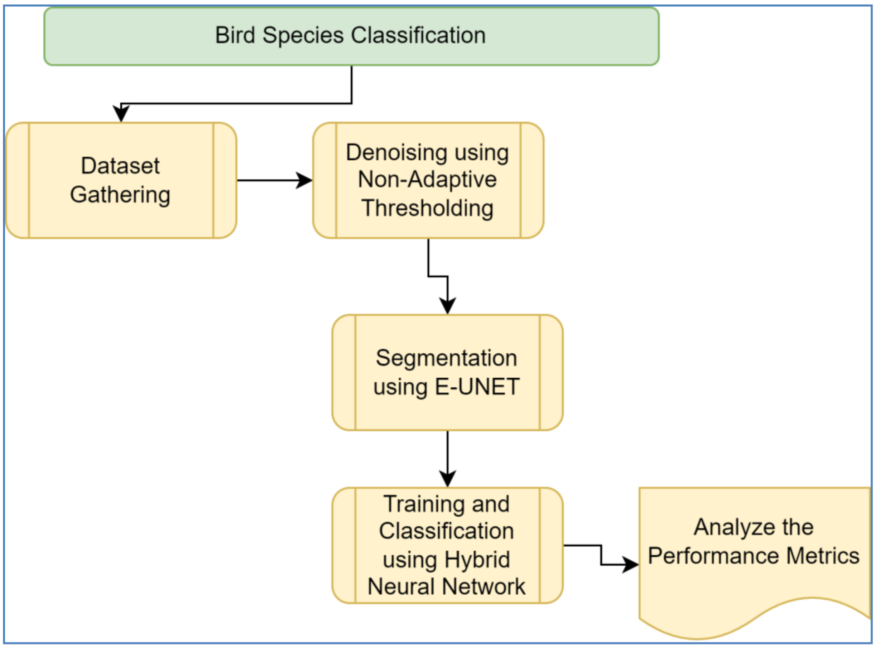

Fig 1: BSI-EUnet Flow diagram

3.2 Image denoising using Non-Adaptive Threshold

By putting a definite upper limit on the number of tests and an asymptotic limitation on the decoding time for Bird image denoising, this research improves upon previous work by Bui, T. V. et al. (2021) by allowing for detecting broken objects in a noisy environment.

The number of tests for encoding and the time for decoding in a gap may be decreased by constructing $(x, e - i, r; y)$ -disjoint matrix. We get the following theorem when applying Algorithm 1 to the (n, d, u, z]-disjoint matrix presented in Bui, T. V. et al. (2021) work. Let there be integers where $(e-r)\binom{x-r}{d+1}$. We'll call the flawed set S and the positive integer z for simplicity. There is a non-adaptive method that, given $(x, e-i, r; y]$ -model with at most $(x, e-i, r; y]$ erroneous outcomes can effectively identify some sets with only $(x, e-i, r; y]$ tests, where the decoding difficulty is

$$Ds = \left(g(x, e-i, r; y] \times r \left(\binom{x}{r} + (e-r)\binom{x-r}{d+1}\binom{e-1}{g}\binom{e}{r} \right) \right) \text{------ (1)}$$

Algorithm 1: Image Denoising using Non-Adaptive Threshold

Input:

- x, e, i, r, y (Parameters for the (x, e-i, r, y]-model)
- Bird Noisy image data with damaged objects

Output:

- Identified sets of damaged objects

Steps

- Set decoding difficulty DsDs using the formula in equation (3).
- Apply Non-Adaptive Thresholding to the noisy image data to enhance the signal-to-noise ratio.
- This involves setting a fixed threshold to distinguish between noise and actual image features.
- Construct a (x, e-i, r, y]-disjunct matrix based on the thresholded image data.
- This matrix aids in reducing the number of tests for encoding and decoding.
- Utilize the non-adaptive method to effectively identify sets of damaged objects with only (x, e-i, r, y] tests.

3.3 Segmentation using Enhanced UNet Architecture

After Denoising, the bird image has segmented by using Enhanced UNET architecture. The UNET is U-shaped encoder-decoder structure has been referred from Vinisha, A. ., & Boda, R. (2023). Figure 2 shows the encoder, which employs max-pooling (red arrows) and double convolution (blue arrows) to cut the image size in half while increasing the number of feature mappings by a factor of two. A decoder utilizes a double convolution following a bilinear up sampling operation (green arrows) to increase the feature map to minimize the number of feature maps by a factor of two. The feature maps are then concatenated with the preceding encoder's output using skip connections (grey arrows). Because of the skip connections, the model's output may be generated for a variety of input granularities. Finally, our model employs a single convolution (purple arrow) to build a single feature map reflecting the expected value of the network.

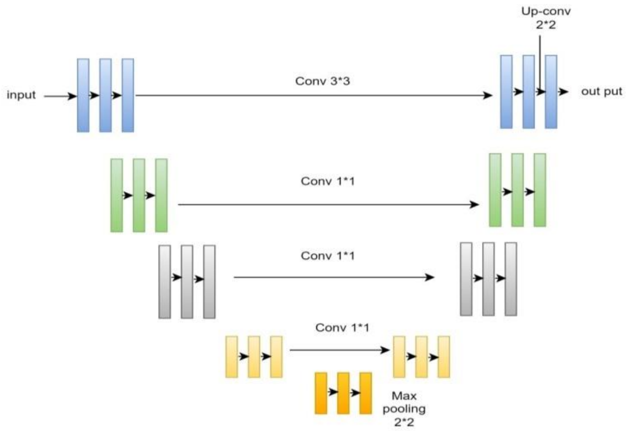

Fig 2: E-UNET architecture

The E-UNET architecture is shown in Figure 2, with the input picture size set at 572x572. However, 128 x 128 x 3 pixels is the standard many studies used. This necessitates thoroughly examining the photographs' various storage sites by professionals for optimum outcomes.

The input may include varied sizes, making it essential for specific uses to use several scales. Since UNets can be trained to predict a class for each pixel, they are often employed for classification and segmentation applications. The precise value of each pixel was anticipated, though, so we utilized it for a time series prediction challenge. Our unique Small Attention-UNet (E-UNet) model improves upon the classic UNet layout in two significant ways.

Algorithm 2: Segmentation using Enhanced UNet Architecture

Input:

- Denoised bird image
- Parameters of the Enhanced UNet architecture (e.g., input size, layer configurations)

Output:

Segmented bird image $Image_{Segmented}$

Steps

- Load the denoised bird image
- Configure the parameters for the Enhanced UNet architecture, such as input size and layer configurations.
- Implement the U-shaped encoder-decoder structure of the Enhanced UNet, as depicted in Figure 2.
- Encoder: H_i=Conv2D(MaxPooling(ReLU(Conv2D(Hi−1))))
- Decoder: H_i=Conv2D(BilinearUpsampling(ReLU(DoubleConvolution(Hi−1))))

Final Output: $Image_{Segmented}$=Conv2D(Feature Concatenation(Encoder Output, Decoder Output))

- Here, H_i represents the feature map at layer *i*, and DoubleConvolution refers to applying two convolutional layers sequentially.
- Examine the E-UNET design, as shown in Figure 2.
- Input Size: 572×572
- Standard Input Size: 128×128×3
- Consider potential variations in input sizes and the need for multiple scales for specific use cases.
- Incorporate attention mechanisms into the Enhanced UNet architecture.
- Attention Mechanism: H_i=Attention(Encoder Output, Decoder Input)
- Specify how attention mechanisms enhance the model's ability to focus on relevant regions.
- Obtain the segmented bird image as the output of the Enhanced UNet architecture.
- $Image_{Segmented}$=Conv2D(Feature Concatenation(Encoder Output, Decoder Output))

3.4 Training and testing Hybrid neural network (HNN) model

The artificial neural network(ANN's) data-driven model is accurate for timely predictions. Nodes in the input layer (data given to the network) are linked with a defined activation function to nodes in the hidden layer, and so on, until the nodes in the output layer are reached. The HNN model keeps the standard ANN's three-layer architecture while adding special activation functions. The following is the formulation of a conceptual function from the input layer to the hidden layer that uses the fuzzy pattern-recognition notion.

$$q_i = \frac{1}{\sum_{j=1}^{k}\left[W_{ij}\left(q_j^{im}-m_i\right)\right]^2} \text{ ---- (2)}$$

3.4.1 Visual Geometric Group-16 Layers

This research provides a novel model that takes advantage of transfer learning using VGG-16 networks referred by Dhulipalla, R. K. ., & Sethuraman, S. C. (2023) . Specifically, VGG-16 intends to spread 16- and 19-layer CNN models. When compared to current market leaders, the VGG-16 is just slightly behind. Even so, they are somewhatpractical and helpful for picture classification and as a basis for future models that use images as input. Seeing how TensorFlow is a background process in VGG-16, we use this library to make bird identifications. We used a 16-layer VGG network. During training using VGG-16, the input images are scaled up to a resolution of 224X224.

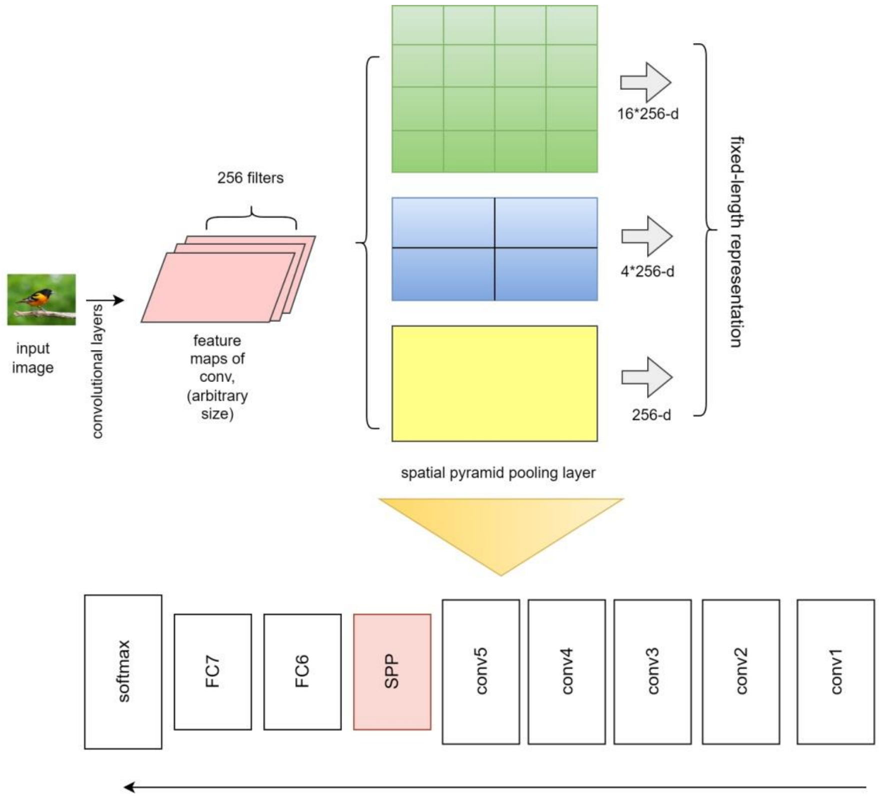

Fig 3: VGG-16 Model

For classifying data and the Softmax layer for optimizing the results. When all layers are discarded, VGG-16 provides feature representations in 4096 dimensions is represented at figure 3.

3.2.3 CNN

Ullah, F., et al. refer to convolutional neural networks (CNNs) as a type of DL system that can take an image as input and differentiate between comparable images by assigning weights to different features of the image (2024). CNN needs far less preprocessing than other classification methods. In earlier methods, filters were frequently manually constructed, but CNN can teach itself to do so after being exposed to enough training data.

Conceptually, CNNs resemble the interconnection of neurons in the human brain, in which individual neurons exclusively react to stimuli that are contained within their designated receptive field. These receptors collaborate to comprehend the entirety of human vision. The initial parameters of convolutional neural networks have a significant impact on their efficacy.

• Input Image

• CNN

• Output Label (Image Class)

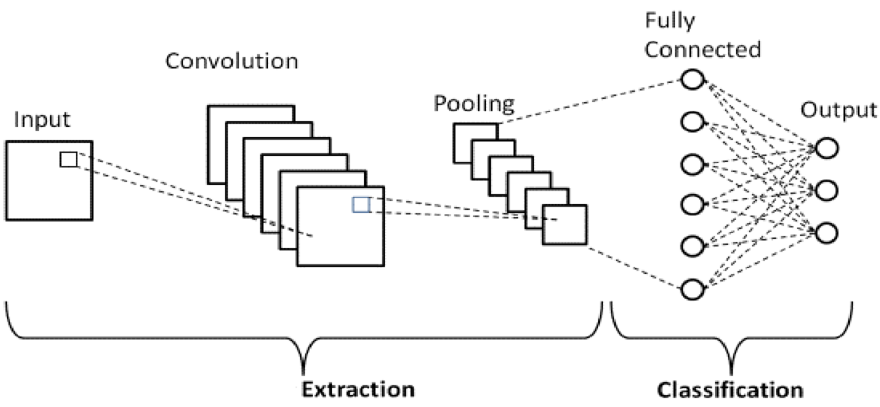

Fig 4: CNN Layers

Convolutional Neural Networks are created using the following stages:

• Convolution followed by the application of the RectifierFunction

• Pooling

• Flattening

• Fully Connection layers

3.2.3 Modified DenseNet algorithm

Enhancing the classification accuracy of the DenseNet network model on the power equipment dataset can be achieved by capitalising on the capabilities of the residual attention mechanism, which serves to emphasise significant attributes while diminishing extraneous ones. The dense architecture has referred by Zhang, K. et al. (2019) Incorporating the residual attention mechanism, DenseNet is a network model. The input, shown by x in the above diagram, is down-sampled by the mask branch on the left to simplify the extraction of more useful information. Then the output and input are up-sampled to make them comparable in size. The leads to an equation for the Attention Mechanism network model output as a result of point-multiplying the feature map outputs from the mask and stem branches:

$$H_{i,c}(x) = \left(1 + M_{i,c}(x)\right) * T_{i,c}(x) \text{ ----- (3)}$$

with $M_{i,c}(x)$ varying from 0 to 1, and $T_{i,c}(x)$ representing the location of the i^{th} pixel point, the result of the trunk branch's feature map matching the c^{th} channel's location.

Algorithm 3: HNN Model

Input:

- Labeled bird image dataset for training
- Labeled bird image dataset for testing

Output:

- Trained HNN model
- Testing accuracy and predictions

Steps

- Initialize the HNN model with its three-layer architecture and special activation functions.
  - $q_i = \frac{1}{\sum_{j=1}^{k}\left[w_{ij}\left(q_j^{im} - m_i\right)\right]^2}$
- Utilize the last two layers of a pre-trained VGG-16 CNN for transfer learning.
- Adapt the VGG-16 model for bird image classification
- Train a Convolutional Neural Network (CNN)
- Improve the classification accuracy of the DenseNet network model using the strengths of the residual attention mechanism.
  - $H_{i,c}(x) = \left(1 + M_{i,c}(x)\right) * T_{i,c}(x)$
- Integrate the trained VGG-16, CNN, and Modified DenseNet models into the Hybrid Neural Network (HNN).
- HNN Output=Activation(W_{HNN}·Concatenate(VGG-16 Output, CNN Output, DenseNet Output))
- Train the HNN on the labeled bird image dataset for improved accuracy.
- Evaluate the testing accuracy of the trained HNN model on the testing dataset.

4. Results and Discussion

The BSI-EUNet model has been implemented by using a Python programming. First, a dataset of bird images is collected and preprocessed. The preprocessing involves techniques such as denoising and filtering to remove any noise or artefacts in the images that may interfere with accurate identification. Next, a HNN is employed to train the dataset. Combining a VGG16 model and a custom convolutional neural network (CNN) with a modified descent method to optimize the network's parameters. Finally, the performance of the automated bird species recognition system is evaluated using various metrics and the results are compared to those of existing methodologies to assess its effectiveness.

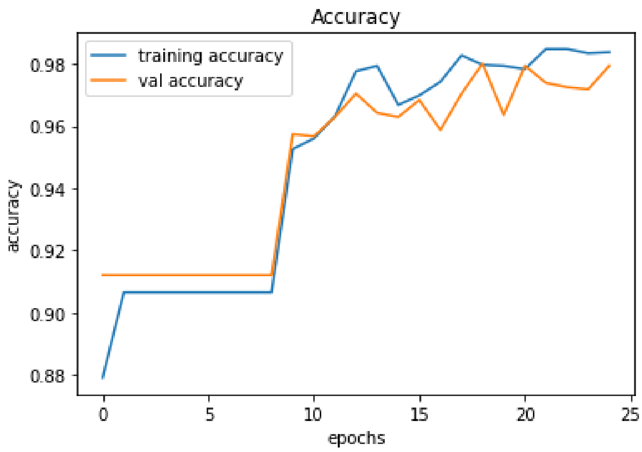

Fig 5: Training accuracy

The graph in Figure 5 illustrates the training accuracy of the model over a series of epochs. The x-axis indicates the epoch number, representing the number of times the model has undergone training. The y-axis shows the precision, which measures the accuracy of the model's predictions.

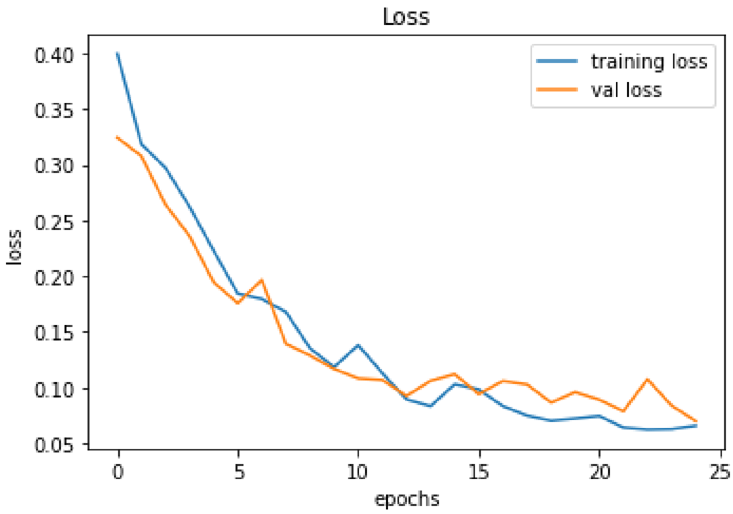

Fig 6: Training loss

The graph in Figure 6 represents the training loss of a machine-learning model across a series of epochs. The x-axis indicates the epoch number, representing the number of times the model has undergone training. The loss, shown on the y-axis, is a metric for evaluating the model's efficacy in reducing the variance between the anticipated and observed results.

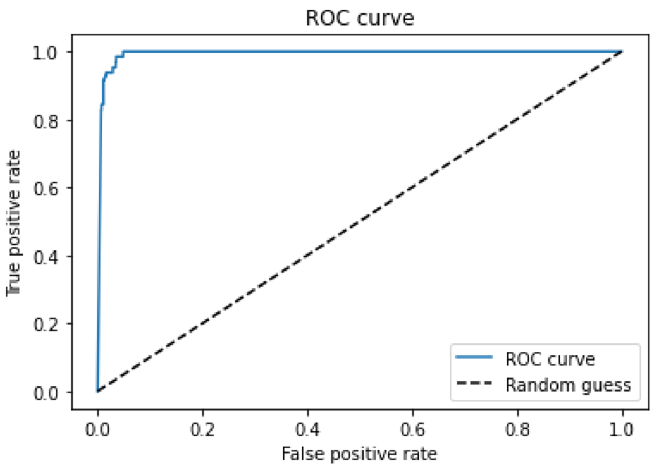

Fig 7: ROC curve

The performance of a binary classifier is shown graphically in Figure 7 as the Receiver Operating Characteristic (ROC) curve. The percentage of really negative samples wrongly projected as positive is the false positive rate, shown along the x-axis. The percentage of real positive samples accurately predicted as positive is shown along the y-axis as the true positive rate.

Fig 8: Original image

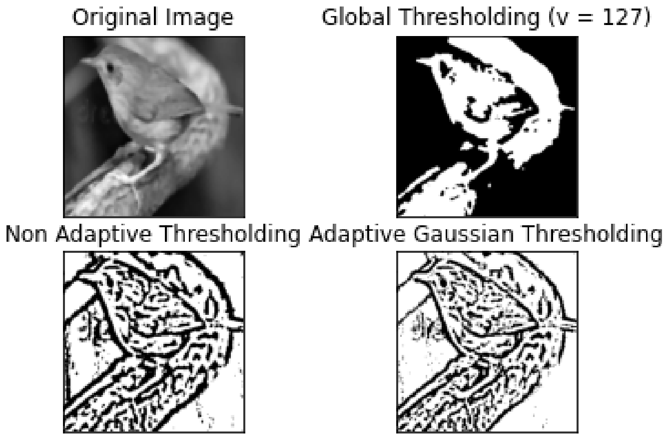

Fig 9: Non-adaptive thresholding and adaptive Gaussian thresholding

Figure 9 compares two commonly used image processing techniques for segmenting images into foreground and background regions based on pixel intensity values. The two techniques being compared are non-adaptive thresholding and adaptive Gaussian thresholding.

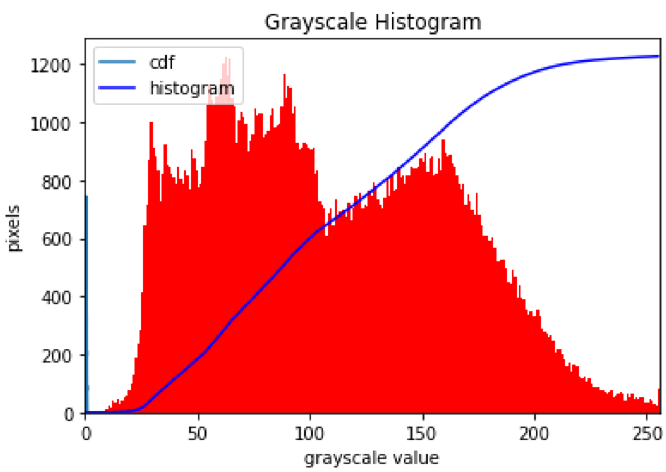

Fig 10: grayscale histogram

The grayscale histogram displayed in Figure 10 visually represents the distribution of pixel values in the image. Specifically, the x-axis indicates the grayscale value of each pixel, while the y-axis represents the number of pixels with that grayscale value.

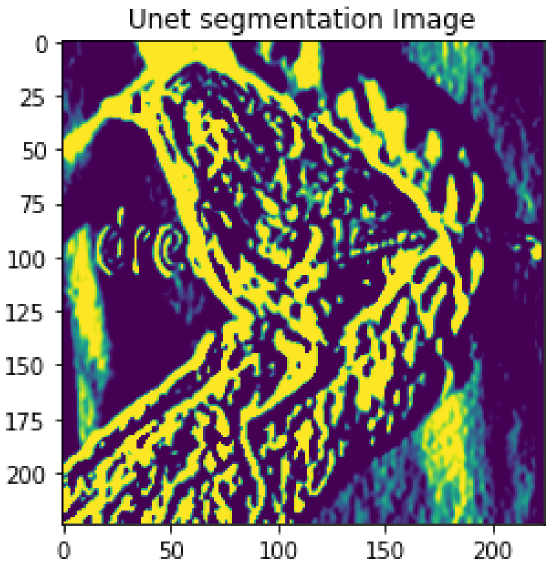

Fig 11: UNet segmentation image

Figure 11 depicts a U-Net segmentation picture, which entails analyzing the correctness of the segmentation border and identifying the various regions of interest that have been segmented. To ensure an appropriate diagnosis or conclusion, the interpretation should always be combined with any other accessible information.

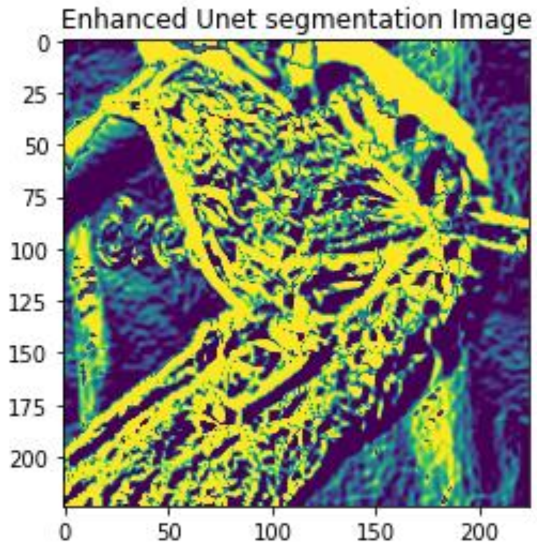

Fig 12: Enhanced UNet segmentation image.

Figure 12 depicts an improved U-Net segmentation image, which entails assessing the accuracy and amount of detail in the segmentation borders and regions and taking into account any extra characteristics or outputs offered by the upgraded model. To ensure an appropriate diagnosis or conclusion, the interpretation should always be combined with the original picture and any other accessible information, as with any segmentation model.

A) Experimental Evaluation

Our findings were analyzed within the context of several taxonomies. In this method, we split the multiclassification data, creating two separate classification problems.

TP_d: Both the forecast and the actual outcome fall within category D.

TN_d: Other subcategories of Category D are predicted and exist in practice.

FP_d: Category D is the expected outcome; however, additional d classes exist.

FN_d: Other D classes are predicted, but only D itself exists.

Each class was a training set for calculating the total accuracy, precision, and recall. The accuracy may be represented by the following Equation (4):

$$Accuracy = \frac{Number of samples correctly classified}{number of samples for all categories} \quad (4)$$

As demonstrated in Equation (5), the accuracy of the sample may be inferred from the precision of a single category:

$$Precision_i = \frac{TP_d}{TP_d + FP_d} \quad (5)$$

As indicated in Equation (6), recall for a given category may be thought of as the percentage by which a correctly predicted sample of category d covers the percentage of category d in the sample set,

$$Recall_i = \frac{TP_d}{TP_d + FN_d} \quad (6)$$

It evaluates the advantages and disadvantages of classifiers within separate categories. Implementing a macro average is recommended. The term macro-averaging refers to the arithmetic mean of all statistical index values. Their calculating formulae are given in Equations (7) through (9):

$$Precision_{macro} = \frac{\sum_{d=1}^{N} Precision_d}{N} \quad (7)$$

$$Recall_{macro} = \frac{\sum_{d=1}^{N} Recall_d}{N} \quad (8)$$

$$P1_{macro} = \frac{2Precision_{macro}Recall_{macro}}{Precision_{macro} + Recall_{macro}} \quad (9)$$

Table 1: Performance metrics

		Accuracy	Precision	Recall	F-Measure
Existing Authors	Gavali, P., & Banu, J. S.	91	90	92	89
	Islam, S. et al.	89	88	89	87
	Ragib, K. M et al.	92	92	93	91
Existing Algorithms	CNN	94	93	95	90
	DNN	95	96	96	92
	DenseNet	93	93	94	92
	VGG16	95	95	94	95
	VGG16 with CNN	96	96	96	92
	UNET	95.2	94	96	95
Proposed Method	BSI-EUNET	97.9	98	97	98

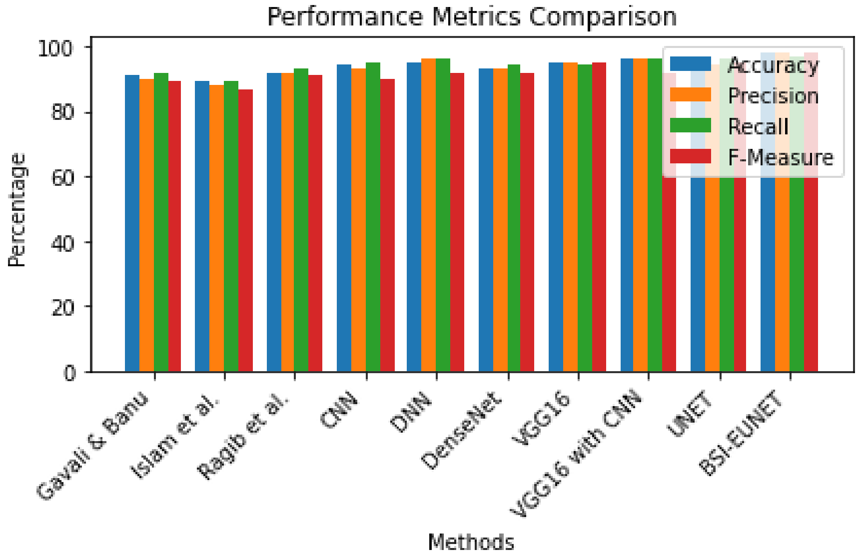

Fig 13: Classification Comparison Chart

Table 1 and image 13 depict the specific metric, encompassing Accuracy, Precision, Recall, and F-Measure. The compared approaches encompass both established algorithms and the proposed BSI-EUNET method. BSI-EUNET surpasses other methods in terms of Accuracy, obtaining an amazing 97.9%. This signifies the model's aptitude in accurately forecasting outcomes across all cases. Other notable performances in terms of accuracy are UNET, VGG16 with CNN, and DNN, all exhibiting values ranging from 94% to 96%. When assessing Precision, which quantifies the model's ability to limit false positives, BSI-EUNET once again exhibits dominance with a precision rate of 98%. Both DNN and VGG16, when combined with CNN, exhibit a high level of precision, with both obtaining an accuracy rate of 96%. The recall statistic of BSI-EUNET, which gauges the model's ability to accurately identify pertinent occurrences, showcases its remarkable performance with a recall rate of approximately 97%. CNN, DNN, VGG16, and UNET demonstrate high recall performance, achieving values ranging from 94% to 96%. The BSI-EUNET technique constantly surpasses in all parameters, showcasing its resilience and efficacy in bird species identification. The findings underscore the capability of the BSI-EUNET model to achieve accurate and exact categorization, demonstrating encouraging advancements in the domain of image-based species identification.

5. Conclusion

To summaries, the BSI-EUNet Model has been introduced for the purpose of segmenting and categorizing bird species. BSI-EUNET exhibits an impressive precision rate of 97.9%, highlighting its proficiency in properly forecasting different instances of avian species. The model's impressive precision rate of 98% showcases its ability to significantly minimize the frequency of false positives, a critical factor in guaranteeing precise species identification. Moreover, the Recall rate of BSI-EUNET, approximately 97%, showcases its ability to precisely detect a substantial number of pertinent instances, thereby emphasizing its robust performance in identifying diverse avian species. Additional well-known methodologies, including as UNET, VGG16 with CNN, and DNN, also demonstrate remarkable efficacy, further highlighting the significance of deep learning approaches in bird species recognition. BSI-EUNET's consistently outstanding performance in all criteria positions it as a promising and advanced model for image-based bird species identification. The findings emphasize the substantial impact of the BSI-EUNET model on improving the dependability and accuracy of bird species identification systems. BSI-EUNET is a reliable and effective solution in the developing field of species identification using images. It provides innovations that can improve our comprehension and conservation efforts regarding bird biodiversity. Additional investigation and verification in various practical situations could strengthen BSI-EUNET's position as a helpful instrument in ornithological research and conservation endeavors.

References

[1] Gavali, P., & Banu, J. S. (2020). Bird Species Identification using Deep Learning on GPU platform. 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE). doi:10.1109/ic-etite47903.2020.85

[2] Islam, S., Khan, S. I. A., Abedin, M. M., Habibullah, K. M., & Das, A. K. (2019). Bird Species Classification from an Image Using VGG-16 Network. Proceedings of the 2019 7th International Conference on Computer and

Communications Management - ICCCM 2019. doi:10.1145/3348445.3348480

[3] Jancovic, P., & Kokuer, M. (2019). Bird Species Recognition using Unsupervised Modelling of Individual Vocalisation Elements. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 1–1. doi:10.1109/tap.2019.2904790

[4] Jian, L., Lei, Z., & Baoping, Y. (2014). Research and Application of Bird Species Identification Algorithm Based on Image Features. 2014 International Symposium on Computer, Consumer and Control. doi:10.1109/is3c.2014.47

[5] Johnston, A., Thaxter, C. B., Austin, G. E., Cook, A. S. C. P., Humphreys, E. M., Still, D. A., … Burton, N. H. K. (2014). Modelling the abundance and distribution of marine birds accounting for uncertain species identification. Journal of Applied Ecology, 52(1), 150–160. doi:10.1111/1365-2664.12364

[6] Lucio, D. R., Maldonado, Y., & da Costa, G. (2015). Bird species classification using spectrograms. 2015 Latin American Computing Conference (CLEI). doi:10.1109/clei.2015.7359990

[7] Marini, A., Turatti, A. J., Britto, A. S., & Koerich, A. L. (2015). Visual and acoustic identification of bird species. 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). doi:10.1109/icassp.2015.7178383

[8] Nanni, L., Costa, Y. M. G., Lucio, D. R., Silla, C. N., & Brahnam, S. (2016). Combining Visual and Acoustic Features for Bird Species Classification. 2016 IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI). doi:10.1109/ictai.2016.0067

[9] Pang, C., Li, H., Cherian, A., & Yao, H. (2017). Part-based fine-grained bird image retrieval respecting species correlation. 2017 IEEE International Conference on Image Processing (ICIP). doi:10.1109/icip.2017.8296812

[10] Qiao, B., Zhou, Z., Yang, H., & Cao, J. (2017). Bird species recognition based on SVM classifier and decision tree. 2017 First International Conference on Electronics Instrumentation & Information Systems (EIIS). doi:10.1109/eiis.2017.8298548

[11] Qiu, Z., Zhu, X., Shi, D., & Kuang, Y. (2020). Recognition of Transmission Line Related Bird Species Based on Image Feature Extraction and Support Vector Machine. 2020 IEEE International Conference on High Voltage Engineering and Application (ICHVE). doi:10.1109/ichve49031.2020.9279508

[12] Ragib, K. M., Shithi, R. T., Haq, S. A., Hasan, M., Sakib, K. M., & Farah, T. (2020). PakhiChini: Automatic Bird Species Identification Using Deep Learning. 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4). doi:10.1109/worlds450073.2020.9210259

[13] Roslan, R., Nazery, N. A., Jamil, N., & Hamzah, R. (2017). Colour-based bird image classification using Support Vector Machine. 2017 IEEE 6th Global Conference on Consumer Electronics (GCCE). doi:10.1109/gcce.2017.8229492

[14] Zottesso, R. H. D., Costa, Y. M. G., Bertolini, D., & Oliveira, L. E. S. (2018). Bird species identification using spectrogram and dissimilarity approach. Ecological Informatics. doi:10.1016/j.ecoinf.2018.08.007

[15] Vinisha, A. ., & Boda, R. . (2023). A Novel Framework for Brain Tumor Segmentation using Neuro Trypetidae Fruit Fly-Based UNet. *International Journal of Intelligent Systems and Applications in Engineering*, *12*(1s), 783–796.

[16] Dhulipalla, R. K. ., & Sethuraman, S. C. . (2023). Hybrid VGG16-Abstract Neural Network Model based Bird Detection and Classification of Images. *International Journal of Intelligent Systems and Applications in Engineering*, *11*(4), 572–582.

[17] Ullah, F., et al. (2024). Deep Hyperspectral Shots: Deep Snap Smooth Wavelet Convolutional Neural Network Shots Ensemble for Hyperspectral Image Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *17*, 14-34. doi: 10.1109/JSTARS.2023.3314900.

[18] Zhang, K., Guo, Y., Wang, X., Yuan, J., & Ding, Q. (2019). Multiple Feature Reweight DenseNet for Image Classification. *IEEE Access*, *7*, 9872-9880. doi: 10.1109/ACCESS.2018.2890127.

[19] Bui, T. V., Cheraghchi, M., & Echizen, I. (2021). Improved Non-Adaptive Algorithms for Threshold Group Testing With a Gap. *IEEE Transactions on Information Theory*, *67*(11), 7180-7196. doi: 10.1109/TIT.2021.3104670.