
ISSN: 2321-2152
IJMECE
International Journal of modern electronics and communication engineering
E-Mail
editor.ijmece@gmail.com
editor@ijmece.com
www.ijmece.com

ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

CNN for Classification of Butterflies

[1] Pradyumna Yambair, [2] T. Pooja,

[1]Assistant Professor, Megha Institute of Engineering & Technology for Women, Ghatkesar.
[2] MCA Student, Megha Institute of Engineering & Technology for Women, Ghatkesar.

Abstract—

In terms of ecological, the butterfly is the most important animal. In order for many ecological processes to function, the connection between plants and butterflies is crucial. You may hear them called "flying flowers" or "flowers of many hues" from time to time. When it comes to pollination, butterflies are crucial. People are killing out this beautiful species for several reasons, such as poisoning plants, destroying their habitat, and not realizing how important butterflies are to the environment. This research is an extension of previous work that used deep learning and image processing to categorize butterflies. Butterflies may be categorized based on their external morphological traits (structure) and genital traits. Preserving the environment and managing pests in forests and farms are only two examples of the numerous real-world applications of studies into the taxonomy of butterfly species. Nevertheless, categorization by type is essential for ease of identification due to the abundance of types and styles. By training and testing on a butterfly dataset, several machine learning approaches are used to get the best possible answer. These techniques include conventional machine learning, deep learning, and transfer learning. By taking a photograph of a butterfly in real-time or selecting an image from a gallery, this application can determine its category.

Keywords—

Image processing, convolutional neural networks, morphological traits, CNN

I. INTRODUCTION

Among the many features shared by insects, butterflies include three sets of legs, an abdomen, a thorax, a head with two compound eyes, and two sets of "scaly" wings. Together with moths, they make up the insect order Lepidoptera. Researchers and students who study lepidoptera make this observation. The name "Lepidoptera," meaning scaly wings, is derived from the Greek words "lepido," meaning scale, and "ptera," meaning wings. India is host to around 1500 species and subspecies out of 17,000 total. India is a global hub for butterfly diversity, with 20% of these species being endemic to the country. The study of butterflies is an important part of biology. An important part of every ecosystem is the butterfly. There are presently no reliable ways for butterfly identification since there are so many species in the lepidoptera family and they all have different shapes and colors. Therefore, convolutional neural networks (CNNs) for picture processing are suggested for the purpose of butterfly species identification. Convolutional neural networks (CNNs) are famous deep learning methods for image segmentation, object recognition, and classification. The most crucial aspect to keep in mind is that training a deep learning model requires a substantial quantity of data and computational capacity. The outcome will be more precise if you input a butterfly photograph, whether it captured in real-time or selected from the gallery. Using training data stored in a database, this research aims to allow the recognition and classification of butterfly photos based on their wings. The data collection includes a wide variety of butterfly species and was collected via Kaggle. This research shows that many species of butterflies may be classified using convolutional neural network (CNN) methods. These species include Danaus plexippus, Heliconiuscharitonius, Heliconiuserato, Junoniacoenia, Lycaenaphlaeas, Nymphalisantiopa, Papiliocresphontes, Pierisrapae, Vanessa atalanta, and Vanessa cardui.

II. LITERATURE SURVEY

Ghazanfar et al. [1] conducted a comprehensive investigation comparing ecological images with photographs of butterfly specimens with the purpose of classifying them. In order to classify ecological photographs, Xin et al. [2] provide a novel classification that uses the squeeze-and-excitation (SE) module, dilated residual network, and spatial attention (SA) module. The SA module is used to better utilize the pictures' long-range dependencies. The SE module, on the other hand, makes use of data from across the world to identify and remove aspects that aren't beneficial. Lastly, the model outperformed the standard approaches in the experiments. To identify various butterfly species, including Grey Pansy, Chocolate Grass Yellow, Black Veined Tiger, and Plain Lacewing, Almryadet al. [3] proposed a pre-trained CNN architecture model called Google Net. Using 120 images representing four distinct butterfly species, the researchers were able to demonstrate that the algorithm achieved an overall accuracy of 97.5%. Research [4] built, implemented, and evaluated a system for butterfly identification and detection using convolutional neural networks (CNN). Because it is consistent, the detection and identification procedure has an accuracy of 92.7%. By a margin of 62.5%, training results correspond to real system installation. The approach may sort groups of butterflies into categories based on their developmental stage, however the accuracy of this classification is heavily dependent on the viewing angle. A different research [5] created a system for automatically detecting and identifying butterfly species by using Faster R-CNN (Region-based Convolutional Neural Network) in various ecological systems. Classification accuracy was achieved, on average, up to 70.4%. The author of research [6] compares two texture descriptors, local binary patterns (LBP) and grey level cooccurrence matrix (GLCM), using 190 photos of butterflies from 19 species of the pieridae family. For GLCM, the proposed method achieved a precision of 98.25% and for LBP, of 96.45%. Data from existing systems are compared with data obtained from selected surveys in study [7]. Weband is building a smartphone app with features like expert butterfly species verification and picture classification. The presence, history, and advantages of butterflies have been the subject of several studies. A field-based dataset was created by [8] using photographs of butterflies captured in their natural habitat and recognized by entomology specialists. Feature extraction techniques are superfluous when using deep learning (DL) architectures to recognize butterflies in input photos. The transfer learning procedure made use of trained models. Results were compared and evaluated using three distinct network topologies in the experiment. Table 1 displays the results of a comparative examination of prior studies that made use of deep learning techniques.

TABLE I. COMPARATIVE ANALYSIS OF PREVIOUS RESEARCH WORK CARRIED USING DEEP LEARNING APPROACHES

Author	CNN MODEL	TEST ACCURACY	Reference
Almryad et al.	VGG16	79.5%	[3]
	VGG19	77.2%	
	RESNET	70.2%	
Zhao et al.	Faster R-CNN	70.4%	[5]

Fig. 1. Various Species Images of Butterflies

"Danaus plexippus," meaning "sleepy transformation" in Greek, is the scientific name for the monarch butterfly. In addition to having less distinct white dot patterns, butterfly wings do not have black wing veins. The zebra longwing butterfly, scientifically known as Heliconiuscharitonius, is easily identifiable by its long, thin wings that are striped black and light yellow. The genus Heliconius includes over 40 different species of neotropical butterflies, one of which being the red postman, or Heliconiuserato. The small postman, red passion flower butterfly, and crimson-patched longwing are some of its alternate names. Butterfly species Junoniacoenia, also called buckeye or common buckeye, belongs to the Nymphalidae family. East of the Rocky Mountains in the US and Mexico is where you may find it. It lives in wide open spaces with little vegetation and occasional bare spots. Little copper, American copper, or common copper butterflies are all members of the Lycaenid family of gossamer-winged butterflies. The specific name phlaeas is thought by Guppy and Shepard to have originated from the Latin floreo, meaning "to flourish," or from the Greek phlaeas, meaning "to burn up." One of the most extensively distributed butterfly species in our country is the gigantic and visually striking mourning cloak, scientifically known as Nymphalisantiopa (Linnaeus). Where it spends the winter, adults may be observed sunbathing on warm days for the better part of every month. When the mourning cloak is at rest, the wings are closed, yet the top surfaces of the wings are still quite lovely. The Camberwell beauty is the name it's given in the United Kingdom.

III. METHODOLOGY

Classifying butterfly photos using convolutional neural network (CNN) algorithms and evaluating the results is the goal of this study. A total of 890 butterfly images were collected. Prior to being downsized to 224x224 pixels, images are acquired in a range of sizes. The Monarch butterfly, Common Buckeye, American Copper, Nymphalisantiopa, Giant Swallowtail, Pieris rapae, Cabbage White, Vanessa atalanta, Vanessa cardui, and Heliconiuscharitonius are examples of the many types of butterflies shown in Figure 1. Giant Swallowtail butterflies, or Papiliocresphontes, are huge and have been seen on rare occasions in the Adirondack Mountains of upstate New York. The Papilionidae family includes this species, which gets its name from the fact that its rear wings have projections that look like swallowtails. The term "cresphontes" is supposedly derived from the legendary Greek figure Cresphonte. The tiny to medium-sized butterfly, Pieris rapae One of the white-and-yellow families, Pieridae, includes Pieris rapae. It goes by the name "little cabbage white" on a number of different continents. It is possible to distinguish this butterfly by its bigger size, the presence of a black band at the tip of its forewings, and the fact that its wings are white with little black dots. The Vanessa atalanta is a large-winged brush-footed butterfly (order Lepidoptera) with a pattern of scales that ranges from reddish orange to pink to brown to white and blue. During the spring, a large number of individuals make the journey from Africa to Europe across the Mediterranean. One of the most widely distributed butterfly species, the Painted Lady (Vanessa cardui) can be found everywhere on Earth with the exception of Antarctica and Australia. Fields and meadows are common places to see Painted Ladies. Part A: A Convolutional Neural Network In addition to representation learning, deep learning is a part of a larger family of approaches based on artificial neural networks. Because it can process very large datasets, deep learning has proven to be an exceptionally useful technology. Hidden layer technology is mostly used for pattern recognition, in contrast to more traditional methods. In the realm of deep neural networks, convolutional neural networks are among the most often used. A specific kind of feed-forward neural network known as a convolutional network processes data in a grid-like arrangement in order to analyze visual pictures.

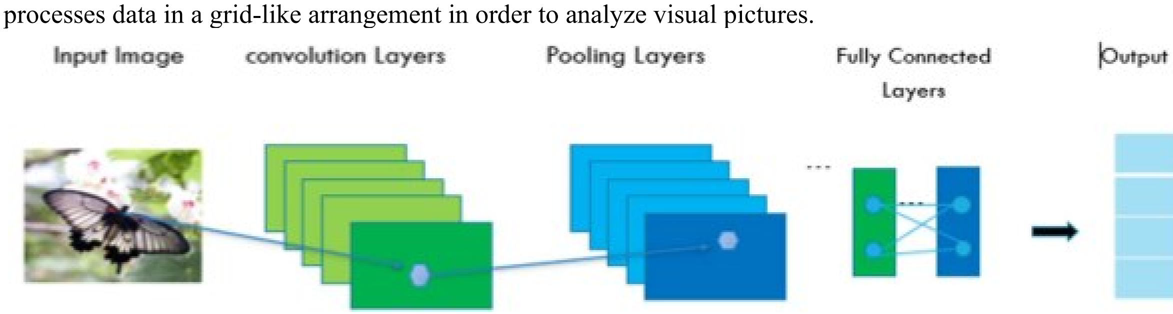

Fig. 2. Architecture of Convolutional Neural Network.

A convolution neural network's several hidden layers aid in the extraction of information from pictures. Each convolutional neural network (CNN) consists of four primary layers: the convolutional layer, the pooling layer, the fully connected layer, and the receiving unit (ReLU). To begin removing relevant details from a picture, the convolution layer is used. A convolution layer has several filters that work together to perform the convolution function. You might say that every picture is a matrix of pixel values. The main goal of convolution is to retrieve characteristics from the source pictures. Producing the desired result is as simple as feeding the input picture into the convolutional layer and applying the filter. The acronym for rectified linear unit is ReLU. The feature maps need to be moved to a ReLU layer after retrieval. Performing an operation element by element, ReLU changes all the negative pixels to 0. The final product is a nonlinear network with a revised feature map. An activation function that goes by the name "ReLU" stands for a rectified linear unit. Down sampling is a part of pooling that reduces the feature map's dimensionality. Next, a pooling layer is applied to the corrected feature map, resulting in a pooled feature map. To isolate certain picture elements—such as beaks, feathers, edges, corners, and bodies—the pooling layer employs a number of filters. The two main varieties of pooling are maximum and average. The flatten layer is responsible for transforming matrices into vectors. When talking about a network, the last several levels are called Fully Connected Layers. Before being used, the output of the last convolutional or pooling layer is flattened and transferred to the fully connected layer. The output layer is connected to the SoftMax activation function. The application function SoftMax shows which class is most likely to be chosen as the final result. Section B: Butterfly Taxonomy: In picture categorization, several groups are created and the user is provided with a prediction accuracy rating. To classify butterflies, one uses the Softmax classifier. When this classifier makes a prediction, it shows the highest possible probability for that class. Among the butterfly species included in the dataset are Lycaenaphlaeas, Junoniacoenia, Lycaenacharitonius, Heliconiuserato, Nymphalisantiopa, Papiliocresphontes, Pieris rapae, Vanessa atalanta, and Vanessa cardui. The suggested approach's general design is shown in Figure 3. . The Training Process: Figure 4 shows the initial data used to train the robots. The goal of training datasets for deep learning algorithms is to teach them how to do tasks or make predictions. Following exposure to increasingly distorted and trained pictures, training is a process that aids in producing improved results. In this work, we took a look at the 6,500-image Butterfly Classification Training Data Set, which covers 75 different classes and was painstakingly assembled. D. Testing: Evaluating data post-model generation confirms the model's forecast accuracy. See Figure 5 for an example of how test data might be used as a last-ditch effort to verify if a new dataset was adequately used to train the CNN algorithm.

IV. RESULTS & DISCUSSIONS

We find out how well the butterfly categorization is. In this research, testing uses 30% of the data and training uses 70%. Table 2 and Figure 6 show the results of the testing and training data. Lycaenaphlaeas, Junoniacoenia, Lycaenacharitonius, Heliconiuserato, Nymphalisantiopa, Papiliocresphontes, Pieris rapae, Vanessa atalanta, and Vanessa cardui were the 10 species of butterflies to be identified.

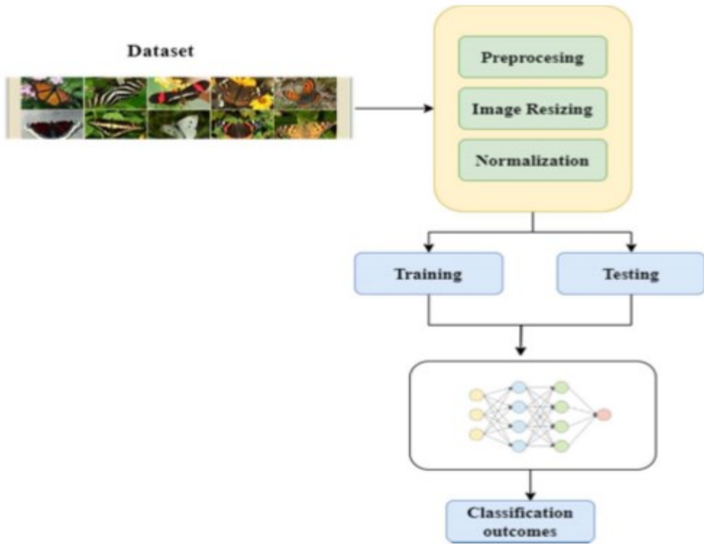

Fig. 3. Overall architecture of Proposed approach

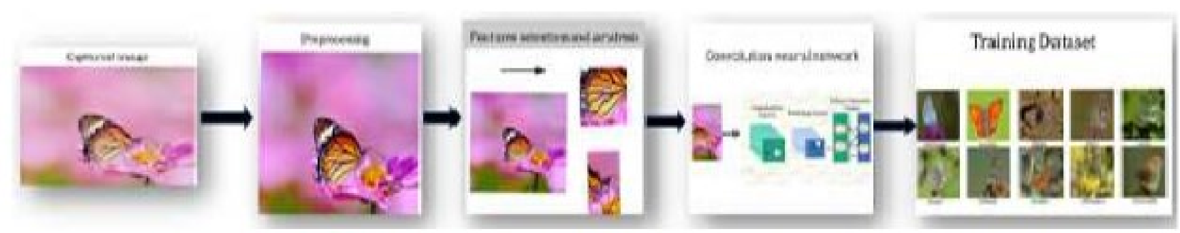

Fig. 4. Training Process

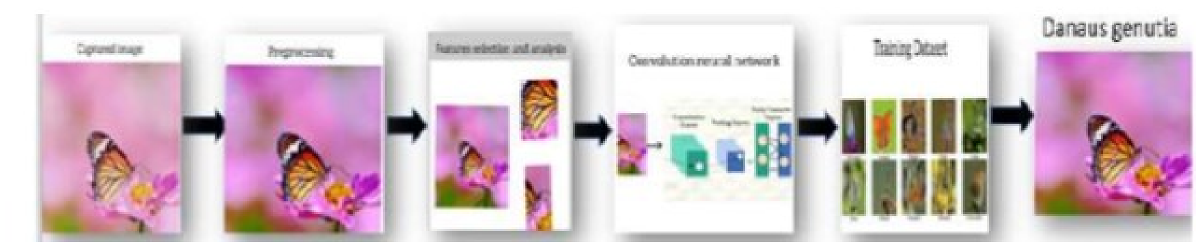

Fig. 5. Testing Process

TABLE II. TRAINING AND TESTING ACCURACY OF PROPOSED APPROACH

Approach	Testing Accuracy observed	Training Accuracy observed
ConvolutionalNeural Network	98%	94%

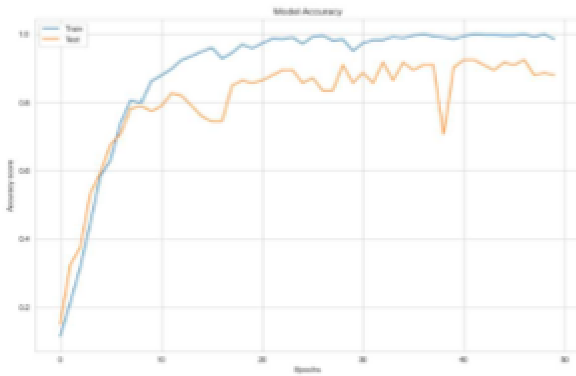

Fig. 6. Accuracy curve of Training and Testing Data

V. CONCLUSIONS

A comprehensive comprehension of the many butterfly species was achieved via reading several research papers and publications. The authors of this paper proposed using CNNs to classify different kinds of butterflies. Several sources (Kaggle) were combed through in order to compile images of various butterfly species for this project. Here are some sample images to get you started. The scales, color, and habitat of a butterfly allow it to be classified into many groups. We also pay close attention to image processing techniques, such as those for filtering and noise reduction. With a training accuracy of 98% and an installation accuracy of 84%, the detection and identification system is consistently performing well. Thanks to ongoing improvements, the system can now identify different types of butterfly groups.

REFERENCES

[1]Ghazanfar, Mobeen, et al. "Butterflies and their contribution in ecosystem: A review." Journal of Entomology and Zoology Studies 4.2 (2016): 115-118.

[2]Xin, Dongjun, Yen-Wei Chen, and Jianjun Li. "Fine-grained butterfly classification in ecological images using squeeze-and-excitation and spatial attention modules." Applied Sciences 10.5 (2020): 1681.

[3]Almryad, Ayad Saad, and Hakan Kutucu. "Automatic identification for field butterflies by convolutional neural networks." Engineering

Science and Technology, an International Journal 23.1 (2020): 189195.

[4]Priya, R. L., et al. "Patang Abhidhani-Convolution Neural Network based Butterfly Research Survey." Journal of Physics: Conference Series. Vol. 1916. No. 1. IOP Publishing, 2021.

[5]Zhao, Ruoyan, et al. "Butterfly recognition based on faster R-CNN." Journal of Physics: Conference Series. Vol. 1176. No. 3. IOP Publishing, 2019.

[6]Selvaraju, Ramprasaath R., et al. "Grad-cam: Visual explanations from deep networks via gradient-based localization." Proceedings of the IEEE international conference on computer vision. 2017.

[7]Howard, Andrew G., et al. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." arXiv preprint arXiv:1704.04861 (2017).

ISSN 2321-2152
www.ijmece.com
Vol 13, Issue 2, 2025

[8]Arzar, Nur Nabila Kamaron, et al. "Butterfly species identification using convolutional neural network (CNN)." 2019 IEEE international conference on automatic control and intelligent systems (I2CACIS). IEEE, 2019.

[9]Hu, Jie, Li Shen, and Gang Sun. "Squeeze-and-excitation networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

[10]Sabri, Nurbaity, et al. "Comparing convolution neural network models for leaf recognition." International Journal of Engineering & Technology 7.15 (2018): 141-144.

[11]Azizah, Laila Ma'rifatul, et al. "Deep learning implementation using convolutional neural network in mangosteen surface defect detection." 2017 7th IEEE international conference on control system, computing and engineering (ICCSCE). IEEE, 2017.

[12]Bakri, Badrul Aiman, Zaaba Ahmad, and Shahirah Mohamed Hatim. "Butterfly family detection and identification using convolutional neural network for lepidopterology." Int. J. Recent Technol. Eng 8.2 (2019): 635-640.

[13]Naidu, Balaka Ramesh, and PVGD Prasad Reddy. "Fusion of face and voice for a multimodal biometric recognition system." International Journal of Engineering and Advanced Technology (IJEAT) 8.3 (2019).

[14]Hu, Benyi, et al. "Multi-label x-ray imagery classification via bottomup attention and meta fusion." Proceedings of the Asian Conference on Computer Vision. 2020.

[15]Tang, Pengjie, Hanli Wang, and Sam Kwong. "G-MS2F: GoogLeNet based multi-stage feature fusion of deep CNN for scene recognition." Neurocomputing 225 (2017): 188-197.