

# Butterfly Family Detection and Identification Using Convolutional Neural Network for Lepidopterology

Badrul Aiman Bakri, Zaaba Ahmad, Shahirah Mohamed Hatim

**Abstract**— *Lepidopterology is a branch of entomology concerning the scientific study of moths and the three superfamilies of butterflies. The project aims to help biology students in identifying butterfly without harming the insect. In the studies of lepidopterology, the students normally need to capture the butterflies with nets and dissect the insect to identify its family types. Computer vision is a study on how computers can be used to make high-level comprehension from the input of digital image and videos. By utilizing the latest Image Processing technique, it can identify the correct species of butterfly with high accuracy by using layers of node in a Convolutional Neural Network (CNN). The work process starts with data acquisition (mining the butterfly image automatically from google image search), pre-processing (converting image format and rotation), analyzing and understanding digital images (group images into folders), and to make assumptions of the high complication data from the real world in the process of producing numerical information that can be comprehend by machines in order to form conclusions. Benefits of using CNN is to reduce the need for human and physical intervention in identifying each of the butterfly characters. This makes it easier to expand the database in the future. The image is acquired using Fatkun Batch Downloader to download large number of images. The project is develop using Tensorflow in Ubuntu operating system and interface is in HTML connected to the Python script via Flask. The results of the experiment show that CNN can identify with 92.7 percent of final accuracy with learning saturation (overfitting) of 500 cycle. While testing results shows 62.5 percent of accuracy in predicting new datasets.*

**Keywords:** Automatic detection, convolutional neural network, flask architecture.

## I. INTRODUCTION

Butterfly research is one of the important studies in the biology field and is widely being research in all over the world. Researchers and students who conducted this observation are known as lepidopterist. Lepidoptera stands for the scientific name of the butterfly. Lepidoptera is an order of Animalia that includes moths and butterflies. This Animalia order exist during the Jurassic era until today. The most interesting features of the lepidoptera is the scales that covers the whole-body including bodies, wings and proboscis [1]. This results in various variance of butterfly wing colour

**Revised Version Manuscript Received on September 16, 2019.**

**Badrul Aiman Bakri**, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Perak Branch, Tapah Campus, 35400 Tapah Road, Perak, Malaysia.

**Zaaba Ahmad**, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Perak Branch, Tapah Campus, 35400 Tapah Road, Perak, Malaysia.

**Shahirah Mohamed Hatim**, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Perak Branch, Tapah Campus, 35400 Tapah Road, Perak, Malaysia.

and shape.

There are a large number of species in the lepidoptera family with various wing shape and color, making identification on this vast numbers of species are difficult [1]. Thus, a tool needs to be introduced in conjunction to reduce the burden for the routine identification of butterfly.

In the process for conventional butterfly identification, the targeted butterfly species needs to be captured manually by trap. This is not only consumed time and energy, but it also harmful to the butterfly as it needs to be confined in small spaces for a long time to identify the differences in the abdomen and wings [3]. Not only that, the characteristics of the butterfly is then needs to be identified through the encyclopedia which is large and not portable to be carried around in a field trip [4]. Through the introduction of a handy mobile application, the problem of bringing an encyclopedia can be solve by only using the database which is included in the application. Since smartphone is a daily companion of human nowadays, they will always keep mobile phone with them. This is convenient for them as they do not need to remind themselves to bring specific gear or tools when they are going on a field trip [3].

This project aims to propose a web application that allows biology students and researchers to identify the captured image of the butterfly by the wings and classify them based on the trained data that is stored in the database. The application will store the information of the classified butterfly and can be further develop for butterfly conservation purposes soon so that the ecosystem of butterfly can be recorded and conserved.

Butterfly recognition is important for several reasons. One of the main reasons is for butterfly surveying and record. Butterfly survey is conducted in camps where the identified butterfly is the recorded and the species is being counted [5]. In order to help in the education department, the introduction of an automated system is needed to reduce the burden of routine identification.

## II. METHODOLOGY

Image processing is a modern way on how the computer sees things and identify them using machine learning [6]. Image processing is a method to perform operations on an image in order to enhance the image and extract useful information from it. It is a type of signal processing where an

input is an image and the output will be feature or characteristics related to the image [7]. Image processing ties closely with the core of research area within engineering and computer science disciplines. Image processing is important in classifying based on what the machine sees and identify in a much faster way [8]. There are roughly two types of image processing, which are analog and digital image processing. Analog processing is taking a hard copy image such as printouts and photographs. In analog processing, image analyst needs to painstakingly use various fundamentals of interpretation using visual techniques. While, digital image processing techniques use the help of computers in manipulating the digital images. Using the digital technique, it will undergo three general phases which are pre-processing, enhancement and information extraction.

This project aims to classify butterfly according to their family based on the image of its wings. It is useful for making survey to count the butterfly family in that population or environment.

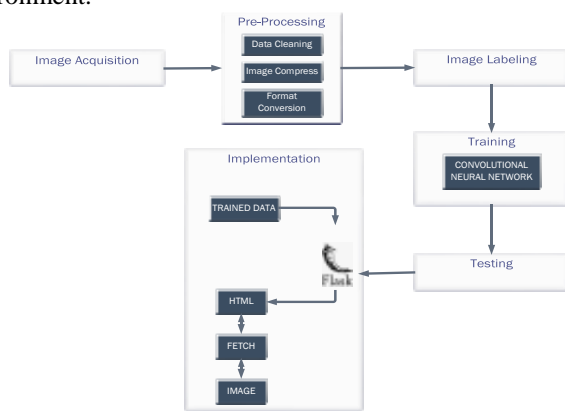


Fig. 1: Proposed process flow of butterfly identification architecture

Fig. 1 shows the process flow of the development for butterfly identification. Image first need to be acquired before pre-processing can take place. Training will then take place to train using the CNN model. After that, it will be tested to determine accuracy of the model. The trained set will then be implemented in web-based interface for identification process to take place.

Python is the programming language to be used, it provides a general-purpose of programming language that was founded by [1]. Python easily became very popular mainly because of the simplicity and reliability provided by the software, it allows the user to code with fewer lines without sacrificing any reliability and readability issues in the meantime. Python features that really stands is that it helps to wrap the codes which are coded in C or C++ so that it can be used in the Python modules. This gives two advantages. The first is that the written code is as fast as the original C/C++ code and second, it is very easy to code in Python. This is how Python works; it is a Python wrapper around original C++ implementation [9]. And the support of Numpy makes the task easier. Numpy is a highly optimized library for numerical operations. It gives a MATLAB-style syntax [9]. All the array structures are converted to-and-from Numpy arrays, so whatever operations that can do in Numpy, it can be combined with Python, which increases the chances of compatibility and diversity [9]. Besides that, several other

libraries like SciPy, Matplotlib which supports Numpy can be used with this. Python is an appropriate tool for fast prototyping of computer vision problems. Python is the most comprehensive open source library for computer vision, and it has a large user community. It has more functions for computer vision than MATLAB because many of its functions are implemented on GPU. The library is being continuously updated. Its programs can be made fast too, however, you will have to rely on high-level functions, max functions, and have to learn how to vectorize code to achieve decent speed. Python is a language for carrying out computer vision tasks. Python language bindings will give a nice scripting interface to do matrix image operations while still having easy access to OpenCV-function [10].

A. Input image

The study focused on butterfly that can be identified by its wing structure and colour. There are four types of Butterfly which commonly referred as Dingy Skipper, Rajah Brooke, Ringlet and Tree Nymph butterfly. There are about 6,000 images collected from field trip and online web scraping. The collected images are barely sufficient for the training process as CNN needs more data as it is unsupervised learning.

The first step in making image identification is acquiring the image itself. It is important that the image feed into the algorithm is digitally processed. Digital image can be acquired in several ways [6]:

1. Offline capture using a digital camera. The captured images will be fed into the identification system.
2. Online capture is taking online images that have been captured by someone else.

This will help to speed up time in acquiring butterfly images. This type of acquisition posed a problem where the butterfly has different characteristics because of their environment and climate [11].

Table 1: Butterfly image collected

Butterfly Species	Image Collected
Dingy Skipper	1401
Large Tree Nymph	1647
Rajah Brooke	1198
Ringlet	1608

Table 1 depicts the number of images collected for the training input. The number of collected image will be kept a minimum of 1000 images to ensure more feature can be extracted without any supervised procedure.

B. Pre-processing

Training is a process that help to make results that better improve after more images is distorted and trained. Simple distortions like crops, scales, and flips helps to reflect the variations that can be expect from real world situations. This can help train the model to comprehend with natural data more accurately. There are some included parameters and network of operations to be applied to an image.

Image pre-processing is the preparation of the images before being segmented and identified. Image pre-processing

can provide benefits and can solve problems that can lead to better feature identification [12]. Raw image data from the camera may have a variety of problems. So, it is most likely to produce the best computer vision results. This is why image pre-processing needs to be considered [12]. There may be some aspects of the image that needs to be adjusted prior for analysis purpose. Some of them are sensor corrections, lighting corrections, and noise color corrections. Image pre-processing also is to enhance the image, for example, the illumination and focus enhancements of the image. The image pre-processing method used for the development are as follows:

### Cropping

Cropping is the method of placing and square at random position of the full image. The square relative size on the input image is controlled by a parameter. If the parameter is zero, then the square is the same size as the image and thus, no cropping is done. If the value is 50%, the crop square will crop half of the input image width and height.

### Scaling

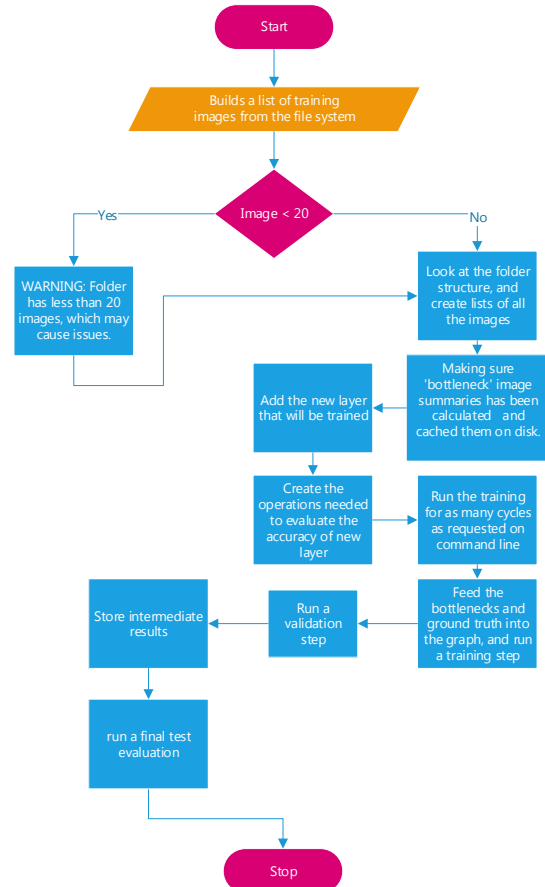
Scaling is similar to cropping, except the square is always at the center of the image. Its size varies depending on the given range. For instance, if the scale is zero, then the square will be the same size as the input and no scaling is used. If it's 50%, then the square will be in a random range between half the width and height of the image.

### Gray-scale

Grayscale is a process where each pixel in a single sample is converted to represent only the amount of light. The intensity ranges from black which the weakest intensity to white which is the strongest. Grayscale is necessary for doing image processing because the processing unit in the computers will struggle to process raw image. This is because of the complexity of the code to find the edges is much higher compared to grayscale images. To make the process much faster, some data collapsing needs to be done so that the computer can process faster. Furthermore, color information does not really help to determine the edges and other details in an image.

### C. Processing

To prepare the image to be trained by the model, the image first must be converted into .jpg format just because to make it uniform and readable by the model. After the image has been converted, the image needs to be scan for any corrupt image which can be identified by deleting images below 100kb which are too small to be read by the model. Then, the image needs to be labelled into correct butterfly family by naming the folder according to their butterfly family.



**Fig. 2: Training process**

Fig. 2 shows the training process conducted by the system. System will first check whether the folder contains more than 20 images. If more, it will continue on making bottlenecks and evaluate each layer and store the results in a graph. The processing consists of two main process which are recognition and web presentation.

### Recognition

Recognition process integrate Convolutional Neural Networks are very similar to how a brain works because the CNN is developed through understanding the structure of the brain. CNN is an algorithm that can identify the visual image for recognition. CNN is made up of neurons that have learnable weights and biases. Each neuron receives some inputs, performs a dot product and optionally follows it with a non-linearity. The whole network still expresses a single differentiable score function from the raw image pixels on one end to class scores at the other. They still have a loss function on the last layer. Convolutional neural network architectures make the explicit assumption that the inputs are images, which allows the developer to encode certain properties into the architecture [2]. These then make the forward function more efficient to implement and vastly reduce the number of parameters in the network.

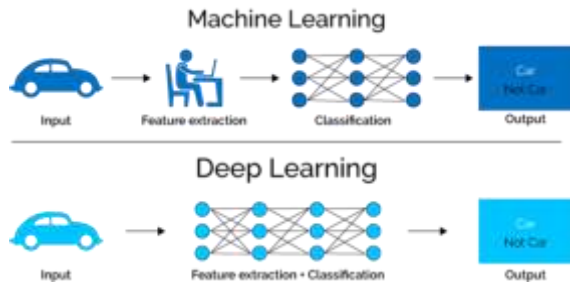


Fig. 3: Representation between SVM and CNN [13]

Based on Fig. 3, CNN does not need manual feature extraction process which eliminates the need for human labor. This makes the process of image identification much simpler. CNN can be calculated using the formula [1]:

$$n_{out} = \left\lfloor \frac{n_{in} + 2p - k}{s} \right\rfloor + 1$$

- $n_{in}$ : number of input features
- $n_{out}$ : number of output features
- $k$ : convolution kernel size
- $p$ : convolution padding size
- $s$ : convolution stride size

(1)

Cross entropy is one out of many possible loss functions. These loss functions are typically used within gradient descent, which is an iterative algorithm to move the parameters towards the optimum values. The gradient descent shows minimum loss function as the model learn. This shows that the model able to extract feature and use it to determine the butterfly category. Cross-entropy is determined using this formula [2]:

$$H(p, q) = - \sum_x p(x) \log q(x).$$

(2)

To develop the project, the programming languages use is important so that the performance can be optimized and readable. The best programming languages can be determined by reading and researching existing papers and articles. Based on the research that has done, Python is the most suitable language because it can be used in any application as well as support for its development. Python also can have many extensions across multiple platforms such as Android. It is hoped that this cross-platform support can help in achieving success in developing the application.

Table 2: Programming languages used

Category	Description
Image Training	Python 3.6 – Tensorflow, Numpy
Web Development	Python 3.6 - Flask HTML

Table 2 shows programming languages that will be used and the significant use of the it. After the language to be use is defined, then the development of the model can be continued based on the language that revolves around it.

Web presentation

The python script is a command-line interface where user will find it difficult to interact with the system. A friendly user interface is needed in order to make user more understand on what to do in order to identify the butterfly.

Python applet named flask is use in order to create host in the python script. This applet also enables the call of html files into the script which will make things easier to collect uploaded image and move the image into a folder named test. The script will then try to identify a single image that has been uploaded what group does the butterfly image belongs. This process of identification took about 5 seconds to be identified and display the results in html page that contains all of the information of the butterfly. User can upload and then submit to be tested for butterfly species. The output is the highest accuracy from the 4 species and display the information of the highest accuracy of butterfly species.



Fig. 4: Website main interface

Fig. 4 shows the interface of the web where user will interact with the system. It is crucial to make the interface as simple as possible as it will be easier for students to make the identification without the help from educators.



Fig. 5: Detected butterfly interface

Fig. 5 shows the result after the uploaded image has been processed. The image of the butterfly along with its information will be displayed to let the user know what the characteristics of the butterfly. The accuracy of the image can be seen at the command line that runs in the background.

III. RESULTS AND DISCUSSION

The image collected is run through 500 cycle of training. Each of the cycle will give a result of accuracy. The results will progress steadily until the learning curve can be plotted



to form a graph that can be used to see the training results to deduce reasonable conclusion for the method use.

The top layer receives as input a 2048-dimensional vector (1001-dimensional for Mobilenet) for each image. The script will train a softmax layer on top of this representation. Assuming the softmax layer contains N labels, this corresponds to learning  $N + 2048 \cdot N$  (or  $1001 \cdot N$ ) model parameters corresponding to the learned biases and weights. By default, the script will use the high accuracy, but comparatively large and slow Inception v3 model architecture. It is recommended that before validation, it is needed to have gathered good training data.

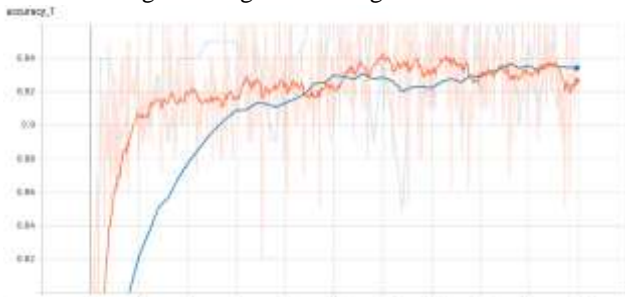


Fig. 6: Training accuracy for 500 epoch

From the training of 500 epochs in Fig. 6, orange lines represent training accuracy while blue lines represent validation accuracy. Results shows validation accuracy starts decreasing while the test accuracy increase at epoch 300. This is a clear sign that overfitting has occur at this cycle. Thus, the learning curve has been saturated at 300 epochs.

By determining the CNN for each layer, then the graph can be visualize for each layer. The receptive field is the region in the input image that a particular CNN algorithm feature is learning. In a receptive field, the nearer a pixel to the center, the more it accumulates to the calculation of the feature learning output. Which means that a feature does not only look at a particular region in the input image, but also relies more on the pixels that lies to the middle of the region. CNN architecture is asymmetric, this means it can independently make calculation of the feature map attributes separately for each layer.

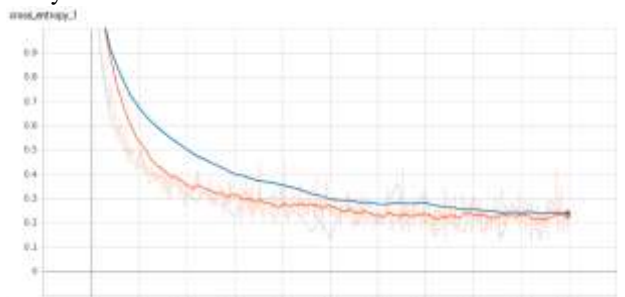


Fig. 7: Cross-entropy

Fig. 7 shows the cross-entropy of the model. Orange lines represent training loss while blue lines show validation loss. Results proof that training loss descent equally to validation loss. There no sudden increase in percentage loss during training. The graph depicts a gradient descent which shows loss function reduce over the training period. This proves the model learns by reducing the error it makes during classifying that produce more accurate end results that can be implemented in website.

Testing is done through placing a bunch of pictures into a folder named test where the python script will test all images that exist in the folder. The script will then use the bottleneck data that has been done during training to identify which group the butterfly belongs to.

Table 3: Testing results

Species	Tested Images	Identified Correctly	Percentage (%)
Rajah Brooke Butterfly	10	10	100
Large Tree Nymph	10	4	40
Dingy Skipper	10	7	70
Ringlet	10	4	40
<b>Total Accuracy</b>	<b>40</b>	<b>25</b>	<b>62.5</b>

Table 3 shows testing results conducted. Testing is done by uploading ten images for each species of butterfly. Large Tree Nymph and Ringlet butterflies scores the lowest. This does not mean the system unable to detect the butterfly, but maybe there are some anomalies that contribute for the butterfly to be unrecognized. Lack of input images also contribute to the unclassified group. The final accuracy for testing is 62.5 percent.

#### IV. CONCLUSION

We proposed a framework design and implementation for butterfly detection and identification system using convolutional neural network approach. The performance and consistency of the detection and identification system are promising with an accuracy of 92.7 percent from training and 62.5 percent from actual system implementation results. Based on the development carried on, the system able to detect the butterfly groups, but it is highly depended on the captured image and angle of the butterfly. For future works, we plan to investigate on the deeper domain of the lepidopterology studies or a bigger dataset and to explore on the other popular approaches such as recurrent and generative adversarial neural network which has been proven an outstanding achievement in other research fields.

#### V. ACKNOWLEDGEMENT

This research has been supported by Universiti Teknologi MARA Perak Branch, Tapah Campus.

#### REFERENCES

1. J. A. Powell, "Lepidoptera: Moths, butterflies," in Encyclopedia of Insects, V. H. Resh and R. T. Cardé, Eds. Massachusetts: Academic Press, 2009, pp. 559-587.
2. H. F. Nijhout, "Elements of butterfly wing patterns," Journal of Experimental Zoology, 291(3), 2001, pp. 213-225.
3. G. T. Austin and T. J. Riley, "Portable bait traps for the study of butterflies," Tropical Lepidoptera, 6(1), 1995, pp. 5-9.
4. B. H. Slager and S. B. Malcolm, "Evidence for partial

- migration in the southern monarch butterfly, *Danaus erippus*, in Bolivia and Argentina," *Biotropica*, 47(3), 2015, pp. 355-362.
5. R. Leslie, D. M. Debinski, and M. L. Wieland, "Conservation value of roadside prairie restoration to butterfly communities," *Conservation Biology*, 15(2), 2001, pp. 401-411.
  6. K. J. Gaston and M. A. O'Neill, "Automated species identification: Why not?," *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(1444), 2004, pp. 655-667.
  7. J. G. A. Barbedo, "Using digital image processing for counting whiteflies on soybean leaves," *Journal of Asia-Pacific Entomology*, 17(4), 2014, pp. 685-694.
  8. H. P. Yang, C. S. Ma, H. Wen, Q. B. Zhan, and X. L. Wang, "A tool for developing an automatic insect identification system based on wing outlines," *Scientific Reports*, 5, 2015, pp. 1-11.
  9. W. McKinney, *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. California: O'Reilly Media, 2012.
  10. V. D. Walt, J. L. S. Stefan, J. N. Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu. "Scikit-image: Image processing in Python," *PeerJ*, 2, 2014, pp. 1-18.
  11. A. Bakhshipour and A. Jafari, "Evaluation of support vector machine and artificial neural networks in weed detection using shape features," *Computers and Electronics in Agriculture*, 145, 2018, pp. 153-160.
  12. S. Krig, "Image pre-processing," in *Computer Vision Metrics*, Cham: Springer, 2014, pp. 35-74.
  13. G. Seif, I'll tell you why Deep Learning is so popular and in demand. Available: <https://medium.com/swlh/ill-tell-you-why-deep-learning-is-so-popular-and-in-demand-5aca72628780>.