International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 14, No.1 (Jun-24)

http://dx.doi.org/10.12785/ijcds/1501125

Butterfly Image Identification Using Multilevel Thresholding Segmentasi and Convolution Neural Network Classification with Alexnet Architecture

Abdul Fadlil[1], Ainin Maftukhah[2], Sunardi[3] and Tole Sutikno[4]

[1,3,4]*Department of Electrical Engineering, Universitas Ahmad Dahlan, Yogyakarta, Indonesia*
[2]*Master Program of Informatics, Universitas Ahmad Dahlan, Yogyakarta, Indonesia*

Received 31 Jan. 2024, Revised 4 Apr. 2024, Accepted 6 Apr. 2024, Published 1 Jun. 2024

Abstract: Butterflies form a large group called Lepidoptera. Butterflies play an important role in the ecosystem so the lack of knowledge about butterfly species is a problem. Since butterflies are a natural phenomenon and can serve as an educational tool, knowledge about butterflies is an important component of education. The data used totaled 419 butterfly images which were divided into two, namely training data and testing data. First, the dataset is input, and then the dataset is preprocessed such as resizing, converting RGB to grayscale, and segmentation. The output of the preprocessing dataset is classified using CNN with AlexNet architecture. The results of the Alexnet architecture classification stage include ReLu (Convolution, Batch Normalisation, Max Pooling), Flatten, and Danse. After the Alexnet CNN training process is complete, the output data is evaluated using the calculation of Accuracy, Precision, and recall. The final result of the data is classified according to the species, the model without segmentation is able to classify the image with high accuracy, while using multilevel threshold segmentation cannot classify the image with high accuracy. The test results show that the model without segmentation has 83% accuracy, while the model with multilevel threshold segmentation only achieves 62% accuracy. The test results show that the combination of multilevel thresholding segmentation and AlexNet architecture creates a classification model that is less accurate in recognizing butterfly species. Comparing these test results, it can be concluded that the model without segmentation tends to be better at classifying information than the model using multilevel threshold segmentation.

Keywords: Convolution Neural Network, Butterfly, Segmentation, Multilevel, Alexnet

1. INTRODUCTION

Butterflies are a large group called Lepidoptera. Butterflies play an important role in the ecosystem, so the lack of knowledge about butterfly species is a problem. Other threats arrive from pests, diseases, predators, and weather transformation [1] [2] [3] [4]. This research is to identify butterfly species because their populations are declining. Knowledge about butterflies is part of education because butterflies are part of the environment that can be used as learning objects. Butterflies are one of the species that have an important role in the ecosystem to be considered and studied in research. The process of identifying patterns and classifying information is an interesting discussion from time to time. Pattern recognition provides solutions to problems related to recognition or classification, such as speech recognition, document reading classification, handwriting personality classification, and batik pattern recognition.

Introduction to butterflies has the most color schemes and shapes. Computers can be used as a medium to recognize and identify butterfly species. The image segmentation stage, feature extraction stage, classification stage, and outcomes are the stages that the computer implements. The stage of picture segmentation is employed to distinguish between backgrounds and objects. Color feature extraction is used in the feature extraction stage. Convolutional Neural Networks (CNNs) are used in the classification stage to compare feature extraction with training and testing datasets of the butterfly variety. The resulting stage will provide an accuracy value or the results of the previous stages.

One of the deep learning methods capable of identifying objects in images is CNN. CNN features are considered the best approach in object detection and recognition [5] [6]. CNN is one of the artificial intelligence algorithms that has been widely used to process images and assign weights and biases that can be learned on certain aspects of the image that can distinguish one image from another [7]. CNN

E-mail address: fadlil@mti.uad.ac.id, 2207048003@webmail.uad.ac.id, sunardi@mti.uad.ac.id, tole@ee.uad.ac.id

has a multi-layer arrangement by pooling layers including fully connected layers. CNN layers organize neurons so that they have three dimensionsi [8] [9]. There are several CNN architectures VGG-16, VGG-19, GoogLeNet, ResNet, and AlexNet that have been applied to image classification without involving or using the segmentation process [10].

This research lacks an understanding of the effectiveness of multilevel thresholding segmentation in the context of butterfly image classification. Although segmentation has become an important part of image pre-processing, most notably in extracting meaningful features for further analysis, few studies have specifically evaluated its effect on butterfly image classification. Understanding the role of segmentation in butterfly image pre-processing can share valuable knowledge on how to improve the accuracy and efficiency of image classification. Expanding the description of the effectiveness of multilevel thresholding segmentation in butterfly image classification can contribute significantly to filling the knowledge gap in the literature and broadening the description of image pre-processing techniques. Therefore, this research has significant value in enriching the field of pattern recognition and image processing, and its implementation in the fields of nature conservation and learning.

This research conducted image segmentation with Multilevel Thresholding to separate objects and backgrounds in butterfly images. The type of butterfly can be determined based on the name of the species using CNN classification. Multilevel Thresholding segmentation method to get object separation with background and CNN classification to get the best accuracy results.

This research lacks an understanding of the effectiveness of multilevel thresholding segmentation in the context of butterfly image classification. Although segmentation has become an important part of image pre-processing, most notably in extracting meaningful features for further analysis, few studies have specifically evaluated its effect on butterfly image classification. Understanding the role of segmentation in butterfly image pre-processing can share valuable knowledge on how to improve the accuracy and efficiency of image classification. Expanding the description of the effectiveness of multilevel thresholding segmentation in butterfly image classification can contribute significantly to filling the knowledge gap in the literature and broadening the description of image pre-processing techniques. Therefore, this research has significant value in enriching the field of pattern recognition and image processing, and its implementation in the fields of nature conservation and learning.

2. LITERATURE REVIEW

There are several previous studies related to the objects and methods in this study.

Marshal [11] CNN VGG16 and Alexnet models are used in classifying mango images. The classification results from CNN VGG16 get a high accuracy of 92.50%, while the use of Alexnet gets an accuracy of 79.64%.

Sowjanya and Injeti [12] utilizing a multilevel thresholding picture optimization technique. The suggested method works better than other algorithms, according to the results. Multilevel Thresholding segmentation using the Harmony Search Algorithm was researched by Srikanth and Bikshalu. Images of the cameraman, brain, Lena, and other objects were used as objects. The comparison illustrates why the suggested method's outcomes outperform those of the histogram-based approach.

Dave Jonathan et al [13] compared the location of polyps using CNN with RetinaNet architecture. The object used is a large uss image. The best model results without data extension with a value of 0.8415 and augmentation data with a matrix value of Ap25 = 0.9308.

Ibrahim et al [14] performed classification using the CNN method. The best accuracy result is 97.5% using VGGNET19 architecture. The difference from previous research is in the method and object.

Akmal Hariz [15] used the Convolutional Neural Network approach with MobileNet architecture to investigate human activity detection based on camera captures. The tool utilized is random rotation data augmentation applied to video. An ideal model with hyperparameters of 20 epochs, early termination with patience 10, learning speed 0.0001, stack size 16, and dense layer 5 was generated by the test results. After testing the confusion matrix and running the model with cross-validation, the F1 final result performance was 84.52%. Using authentic COPD CT-scan images.

Zhao [16] performed a multilevel threshold image segmentation study for chronic obstructive pulmonary disease using the diffusion association lender mold algorithm and entropy renyi. The experimental findings of the valuation by image quality matrix demonstrate the superior performance of the used algorithms. These findings can aid medical professionals in the qualitative and quantitative analysis of the lesion network, hence enhancing the diagnostic precision of the network.

Amalia [17] used the CNN approach with Alexnet architecture to classify brain tumor disease in MRI images. With an image size of 224 x 224 pixels, a training data ratio of 80%, validation data of 10%, and test data of 10% using the Adam optimizer, learning rate of 0.0001 utilizing batch size 8 and epoch value 50, the results obtained are the optimal parameters that determine system performance. The precision of 98.84%, recall of 97.65%, precision value of 97.65%, loss of 0.1616, and F1 score of 97.65% are the top results.

Benign and malignant skin cancer kinds were categorized by Saputra [18] using the Alexnet architectural model. The Adaptive Moment Estimation (Adam) optimization function and the AlexNet architecture are used in the construction of the model. According to the findings, the AlexNet model using the Adam Optimizer has an 81.26% classification accuracy for cancer types.

Similarities from previous research both use the Multi-level Thresholding method and butterfly objects. The similarities use CNN and Multilevel Thresholding methods.

3. RESEARCH METHODS

Tools and some materials used to support the research can be seen in Table I and Table II.

TABLE I. Research Tools

No	Device	Usability
1	Laptop Lenovo Ideapad Slim 3	Workstation to complete the research
2	Intel® Core™ i5-1035G1 CPU @1.00GHz 1.19 GHz	To perform turbo boost
3	Windows 11 Version 64 bit	Workstation operating system
4	Mouse and Keyboard	Tools for typing and hovering
5	Google Colab	Software Tools
6	Web Browser 64 bit	Software Tools

TABLE II. Supporting Materials

No	Research Materials	Number of Types	Total Data	Usability
1	Image Test Dataset	4 Types	41	Butterfly image materials for research
2	Image Train Dataset	4 Types	378	Butterfly image materials for research

The flow of research conducted from start to finish is shown in Figure 1.

The research flow shown in Figure 1 shows that the first step is to input the dataset taken from the Kaggle Website, and then read all the datasets that have been input. After that, data preprocessing is carried out such as Resizing, image conversion, and segmentation. The results of dataset preprocessing are divided into two, namely training data and testing data. Training data is done by processing according to the AlexNet architectural CNN model which includes Convolution, ReLu, Flatten, and Danse. The results of the training model, then testing data is carried out and the testing performance is seen. After the testing performance is processed and the accuracy results are obtained, the last stage of the model is completed by displaying the accuracy value.

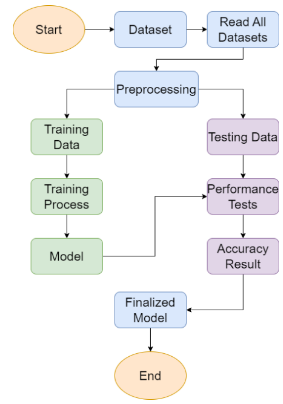

Figure 1. Research Flow

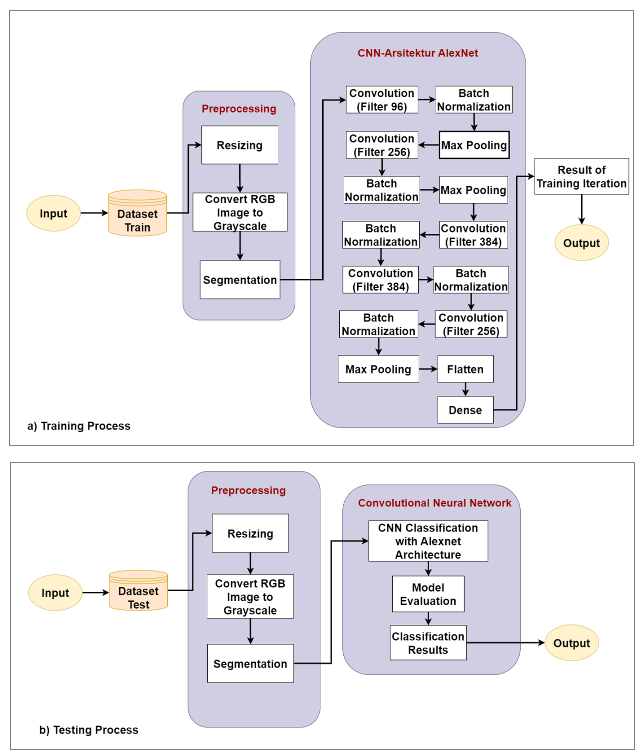

Figure 2. Training and Testing Process

The training and testing process shown in Figure 2(a) starts with the input of the training dataset. Output data is preprocessed including resize, image conversion, and segmentation. The preprocessing results are convolved with filter 96. The convolutional output is subjected to batch normalization, i.e. ReLU, after which max-pooling is applied to the batch normalization result. The convolution, batch normalization, and max-pooling steps are performed 5 times. Then the last max-pooling result is flattened, and the flattened result is dense twice. The dense result is made with an iteration training model to get the best result. Figure 2(b) starts with inputting the training dataset. Output data is preprocessed including resize, image conversion, and segmentation. The preprocessing results are classified using CNN Alexnet architecture. Furthermore, the results of the evaluation model are carried out to obtain the classification results at the last stage.

A. Datasets

About 419 butterfly photos from the Kaggle website (https://www.kaggle.com/datasets/gposenka/butterfly-images40-species) are included in the dataset. The photos represent a variety of species, including common, rare, and vulnerable ones. A species is a collection of creatures with similar traits, and the term "species" refers to individuals rather than classes or groupings. The lowest rung of toxins are species. Butterfly attributes are physical characteristics, such as color, pattern, wing form, and size, that can be seen or recognized in different species of butterflies. Building a recognition model with a big range of butterfly species will yield more valid findings when using a dataset with a sufficient number of data. Figure 3 is the butterfly image, the category dataset can be seen in Figure 4, and Figure 5 is the entire image dataset.

Figure 3. Butterfly Image

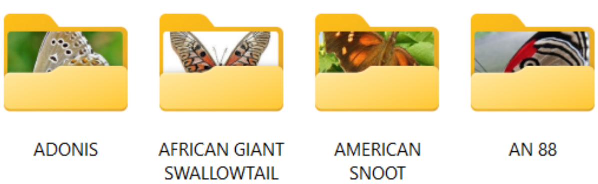

Figure 4. Datasets Categories

Figure 5. Image Datasets

B. Preprocessing

Preprocessing is a step used to prepare image data before further processing [19]. Pre-processing methods to reduce these differences rely on contrast enhancement and data focusing [20],[21]. The preprocessing stage is where all unequal image sizes are resized to 150×150 pixels since CNNs accept the same size. After resizing, the RGB images are converted to grayscale for model processing and training [22]. A two-dimensional matrix can be used as a digital representation of a greyscale image. The table's elements each display the image's intensity (or grey level) at the associated coordinate point. An image with an 8-bit representation has 2^8 or 256 grey scale levels, often ranging from 0 to 255, where 0 denotes the lightest intensity and 255 is the darkest. All the elements in the preceding table are referred to as image elements, or frequently, pixels. Each pixel's intensity can be altered to alter the image's overall depiction. A specific pixel in an image represented as a M x N matrix has a certain intensity. The shadow elements have distinct coordinates (x,y) and positions (i,j). Pixel counting starts from the top left corner, while the x and y coordinates are at the bottom left corner [23]. To assess the influence of the pre-processing on the classification outcomes of each CNN model, the images were normalized. How to use the Tensorflow Preprocessing Library to preprocess data [24].

C. Segmentation

The process of dividing a picture into distinct areas is called image segmentation. Various segmentation techniques rely on searching inside or outside of a region. Edge detection can be used to locate region borders. The characteristics of the pixels that comprise an area define its interior. One well-liked segmentation approach is the region-based thresholding method [25].

D. Thresholding

In image processing, thresholding procedures are frequently employed for compression, segmentation, and enhancement. Thresholding shows how important a specific level is, in terms of magnitude. In image processing, thresholding is applied in a variety of ways. The first method involves creating a binary image by applying grayscale image thresholding. Setting all gray levels with values less than or equal to T to 0 and setting the remaining gray levels to 1 establishes the threshold $T > 0$ [25]. Multilevel thresholding is a method that uses a set of thresholds to classify the pixels of an image [26]. Color images are divided into foreground and background using more than two thresholds (tri or quad-level) to separate the three components R, G, and B. This method provides good specificity [27]. Several threshold points divide the image into distinct classes, giving target region analysis possibilities.

$$O_1(x, y) = \{i(x, y) \in I \,|\, 0 \leq I(x, y) \leq m_1 - 1\} \tag{1}$$

$$O_2(x, y) = \{i(x, y) \in I \,|\, m1 \leq I(x, y) \leq m_2 - 1\} \tag{2}$$

$$O_i(x, y) = \{i(x, y) \in I \,|\, m_i \leq I(x, y) \leq m_i - 1\} \tag{3}$$

$$O_r(x, y) = \{i(x, y) \in I \,|\, m_r \leq I(x, y) \leq L - 1\} \tag{4}$$

Where t_1, t_2, t_3, t_4,...t_i,...t_r represent different thresholds. Based on the intensity value, distinct groups of gray pixels are assigned, and within the given range, each group has a separate set of pixel values [28].

E. Classification CNN-Architecture Alexnet

Classification is a task that involves assigning class labels to input models. The class label indicates the class of a particular set of classes. Classification is performed using a model obtained using a supervised learning procedure. Depending on the type of learning used there are two classifications, one using supervised learning and one using unsupervised learning [29]. Finding an image's traits, structures, or patterns and using them to place it in a specific class is the aim of categorization. Human observers frequently complete specific classification tasks using images and visual stimuli with a high degree of accuracy [30]. CNN is a multi-layered deep education method that extends Artificial Neural Network (ANN). CNNs are processed by network arrays and create outputs of specific classes. Each level carries out training and the output of each level is used as input for the next level. Initially, CNNs create simple features such as color, brightness and edges, whereas later levels create more environmental features. CNNs consist of three fully connected layers and five convolutional layers. The first AlexNet layer is used as an input filter image for width, height and depth fields (red, green, blue) with dimensions of 227 × 227 × 3. The last composite layer combines fully 1000 composite layers, and the others. The layers serve as dividers between features. AlexNet creates a 4096-dimensional feature vector for every input image, with a hidden layer that activates immediately before the output layer. With 60 million parameters and 650,000 neurons, AlexNet is a massive architecture. A total of 1.2 million training photos and 150,000 test images from the ImageNet dataset were used to test the model. The model reduces the oversampling problem very effectively by removing and adding data [31].

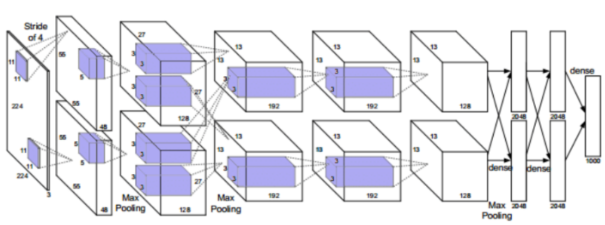

Figure 6. Alexnet Architecture

Figure 6 shows the architecture of AlexNet starting with a pre-processed input image with a result size of 150x150, then performing filtering or kerneling. kernel results using convolution which is pooled. Convolution and max-pooling are performed five times. Each is considered to have characteristics due to convolution. Then a flattening process is performed to convert the image into a one-dimensional matrix. Then it is fully connected three times. As a final step, the image can be classified.

F. Model Evaluation

Evaluation of the Alexnet model is used to determine its effectiveness. This evaluation uses accuracy, recall, precision, and F1-score matrices which can be calculated using the following formula:

$$Accuracy(\%) = \frac{TP + TN}{TotalTestingData} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$Precission = \frac{TP}{TP + FP} \quad (7)$$

$$F1 - Score = 2 * \frac{recall * precission}{recall + precission} \quad (8)$$

4. RESULTS AND DISCUSSION

The results of this study use Google Colab tools using 419 butterfly image data from the Kaggle website. Image data is divided into 90% training data and 10% testing data. The first process is carried out before classification to get the best results with preprocessing and segmentation thresholding stages that have been tested on butterfly image datasets. The results of the study discuss the analysis of the comparison of butterfly image classification with Alexnet architecture using multilevel threshold segmentation and without segmentation.

Figure 7 is the original image that will be processed for preprocessing. at the preprocessing stage, the technique of changing the size and changing the RGB color to grayscale is carried out. The result of 150x150 resizing can be seen in Figure 8 because Alexnet only accepts the same image size. The conversion of RGB to grayscale to improve the efficiency of analysis can be seen in Figure 9. After the RGB to the grayscale stage, the segmentation stage is carried out using multilevel thresholding segmentation techniques. Figure 10 (a) is the initial image, the segmentation results can be seen in Figure 10 (c) and without segmentation can be seen in Figure 10 (b).

The results of preprocessing are used in classification using the CNN method with Alexnet architecture. Next, create a graph using the sequential model. It first adds a convolutional layer with a total of 96 11×11 filters and an activation function as a rectified linear unit (ReLU). Next, apply dropout. Before inputting the output generated at the dense layer, it is necessary to flatten the variables to fit the shape of the dense layer input. The output produced by the dense layer is set according to the dropout, the resulting output is then used as input to the last dense layer for classification. Softmax serves to convert the result into something that is interpreted in the form of probability. Dropout was applied after the Dense layer with a dropout level of 0. 5. Dropout is a regularisation technique that randomly ignores some units during the training process, helping to reduce the model's dependence on certain features and avoid

Figure 7. Original image of a butterfly

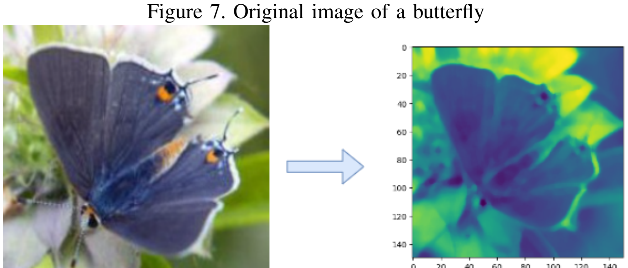

Figure 8. Resizing Result

overfitting. Next, batch normalization is applied after the partial convolution layer. Batch normalization helps speed up the training process by reducing the internal covariate shift and helps avoid overfitting by normalizing the input to each layer. Finally, the learning rate is reduced using the ReduceLROnPlateau callback. This reduces the learning rate when progress in training stops, which can help avoid overfitting and improve training stability.

After 20 epochs were run, the model with no segmentation and using multilevel thresholding segmentation on the butterfly image data established satisfactory results for the simple model. Table III which includes the training accuracy and loss columns are the results of the epochs run, training the model using preprocessing without segmentation. Although these results are quite good, there are still some errors in butterfly image classification, especially in images with high complexity. The model training used image segmentation processed with a multilevel thresholding technique. This technique is used to separate the object from the background based on the brightness of the pixels, with a multilevel thresholding segmentation technique, the model cannot classify the image accurately. The epoch

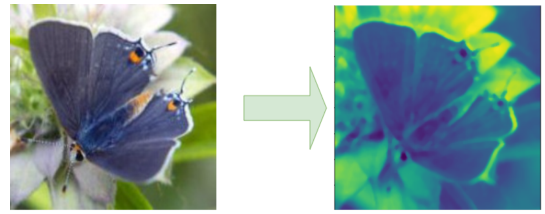

Figure 9. RGB to Grayscale Result

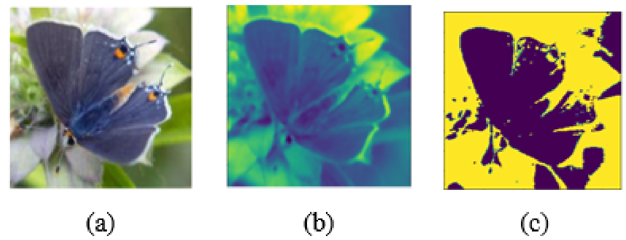

Figure 10. (a) Input Image (b) result without segmentation (c) thresholding segmentation result

TABLE III. Training Epoch Process

Epoch	No Segmentation		Threshold Multilevel Segmentation	
	Accuracy	Training Loss	Accuracy	Training Loss
1	0.49	9.10	0.44	14.76
2	0.73	2.36	0.65	4.37
3	0.84	1.63	0.71	2.33
4	0.89	0.65	0.77	1.25
5	0.87	0.48	0.79	0.78
6	0.89	0.37	0.86	0.48
7	0.89	0.52	0.87	0.42
8	0.89	0.64	0.87	0.33
9	0.88	0.32	0.88	0.31
10	0.91	0.27	0.90	0.27
11	0.91	0.23	0.90	0.32
12	0.91	0.28	0.89	0.26
13	0.92	0.21	0.89	0.26
14	0.94	0.19	0.91	0.24
15	0.92	0.22	0.90	0.29
16	0.94	0.19	0.90	0.28
17	0.91	0.24	0.90	0.25
18	0.94	0.16	0.89	0.28
19	0.92	0.20	0.91	0.25
20	0.94	0.14	0.90	0.25

process results displayed from the training loss and training accuracy function values help evaluate the effectiveness of the model in classifying the train data, as well as improving each epoch process during training.

Figure 11 and Figure 12 can be analyzed that the accuracy and loss graphs are generated by testing using the Alexnet architecture without segmentation and using multilevel thresholding segmentation. The accuracy and loss graphs are used to understand the performance of the model

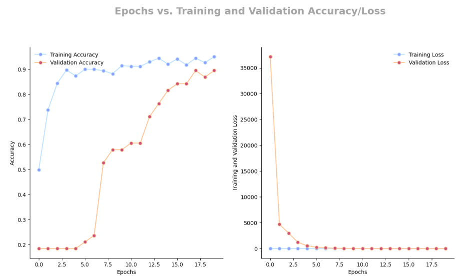

Figure 11. Accuracy Graph Without Segmentation

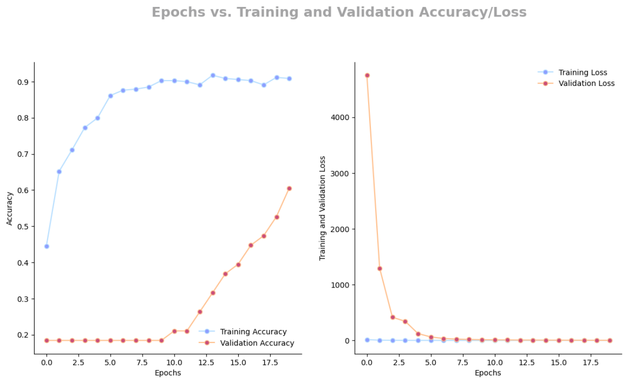

Figure 12. Accuracy Graph with Multilevel Thresholding Segmentation

in classifying butterfly images. It can be seen from the accuracy graph that the accuracy of the model increases with the number of epochs performed. At the beginning of training, the accuracy of the model tends to be low, but over time the accuracy of the model increases significantly. This shows that the model is gradually practicing to classify the image better. At the same time, in the loss graph, it can be seen that the loss of the model decreases as the number of execution epochs increases. Loss is a measure of how far the model's prediction is from the true value, so the lower the loss value, the better the quality of the model's prediction. This study shows that the model succeeds in reducing the loss significantly over time.

Figure 13 shows the output results of the training process.

	precision	recall	f1-score	support
0	0.91	0.91	0.91	11
1	0.86	0.80	0.83	15
2	0.86	0.75	0.80	8
3	0.80	1.00	0.89	8
accuracy			0.86	42
macro avg	0.86	0.86	0.86	42
weighted avg	0.86	0.86	0.86	42

Figure 13. Alexnet Training Process of Multilevel Thresholding Segmentation

The results are shown in Figure 14. The results of the confusion matrix show that 5 Adonis images could be correctly predicted from the original data. 5 images featured American snoot however, 1 photo with Adonis was recognized as an African giant swallowtail. 7 images of African giant swallowtails were judged to be good out of the raw data. AN88 is 1 of the 7 images that African Snoot should be used for. 7 of the top-rated photos of African snoot from the original information. 7 images should be considered with African giant Swallowtail. 1 image should be considered with African giant Swallowtail. 7 images of AN88 are accepted as original data. 1 image should be considered American Snoot.

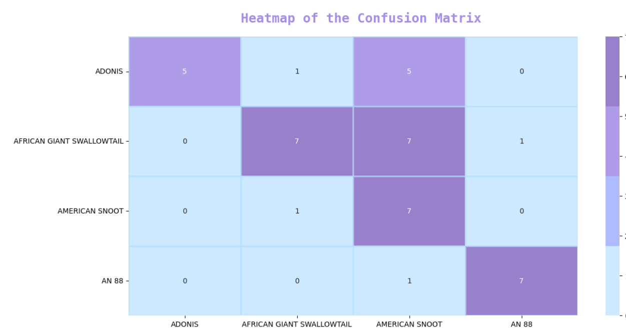

Figure 14. Alexnet's Confusion Matrix of Multilevel Thresholding Segmentation Image

Without segmentation, the output results of the training process are displayed in Figure 15.

	precision	recall	f1-score	support
0	0.91	0.91	0.91	11
1	0.86	0.80	0.83	15
2	0.86	0.75	0.80	8
3	0.80	1.00	0.89	8
accuracy			0.86	42
macro avg	0.86	0.86	0.86	42
weighted avg	0.86	0.86	0.86	42

Figure 15. Alexnet Training without Segmentation

Figure 16 presents the findings. Based on the original data, 10 images of Adonis can be accurately anticipated, according to the confusion matrix result. One picture featuring Adonis was thought to be an African giant swallowtail. Out of the original data, 12 African giant swallowtail images were deemed to be good. Two AN88 photos and one American Snoot image should be taken into consideration. Six of the original information's American Snoot photos were deemed to be of high quality. One Adonis image and one African giant swallowtail image are to be taken into consideration. As raw data, 8 AN88 pictures were approved.

Analyzing from Table IV the test results of the testing data, 10% final accuracy is obtained with the Alexnet architecture without segmentation and using multilevel threshold segmentation. Table IV shows that the model without

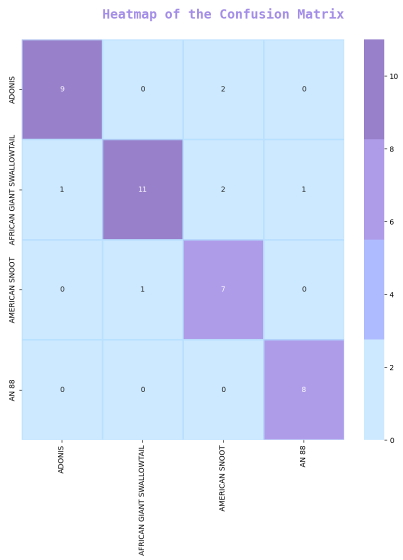

Figure 16. Alexnet Confusion Matrix of Unsegmented Image

TABLE IV. Butterfly Image Classification Test Results

Class	Amount of Data	No Segmentation		Threshold Multilevel Segmentation	
		True	False	True	False
Adonis	11	0.49	9.10	0.44	14.76
African Giant Swallowtail	15	0.73	2.36	0.65	4.37
American Snoot	8	0.84	1.63	0.71	2.33
AN88	8	0.89	0.65	0.77	1.25
Accuracy (%)		83%		62%	

segmentation cannot classify images with high accuracy while using multilevel threshold segmentation cannot classify images with high accuracy. The test results show that the model without segmentation has an accuracy of 83%, while the model with multilevel threshold segmentation only achieves an accuracy of 62%. The model without segmentation achieves greater accuracy but must consider the better classification quality obtained from the model with segmentation. Although the accuracy of the model without segmentation is greater, it may be more prone to overfitting or not be able to generalize well to data that has not been seen before. Although multilevel threshold segmentation performed lower in this problem, this does not mean that segmentation in general is not useful. Segmentation methods remain a significant component of image processing for various implementations, and their usefulness can vary greatly depending on the context and dataset characteristics.

This accuracy comparison can be caused by several factors, such as a model without segmentation can more easily identify and distinguish important features in the image, so that the model can produce more accurate predictions. At the same time, models with multilevel thresholding segmentation may have difficulty in identifying and distinguishing important features, resulting in less accurate predictions. The use of multilevel thresholding segmentation may cause the loss of important data in the image, making it difficult for the model to classify the information correctly. Meanwhile, models without segmentation can utilize the complete data in the image to create more accurate predictions.

The benefit of the AlexNet design is that the model can automatically extract hierarchical features because convolutional neural networks are used in it. AlexNet has demonstrated efficacy in resolving intricate picture categorization assignments. The requirement for vast volumes of data for ideal training, however, and increased computing complexity in comparison to simpler models are possible drawbacks.

Table V lists the accuracy results found in previous research using the Alexnet architecture CNN method with butterfly images and other datasets. Based on Table V, it can be seen that the highest accuracy of the previous research reached 98.84%. These results are higher than the accuracy obtained in this study, using AlexNet architecture without segmentation with brain tumour MRI image datasets.

Table V states that starting in 2023, research related to the use of the CNN Alexnet method with Manggo image datasets has reached 79.64% accuracy. In the following year, the use of the RetinaNet architectural CNN method with the colon image dataset achieved an accuracy of 0.9308. In 2023 the use of the Alexnet method obtained an accuracy of 81.26%. This research with the Alexnet method without segmentation obtained higher accuracy than related research in 2019 with butterfly image datasets. This factor shows that the accuracy results obtained in this study are indeed low. However, there are several factors such as differences in image types and datasets used that can affect classification performance. In addition, differences in data preprocessing such as segmentation also have a significant impact on classification accuracy.

This study shows that using the Alexnet method with segmentation resulted in lower accuracy than previous studies using butterfly image datasets. However, this study highlights the importance of using proper segmentation in improving image classification accuracy. Although segmentation is considered to play an important role, the success of segmentation is highly dependent on the suitability of the technique and parameters used for the particular image data.

Although this study did not achieve the same level of accuracy as previous studies, it highlights the importance of using segmentation techniques to improve input data quality and classification performance. This research provides an opportunity to further develop the combined use of segmentation and image processing techniques with machine learning techniques to significantly improve butterfly image recognition accuracy.

TABLE V. Previous Research Comparison

Author	Year	Methods	Dataset	Accuracy Result
Marshal	2023	CNN Alexnet	400	79.64%
Dave Jonathan	2022	CNN arsitektur RetinaNet	1448	0.9308
Amalia	2022	Alexnet	60	98.84%
Saputra	2023	Alexnet	25331	81.26%
Our study	2024	Alexnet without segmentation	419	83%
Our Study	2024	Alexnet with multilevel thresholding segmentation	419	62%

Previous research tends to use approaches that only rely on CNN, KNN or single segmentation methods for butterfly image classification. Not only that, this research combines the multilevel thresholding segmentation method with CNN to increase the accuracy of image identification. This approach introduces innovation by using the advantages of both methods together. While some studies achieve a high level of accuracy, others may experience challenges in overcoming the diversity in the image. In this study, it was found that the use of multilevel thresholding segmentation displays the inefficiency of this approach in handling complex image variations. This approach makes a meaningful contribution to the development of a more efficient and accurate butterfly image identification method. This comparative analysis helps highlight the novelty as well as the advantages of this approach. This experiment provides deeper insight into the impact and significance of the research findings on butterfly image identification using multi-level threshold segmentation and CNN classification with AlexNet architecture. The implications of the findings are significant in several ways, particularly in the area of conveying knowledge about butterfly species. The ability to identify and classify butterfly status efficiently and accurately. In addition, this finding has implications in the fields of technology and artificial intelligence, specifically the use of image analysis and machine learning techniques to detect and classify butterfly species. The importance of this finding lies in its potential to improve the understanding and management of butterfly species through image analysis and to demonstrate how technology and artificial intelligence can be used to classify butterflies.

5. Conclusions and Future Work

In this research, butterfly image identification using multilevel thresholding segmentation and Convolutional Neural Network (CNN) classification with AlexNet architecture is conducted. The limitations of this research include the use of 4 types of butterflies as datasets with the CNN Alexnet model using multilevel thresholding segmentation techniques and without segmentation. The test results show that the combination of multilevel thresholding segmentation and AlexNet architecture creates a classification model that is less accurate in recognizing butterfly species. Comparing the results of this test, it can be concluded that the model without segmentation tends to be better at classifying image data by getting 83% accuracy, compared to the model using multilevel thresholding segmentation getting 62% accuracy. Classification of butterfly images using AlexNet architecture, the approach without segmentation produces better results than using multilevel thresholding segmentation. Although segmentation aims to improve the quality of the input data in this problem, the use of segmentation reduces the classification accuracy. However, it should be noted that these results may vary depending on the type of dataset and the complexity of the problem. Therefore, choosing the right segmentation method is very important in the development of CNN classification models. Further research should be attempted to master the factors that affect segmentation performance in this particular context and explore possible alternatives and adjustments to improve the results.

References

[1] A. Rajab and D. Asriady, "Species diversity of papilionidae butterflies," 2015.

[2] A. Maftukhah, A. Fadlil, and S. Sunardi, "Butterfly image segmentation using multilevel thresholding method," *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, vol. 7, no. 2, pp. 545–554, 2023.

[3] A. D. Gonggoli, S. Sari, H. Oktofiani, N. Santika, R. Herlina, T. Agatha, and Y. E. Gunawan, "Identification of butterfly species (lepidoptera) at palangka raya university," *Bioeksperimen: Jurnal Penelitian Biologi*, vol. 7, no. 1, pp. 16–20, 2021.

[4] A. Maftukhah, A. Fadlil, and S. Sunardi, "Butterfly image classification using convolution neural network with alexnet architecture," *JURNAL INFOTEL*, vol. 16, no. 1, 2024.

[5] H. Hendriyana, Y. H. Maulana *et al.*, "Identification of types of wood using convolutional neural network with mobilenet architecture," *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, vol. 4, no. 1, pp. 70–76, 2020.

[6] Y. S. Hariyani, S. Hadiyoso, and T. S. Siadari, "Covid-19 disease detection based on x-ray image using deep residual network," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 8, no. 2, p. 443, 2020.

[7] M. I. B. Kerthyayana and B. V. Jenny, "Comparison study of image augmentation on modified cnn architecture for indonesian lasem-batik's motifs," *ICIC Express Letters*, vol. 17, no. 04, p. 409, 2023.

[8] M. D. Chaniago, A. A. Sugiharto, Q. D. Khatulistiwa, Z. Sari, A. Eko *et al.*, "Covid-19 detection using convolutional neural

networks (cnn) classification algorithm," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 2, pp. 190–197, 2022.

[9] F. Han, J. Yao, H. Zhu, and C. Wang, "Marine organism detection and classification from underwater vision based on the deep cnn method," *Mathematical Problems in Engineering*, vol. 2020, pp. 1–11, 2020.

[10] H. Akbar and S. Sandfreni, "Cervical cancer classification using the alexnet convolutional neural network model," *JIKO (Jurnal Informatika dan Komputer)*, vol. 4, no. 1, pp. 44–51, 2021.

[11] A. L. Marshal and A. N. F. A. N. Fajar, "Image classification of mangoes using cnn vgg16 and alexnet," *Journal of Social Science (JoSS)*, vol. 2, no. 8, pp. 694–703, 2023.

[12] K. Sowjanya and S. K. Injeti, "Investigation of butterfly optimization and gases brownian motion optimization algorithms for optimal multilevel image thresholding," *Expert Systems with Applications*, vol. 182, p. 115286, 2021.

[13] R. D. JONATHAN, M. J. HASUGIAN, and E. M. SARTIKA, "Comparison of polyp location detection on colonoscopy image using cnn with retinanet architecture," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 10, no. 4, p. 946, 2022.

[14] N. Ibrahim, G. A. LESTARY, F. S. HANAFI, K. SALEH, N. K. C. PRATIWI, M. S. HAQ, and A. I. MASTUR, "Classification of tea leaf maturity level using convolutional neural network method," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 10, no. 1, p. 162, 2022.

[15] F. A. Hariz, I. N. Yulita, and I. Suryana, "Human activity recognition berdasarkan tangkapan webcam menggunakan metode convolutional neural network (cnn) dengan arsitektur mobilenet," *JITSI: Jurnal Ilmiah Teknologi Sistem Informasi*, vol. 3, no. 4, pp. 103–115, 2022.

[16] S. Zhao, P. Wang, A. A. Heidari, H. Chen, H. Turabieh, M. Mafarja, and C. Li, "Multilevel threshold image segmentation with diffusion association slime mould algorithm and renyi's entropy for chronic obstructive pulmonary disease," *Computers in Biology and Medicine*, vol. 134, p. 104427, 2021.

[17] K. Amalia, R. Magladena, and S. Saidah, "Classification of brain tumour disease on mri image using cnn method with alexnet architecture," *eProceedings of Engineering*, vol. 9, no. 6, 2023.

[18] T. Saputra and M. E. Al Rivan, "Classification of benign and malignant skin cancer types using alexnet architecture model," in *MDP Student Conference*, vol. 2, no. 1, 2023, pp. 158–165.

[19] L. Hakim, H. R. Rahmanto, S. P. Kristanto, and D. Yusuf, "Image classification of banyuwangi batik motifs using convolutional neural network," *Jurnal Teknoinfo*, vol. 17, no. 1, pp. 203–211, 2023.

[20] L. F. Rodrigues, M. C. Naldi, and J. F. Mari, "Comparing convolutional neural networks and preprocessing techniques for hep-2 cell classification in immunofluorescence images," *Computers in biology and medicine*, vol. 116, p. 103542, 2020.

[21] L. C. de Faria, L. F. Rodrigues, and J. F. Mari, "Cell classification using handcrafted features and bag of visual words," in *Anais do XIV Workshop de Visão Computacional*, 2018, pp. 68–73.

[22] M. M. Khodier, S. M. Ahmed, and M. S. Sayed, "Complex pattern jacquard fabrics defect detection using convolutional neural networks and multispectral imaging," *IEEE Access*, vol. 10, pp. 10 653–10 660, 2022.

[23] E. Maria, Y. Yulianto, Y. P. Arinda, J. Jumiaty, and P. Nobel, "Digital image segmentation of leaf shape on plants at politani samarinda using thresholding method," *Jurnal Rekayasa Teknologi Informasi (JURTI)*, vol. 2, no. 1, pp. 37–46, 2018.

[24] R. Soekarta, N. Nurdjan, and A. Syah, "Classification of tomato plant diseases using the convolutional neural network (cnn) method," *Insect (Informatics and Security): Jurnal Teknik Informatika*, vol. 8, no. 2, pp. 143–151, 2023.

[25] D. Sundararajan, *Digital image processing: a signal processing and algorithmic approach*. Springer, 2017.

[26] X. Huang, L. Shen, C. Fan, S. Chen *et al.*, "Multilevel image thresholding using a fully informed cuckoo search algorithm," *arXiv preprint arXiv:2006.09987*, 2020.

[27] M. Abd Elaziz, S. Bhattacharyya, and S. Lu, "Swarm selection method for multilevel thresholding image segmentation," *Expert systems with Applications*, vol. 138, p. 112818, 2019.

[28] R. Kalyani, P. D. Sathya, and V. P. Sakthivel, "Multilevel thresholding for medical image segmentation using teaching-learning based optimization algorithm," *International Journal of Intelligent Engineering and Systems*, vol. 14, no. 2, pp. 11–21, 2021.

[29] M. N. Murty and D. V. S. Devi, *Introduction to Pattern Recognition and Machine Learning*. World Scientific, 2018, vol. 3, no. 1.

[30] C. Solomon and T. Breckon, *Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab*. Wiley, 2018.

[31] F. R. Mashrur, A. D. Roy, and D. K. Saha, "Automatic identification of arrhythmia from ecg using alexnet convolutional neural network," in *2019 4th international conference on electrical information and communication technology (EICT)*. IEEE, 2019, pp. 1–5.

Abdul Fadlil is a lecturer at the Department of Electrical Engineering, Ahmad Dahlan University, Yogyakarta. He has published 37 national and international research journals. His current research interests include Intelligent Systems and Signal Systems.

Ainin Maftukhah is an undergraduate at Ahmad Dahlan University in Informatics and earned her bachelor's degree in 2021. Currently, she is continuing her Masters in Informatics from Ahmad Dahlan University. Her current research interests include pattern recognition.

Sunardi is a lecturer at the Department of Electrical Engineering, Ahmad Dahlan University, Yogyakarta. He has published 89 research journals in national and international journals. His research interests include data, breakthrough information, and communication.

Tole Sutikno is an Associate Professor at the Department of Electrical Engineering, Ahmad Dahlan University, Indonesia, and is the Chair of the Embedded Systems and Power Electronics Research Group (ESPERG). He has published 89 research journals in national and international journals. His current research interests include power electronics, motor drives, renewable energy, FPGA, and intelligent control systems.