Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 5, No.3, October 2023, pp. 0230302-01 ~ 0230302-08
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v5i3.16854

A Good Performance of Convolutional Neural Network Based on AlexNet in Domestic Indonesian Car Types Classification

Ervira Aliya Nabila[1*], Christy Atika Sari[1], Eko Hari Rachmawanto[1], Mohamed Doheir[2]

[1]University of Dian Nuswantoro, Jl. Imam Bonjol 207, Semarang, Central Java 50131, Indonesia

[2]University Technical Malaysia Melacca, Ayer Keroh, Durian Tunggal, Melacca 75450, Malaysia

*erviranabila42@gmail.com

Abstract.

Classification of car vehicle types has been carried out using CNN. There are weaknesses in the CNN algorithm so that it can be continued in the research we propose. This study aims to improve the previous accuracy by using the Alexnet architecture. To improve the results of the data set used we use threshold and brightness adjustment and data augmentation techniques for Reflection, Rotation, and Translation. Sample images with a resolution of 227x227x3 totaling 840 images used to represent 8 class types of cars, including Avanza, Fortuner, Freed, Inova, Pajero, Terios, Xenia, and Xpander. Alexnet with 10 epochs consisting of a total of 760 iterations, and validation is carried out every 30 iterations, the test results show that the use of the "sgdm" optimization function achieves a training accuracy of 99.74%, while the use of the "adam" optimization function produces an accuracy of 96.85%. This experiment shows the model's ability to classify the types of trainers after a success rate of 100%.

Keywords: *Alexnet, Classification, Cars, CNN, Image Processing, Car*

(Received 2023-08-29, Accepted 2023-09-05, Available Online by 2023-09-30)

1. Introduction

With the development of technology, human life has been greatly helped by the presence of machine vision. Recognition of vehicle types has always been a hot topic in the field of object detection, especially cars [1]. In 2023, cars will become increasingly popular and have an important role in supporting the development of the automotive and technology industries. Apart from that, it is impossible to find information directly from one person to another to find out the type of car used. This type of car recognition system can help identify and analyze car usage trends [2] and consumer preferences which can help classify car types based on market demand and it is hoped that the results of this research can help car manufacturers plan their marketing and products. strategy development. Data taken from Gaikindo [3] regarding the top 10 retail car sales in 2023, Avanza and its sibling Veloz appeared as Toyota's best-selling vehicles at the 2023 Gaikindo Indonesia International Auto Show (GIIAS) which took place from 10-20 August 2023. These two cars topped sales with a total difference in orders of more than 1,000 units through Vehicle Purchase Orders (VPO) of 1,213 units (20.9%).

Proving that in Indonesia the majority of people choose Toyota for their cars [4]. With this data, a car model recognition system is needed to classify the types of car vehicles in Indonesia using digital image processing. The Convolutional Neural Network (CNN) algorithm is an important algorithm that helps many machine vision tasks in image classification, semantic segmentation, and object detection and is often used for car model recognition [5]. Previous research has utilized CNNs to recognize distinctive features in car images, thereby enabling more precise classification. Some approaches have used data augmentation techniques, such as rotation, reflection, and brightness adjustment, to increase variation in the training data set, while others have considered various CNN architectures, such as AlexNet. Nonetheless, there is still room for further development in improving the accuracy and efficiency of car type classification. In this context, this paper contributes by introducing a method that combines data augmentation techniques with the AlexNet architecture, which is expected to open up new opportunities to significantly improve the accuracy of car type classification [6].

Data augmentation is a technique commonly used in machine learning to increase variety in a training data set by making small variations on existing images, such as rotating, mirroring, or panning them [7]. This helps machine learning models become more robust and able to recognize a wider variety of patterns. This study investigates the use of the AlexNet architecture on the CNN network [8]. The main focus of this research is centered on two main problem aspects as follows:

a. Found a method of implementing data augmentation on the AlexNet architecture which aims to identify car models.
b. Make a comparison between the Alexnet architecture and the implementation of data augmentation techniques such as Reflection, Rotation, and Translation.

This research begins with the collection of a dataset that includes various types of cars from various points of view. This car-type dataset is then pre-processed to ensure image quality and consistency. CNN Alexnet is proposed as the main model for performing automatic feature extraction from car images. There are weaknesses in the CNN algorithm [9] that can be continued in the research we propose. This study aims to improve the previous accuracy by using the Alexnet architecture. The research that has been built focuses on the use of data augmentation techniques, such as reflection, rotation, and translation, along with image segmentation using thresholding, brightness adjustment, and CNN Alexnet. This approach can provide greater variation in the training dataset, which in turn can improve classification performance and help increase the accuracy of car type recognition and produce good accuracy. After this training was completed, it resulted in an accuracy of 99.74% and proved to help recognize the type of car.

2. Methods

2.1. Data Collection

The dataset used in this training is a private dataset taken in real time on roads around Semarang by taking pictures of passing vehicles. Car images with dimensions of 227x227 pixels total 840 data sets of car types which are categorized into two different data sets, namely training data and test data which includes 8 classes of car type images consisting of 108 Avanza, 100 Fortuner, 100 Freed, 125 Inova, 108 Pajero, 102 Terios, 103 Xenia, 100 Xpander. The training data is used to understand the model for each type of car that needs to be categorized, while the test data is used to evaluate the effectiveness of the proposed method in accurately identifying the desired car type category.

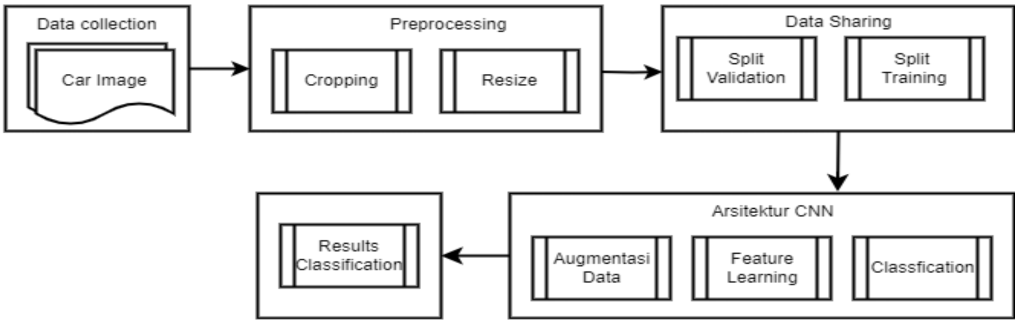

Figure 1. Our Proposed Method

2.2. Preprocessing

The initial data processing (preprocessing) of the vehicle image dataset is carried out to obtain a better image so that the results of the process to be carried out next have better results. The following steps are performed for pre-proccessing

1. Cropping the RGB image to adjust the center of the object so that the image on the car dataset can be more focused and uniform. This cutting process is done manually.
2. Resize the RGB image to 227x227 pixels. This is necessary so that the image of the car can be processed in practice. The car model dataset must be resized according to the input dimensions specified by the CNN AlexNet network, namely 227x227x3. Figure 3 shows the three RGB color channels namely Red (R), Green (G), and Blue (B). The image resizing process is done automatically using Matlab R2022a.

The preprocessing process of car image segmentation is carried out with 2 techniques, namely Thresholding brightness adjustment to help clarify objects [10] in the image so that the model can focus more on recognizing relevant features and classifying car types better. Divide the collected data into two parts: training data and test data[11]. Each vehicle is labeled to distinguish these two data, for each data is divided into 90% training data and 10% test data.

In car classification with CNN AlexNet, a rich and comprehensive feature extraction process occurs. This algorithm identifies complex visual patterns and statistically significant features that characterize car types. Feature extraction includes Mean, Variance, Kurtosis, Minimum Value, Maximum Value, Standard Deviation, Entropy, Skewness [12] as in (1) until (8). Calculations in finding the Mean value use the formula (1), where n is the number of elements in the dataset, x_i is the value of the to-i in the dataset. Calculations in finding the standard deviation value use the formula (2), where n is the number of elements in the dataset, x_i is the value of the to-i dalam dataset, m is the average value of the dataset. Calculations in finding the Kurtosis value use the formula (3), where n is the number of elements in the dataset and t, x_i is the value of the to-i dalam dataset, m is the average value of the dataset, std dataset standard deviation. Calculations in finding the Variance value use the formula (4), where n is the number of elements in the dataset, x_i is the value of the to-i dalam dataset, m is the average value of the dataset. Calculations in finding the Skewness value use the formula (5), where n is the amount of data in the sample, $\mathcal{X}_i$ adalah nilai data pada posisi to-i, $\overline{\mathcal{X}}$ is the average of the data, S is is the standard deviation of the data. Calculations in finding the Entropy value use the formula (6), where N is the number of possible distinct values in the data set, x_i is the intensity value of the pixel value at position to-i, px_i is the probability of occurrence of the value x_i in the dataset. Calculations in finding the Minimum value use the formula (7), where n is the number of elements in the dataset, x_i is the value of the to-i in the dataset. Calculations in finding the Maximum value use the formula (8), where n is the number of elements in the dataset, x_i is the value of the to-i in the dataset.

$$Mean = \frac{\sum_{i=1}^{n} x_i}{n} \qquad (1)$$

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(x_i - Mean)^2}{n}} \quad (2)$$

$$the\ k = \frac{\sum_{i=1}^{n}(x_i - Mean)^4}{n \times SD^4} \quad (3)$$

$$v = \frac{\sum_{i=1}^{n}(x_i - Mean)^2}{n} \quad (4)$$

$$Skewness = \frac{\sum_{i=1}^{n} (\mathcal{X}_i - \bar{\mathcal{X}})^3}{n \cdot SD^3} \quad (5)$$

$$-\sum_{i=1}^{N} p(x_i) \,.\, log_2(p(x_i)) \quad (6)$$

$$Min = \min\,(x_1, x_2, \ldots, x_n) \quad (7)$$

$$Max = \max\,(x_1, x_2, \ldots, x_n) \quad (8)$$

2.3. Dataset and Data Augmentation

The dataset collected is images of different types of cars as many as 840 images divided into 8 classes. Each class has the same size of 227 x 227 pixels. Used for car samples separated into 2 different sets of 90% or 760 images for training data and 10% or 120 images for test data, with equal distribution, and all images are given .jpg image format and include RGB colors. Furthermore, preprocessing, namely cropping and resizing, is applied to the image, after which segmentation is carried out using thresholding and brightness adjustment, and data augmentation techniques that have various functions and parameters. The results of segmented imagery are illustrated in Figure 2. Data augmentation is an approach that allows professionals to significantly increase the variety of data available for training models, without the need to collect new data [13]. This data augmentation technique is useful for increasing variety in a training dataset by making small variations on existing images, such as rotating, mirroring, or panning images as illustrated in Figure 3.

(a) (b) (c)

Figure 2. Example of Raw Dataset: (a) Terios, (b) Fortuner, (c) Avanza

(a) (b) (c) (d)

Figure 3. Example of the results of the Data Augmentation process: (a) original image, (b) brightness adjustment, (c) flip horizontal, (d) rotation using 30^0

The parameters used in this study are:

1. 'RandXReflection', true: This option indicates whether to perform a random horizontal mirror (create a mirrored version of the image). With the value true, the image will be mirrored horizontally randomly.

2. 'RandRotation', [-10, 10]: This option indicates the range of random rotation to be applied to the image. [-10, 10] range means that the image can be randomly rotated between -10 degrees to 10 degrees.
3. 'RandXTranslation', [-10, 10]: This option shows a random range of horizontal shifts applied to the image. This [-10, 10] range means the image can be shifted randomly between -10 pixels to 10 pixels in the horizontal direction.
4. 'RandYTranslation', [-10, 10]: This option shows the random vertical shift range applied to the image. This [-10, 10] range means the image can be shifted randomly between -10 pixels to 10 pixels in the vertical direction.

2.4. Classification with CNN Alexnet

The AlexNet architecture approach is used for the classification of car types in the context of visual pattern recognition. In this task, a collection of input images with a resolution of 227 x 227 pixels and 3 color channels (RGB) is needed and pays attention to the structure of the alexnet architecture and the parameters used. AlexNet has layers such as convolution, ReLU, and pooling. Convolution layers take features from the image through filters. ReLU adds non-linearity, while pooling reduces the dimensionality of features. This process is repeated with subsequent layers to recognize increasingly complex features. Finally, a fully connected layer connects these features to classification classes via a softmax layer. In the image above, Alexnet connects Alexnet features to the car-class imagery used for the training and testing process of this study. Accuracy is a measure used to measure the extent to which a model or system can provide correct results or conform to expected data. In the context of pattern recognition or classification, accuracy aims to measure how well the model can correctly identify and classify data based on the labels or categories it should be.

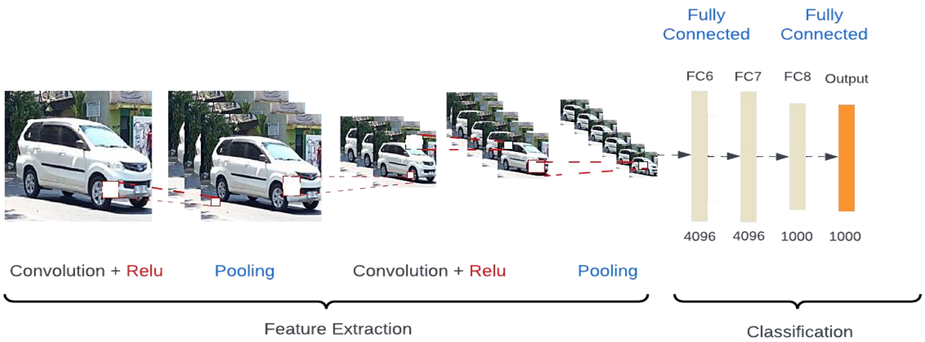

Figure 4. Proposed Structure of AlexNet CNN

3. Results and Discussion

Transfer learning techniques are used in building this type of car recognition model which utilizes a pre-trained model for use in classifying a new data set that undergoes a testing process using the Alexnet architecture to increase accuracy. The dataset collected is in the form of images of various types of cars totaling 840 images which are divided into 8 classes. Each class has the same size, namely 227 x 227 pixels. Used for car samples separated into 2 different sets of 90% or 760 images for training data and 10% or 120 images for test data, with even distribution and all images are rendered in .jpg image format and include RGB colors. In this image processing, there are limitations because some image datasets are taken in real-time on the highway, so the images have to go through a cropping process because there are objects other than cars. There are also some blurry pictures because the car pictures were taken while the car was passing by. Therefore, preprocessing is carried out, namely cropping and resizing the image, after which segmentation is carried out using thresholds and brightness adjustments, as well as data augmentation techniques that have various functions and parameters. The results of the segmented image can be seen in Figure 4.

(a) Original Image (b) Brightness (c) Thresholding

Figure 5. Testing Process

Based on the image above, the next step after preprocessing and segmentation is classification using data augmentation techniques and CNN to classify car types. Data augmentation techniques to increase variation in the training process [14] [15]. The classification process is carried out with an enhanced AlexNet using several variables that are hyperparameters as a result; the optimation function is based on sgdm and adam; the hardware resource uses a single CPU; MaxEpochs is 10; Validation frequency is 30; and the MiniBatch is 10. In this stage, the classification process succeeded in issuing the appropriate results from the input image class. In running this dataset training using the Matlab2022a tool and for training options. Several features a more concise numerical form that reflects the important characteristics of the image. Table 1 shows accuracy and validation visualize the graph and plot the accuracy of the number of iterations. After undergoing training, the model using the 'sgdm' optimization function obtained 99.74% cursive results in Figure 6 and Figure 7. An experiment using the 'Adam' optimization function yielded an accuracy of 96.85%.

Table 1. Extraction Features

Car Type Name	Mean	Standard Deviation	Kurtosis	Skewness	Variance	Entropy	Min	Max
Avanza	143.981	79.6451	1.55745	0.0660755	6343.34	7.71607	1	255
Fortuner	99.7077	58.7835	2.55044	0.317762	3455.5	7.72498	0	255
Freed	105.944	57.8812	2.77128	0.65611	3350.23	7.51721	0	255
Inova	109.519	71.5589	2.22396	0.672527	5120.67	7.57052	0	255
Pajero	113.833	83.9327	1.73023	0.245698	7044.7	7.79173	0	255
Terios	107.178	63.7941	2.51832	0.388208	4069.68	7.76622	1	255
Xenia	107.19	69.4404	2.29854	0.704991	4821.97	7.56934	1	255
Xpander	69.9937	52.9429	4.32262	1.27483	2802.95	7.27868	1	255

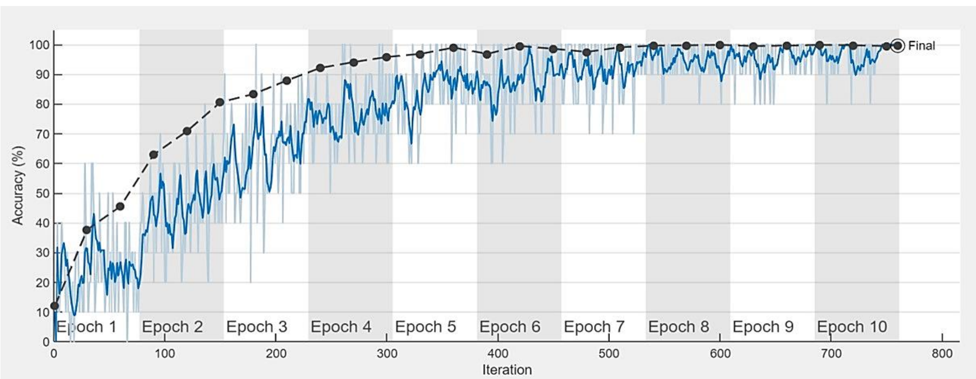

Figure 6. Training accuracy progress

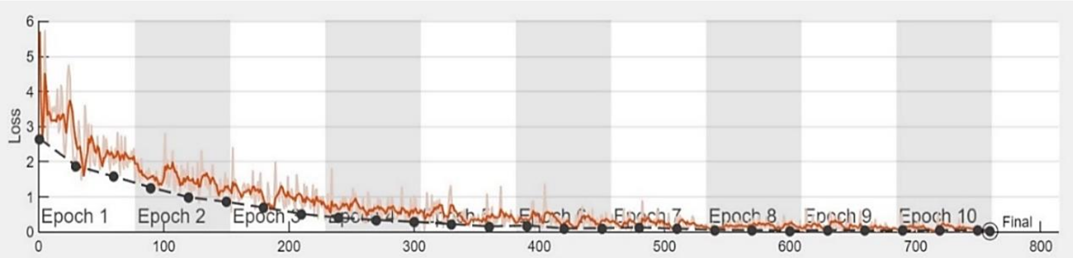

Figure 7. Training loss progress

Based on the results of training and testing, researchers managed to achieve high-accuracy validation and were able to optimize and minimize the value of the loss function in classifying the image of car types. Based on Table 3, the classification of car type classes using CNN Alexnet and data augmentation techniques produces excellent accuracy. All samples that passed the test were classified correctly, resulting in a training accuracy of 99.74%. In 18 experimental attempts per class with 2 samples each, the test results consistently indicated very high accuracy, reaching 100%. In this research, we also conducted experiments using the Deep Convulsive Neural Network (DCNN) method with the same number of datasets and training options but did not use the augmentation technique and AlexNet architecture, resulting in an accuracy of 35.71% with 'Adam' and 22.62% with 'sgdm'. With these results, it can be proven that the CNN method processed through data augmentation techniques using the Alexnet architecture can produce better accuracy.

Table 3. Classification Of Car Types

Image file.jpg	Car Type Class	Car classification	True or False
5.jpg	Avanza	Mobil_Avanza	T
21.jpg	Avanza	Mobil_Avanza	T
1.jpg	Fortuner	Mobil_Fortuner	T
16.jpg	Fortuner	Mobil_Fortuner	T
23.jpg	Freed	Mobil_Freed	T
20.jpg	Freed	Mobil_Freed	T
22.jpg	Inova	Mobil_Inova	T
44.jpg	Inova	Mobil_Avanza	T
17.jpg	Pajero	Mobil_Pajero	T
70.jpg	Pajero	Mobil_Pajero	T
26.jpg	Terios	Mobil_Terios	T
49.jpg	Terios	Mobil_Terios	T
9.jpg	Xenia	Mobil_Xenia	T
42.jpg	Xenia	Mobil_Xenia	T
2.jpg	Xpander	Mobil_Xpander	T
21.jpg	Xpander	Mobil_Xpander	T

4. Conclusion

This study uses the AlexNet architecture to improve the accuracy of car vehicle classification by applying data augmentation techniques. The results of the research achieved a training accuracy of 99.74%. The main finding of this research is to provide valuable insights regarding the use of machine vision technology in identifying car types, which is relevant to automotive industry trends and consumer analysis. Gaikindo's data showing the popularity of Toyota cars in Indonesia reinforces the value of this research contribution for car manufacturers, enabling them to understand market trends and preferences in greater depth. In addition, the implementation of a car model recognition system in the technology field can help the automotive industry develop better solutions for identifying and analyzing trends in car usage. The implications of this research can be used as a basis for further development in optimizing the use of machine vision technology in the automotive industry and related sectors.

References

[1] J. qing Luo, H. sheng Fang, F. ming Shao, Y. Zhong, and X. Hua, “Multi-scale traffic vehicle detection based on faster R–CNN with NAS optimization and feature enrichment,” *Def. Technol.*, vol. 17, no. 4, pp. 1542–1554, 2021, doi: 10.1016/j.dt.2020.10.006.

[2] S. Secinaro, V. Brescia, D. Calandra, and P. Biancone, “Employing bibliometric analysis to identify suitable business models for electric cars,” *J. Clean. Prod.*, vol. 264, p. 121503, 2020, doi: 10.1016/j.jclepro.2020.121503.

[3] F. A. Arsy, “Demand Forecasting of Toyota Avanza Cars in Indonesia: Grey Systems Approach,” *Int. J. Grey Syst.*, vol. 1, no. 1, pp. 38–47, 2021, doi: 10.52812/ijgs.24.

[4] I. Lestari, I. Sadalia, E. S. Rini, and ..., “The Influence of Brand Trust and Product Quality on Customer Loyalty Through Customer Engagement on Users Toyota Cars in Medan City,” *Proceeding ...*, vol. 1, no. 2, pp. 43–61, 2023, [Online]. Available: https://jurnal2.untagsmg.ac.id/index.php/icbe-untagsmg/article/view/729%0Ahttps://jurnal2.untagsmg.ac.id/index.php/icbe-untagsmg/article/download/729/691

[5] L. Zhang, Z. Sheng, Y. Li, Q. Sun, Y. Zhao, and D. Feng, “Image object detection and semantic segmentation based on convolutional neural network,” *Neural Comput. Appl.*, vol. 32, no. 7, pp. 1949–1958, 2020, doi: 10.1007/s00521-019-04491-4.

[6] S. Montaha *et al.*, “MNet-10: A robust shallow convolutional neural network model performing ablation study on medical images assessing the effectiveness of applying optimal data augmentation technique,” *Front. Med.*, vol. 9, no. August, 2022, doi: 10.3389/fmed.2022.924979.

[7] K. Maharana, S. Mondal, and B. Nemade, “A review: Data pre-processing and data augmentation techniques,” *Glob. Transitions Proc.*, vol. 3, no. 1, pp. 91–99, 2022, doi: 10.1016/j.gltp.2022.04.020.

[8] A. Kherraki and R. El Ouazzani, “Deep convolutional neural networks architecture for an efficient emergency vehicle classification in real-time traffic monitoring,” *IAES Int. J. Artif. Intell.*, vol. 11, no. 1, pp. 110–120, 2022, doi: 10.11591/ijai.v11.i1.pp110-120.

[9] M. Faisal *et al.*, “Faster R-CNN Algorithm for Detection of Plastic Garbage in the Ocean: A Case for Turtle Preservation,” *Math. Probl. Eng.*, vol. 2022, 2022, doi: 10.1155/2022/3639222.

[10] A. M. Reddy *et al.*, “An Efficient Multilevel Thresholding Scheme for Heart Image Segmentation Using a Hybrid Generalized Adversarial Network,” *J. Sensors*, vol. 2022, 2022, doi: 10.1155/2022/4093658.

[11] V. R. Joseph, “Optimal ratio for data splitting,” *Stat. Anal. Data Min.*, vol. 15, no. 4, pp. 531–538, 2022, doi: 10.1002/sam.11583.

[12] Hilman Fauzi, Achmad Rizal, Mazaya ’Aqila, Alvin Oktarianto, and Ziani Said, “Classification of Normal and Abnormal Heart Sounds Using Empirical Mode Decomposition and First Order Statistic,” *J. Electron. Electromed. Eng. Med. Informatics*, vol. 5, no. 2, pp. 82–88, 2023, doi: 10.35882/jeeemi.v5i2.287.

[13] B. Shetty, R. Fernandes, A. P. Rodrigues, R. Chengoden, S. Bhattacharya, and K. Lakshmanna, “Skin lesion classification of dermoscopic images using machine learning and convolutional neural network,” *Sci. Rep.*, vol. 12, no. 1, pp. 1–11, 2022, doi: 10.1038/s41598-022-22644-9.

[14] M. Xu, S. Yoon, A. Fuentes, J. Yang, and D. S. Park, “Style-Consistent Image Translation: A Novel Data Augmentation Paradigm to Improve Plant Disease Recognition,” *Front. Plant Sci.*, vol. 12, no. February, pp. 1–16, 2022, doi: 10.3389/fpls.2021.773142.

[15] P. Kollapudi, S. Alghamdi, N. Veeraiah, Y. Alotaibi, S. Thotakura, and A. Alsufyani, “A New Method for Scene Classification from the Remote Sensing Images,” *Comput. Mater. Contin.*, vol. 72, no. 1, pp. 1339–1355, 2022, doi: 10.32604/cmc.2022.025118.