Jurnal Teknik Informatika (JUTIF)
Vol. 5, No. 4, August 2024, pp. 1165-1171

DOI: https://doi.org/10.52436/1.jutif.2024.5.4.2160
p-ISSN: 2723-3863
e-ISSN: 2723-3871

VGG-16 ARCHITECTURE ON CNN FOR AMERICAN SIGN LANGUAGE CLASSIFICATION

Mutiara Dolla Meitantya[*1], Christy Atika Sari[2], Eko Hari Rachmawanto[3], Rabei Raad Ali[4]

[1,2,3]Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang, Indonesia
[24]Faculty of Computer Science, Northern Technical University, Mosul, Iraq
Email: [1]111202113745@mhs.dinus.ac.id, [2]christy.atika.sari@dsn.dinus.ac.id, [3]eko.hari@dsn.dinus.ac.id, [4]rabei@dsn.dinus.ac.id

(Article received: May 30, 2024; Revision: July 03, 2024; published: July 29, 2024)

Abstract

Every country has its sign language such as in Indonesia there are 2 types namely Indonesian Sign Language System called SIBI and BISINDO (Indonesian Sign Language). American Sign Language (ASL) is a sign language that is widely used in the world. In this research, the classification of American Sign Language (ASL) using the Convolutional Neural Network (CNN) method using VGG-16 architecture with Adam optimizer. The data used is 14000 ASL image data with 28 classes consisting of letters A to Z plus space and nothing with a division of 90% training data and 10% validation data. From this research, the overall accuracy is obtained with a value of 98% and the accuracy value of validation data evaluation is 89.07%.

Keywords: *Adam, American Sign Language, Classification, Convolutional Neural Network, VGG-16.*

1. INTRODUCTION

In general, communication is carried out with spoken or written words, but in some cases, some people cannot communicate verbally due to physical limitations such as deafness and speech impairment [1]. So, usually, they communicate using nonverbal communication such as gestures using both hands and other limbs, this communication is commonly called sign language [1]. Sign Language is a language used by people who have difficulty hearing or speaking [2].

There are many deaf and speech-impaired people in Indonesia. However, many people do not understand their sign language because people's interest in sign language is still not high enough. [2]. So that causes the deaf and speech impaired to have difficulty in expressing feelings or communicating. This is also due to the low level of sign language translators. These problems can be solved with a system that can classify sign language hand movement patterns. Each country has its sign language system. There are 2 types of sign language in Indonesia, the Indonesian Sign Language System called SIBI and BISINDO (Indonesian Sign Language). *American Sign Language* (ASL) is a sign language that is widely used in the world. [3]. ASL is a sign language linguistically the same as spoken language expressed with hand gestures. [4].

Many previous researchers have classified sign language in both the Indonesian Sign Language System (SIBI) and American Sign Language (ASL). In research [5] classification of the Indonesian Sign Language System (SIBI) using the Convolution Neural Network (CNN) method with VGG (Visual Geometric Group)-16 architecture and Alexnet. In this study, the data consists of the letters A to the letter Z, namely 320 test data, 1600 training data, and 320 validation data, and the data will be resized to a size of 224 x 224 pixels, followed by grayscale and augmentation [5]. In the test results, the highest accuracy rate was obtained when using the VGG-16 architecture with Adam optimization which resulted in an accuracy rate of 99.32% for each letter and 91.18% for the whole letter [5]. *American Sign Language* (ASL) sign language classification research was also conducted by [2] using the CNN model with the use of several layers including Conv2D, Dropout, Flatten, and Dense, and produces a fairly good accuracy rate of 82.1%. Other studies that use the CNN method were also conducted by [6] which resulted in an accuracy of 52% for all letters.

Here, the author classifies *American Sign Language* (ASL) hand gesture patterns using the CNN method with VGG-16 architecture. The selection of this architecture itself is because it has been proven by its application in other classifications, for example, the quality classification of salak fruit [7], dermoscopic image classification of skin cancer [8], glioma image classification [9], and classification of the Batik motif [10]. This research was conducted because *American Sign Language* (ASL) is a world sign language and the use of the VGG-16 architecture itself has also been proven accurate by previous research.

2. METHOD

Here, the classification of *American Sign Language* (ASL) will be investigated using the Convolutional Neural network (CNN) method using the VGG-16 architecture. Figure 1 below shows the stages in researching the classification of *American Sign Language* (ASL) will be studied using the Convolutional Neural Network (CNN) method using the VGG-16 architecture.

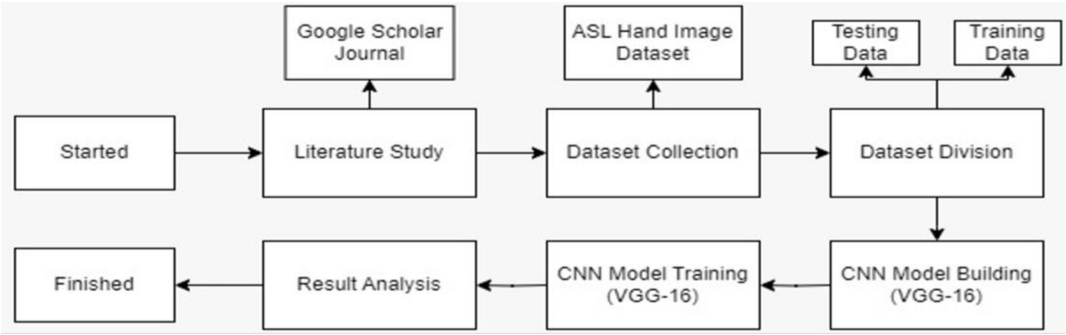

Figure 1. Proposed Research Methods

2.1. Literature Study

At this stage, researchers searched relevant journals related to the classification of *American Sign Language* (ASL) using Convolutional Neural Network using VGG-16 architecture.

1) *American Sign Language* (ASL) is a sign language that is the same as spoken but expressed using hand gestures. ASL is widely used by deaf people in America and is the most widely used sign language in the world.
2) *Convolutional Neural Network* (CNN) is a type of neural network used widely in the field of computer vision [11] *Convolutional Neural Network* (CNN) is also a development of multilayer perceptron (MLP) which is designed to process two-dimensional data in the form of images. [12]. In this method, the classification uses a convolution layer to calculate the convolution between the input and the filter.
3) VGG-16 is a CNN architecture developed by Simonyan and Zisserman in 2014. [5]. This architecture is one of the well-known and effective CNN architectures for classification tasks. The VGG-16 architecture has a total of 16 weight layers consisting of a Convolutional layer and a Fully Connected layer. [13], with a total of 13 Convolution layers and 3 Fully Connected layers. [14]. The default image size that can be accepted by VGG-16 is 224x24 pixels. [15]. However, in its implementation, the size of the input image used can adjust the research needs. This is evidenced by the existence of research [13] which also uses VGG-16 but with various input image sizes, namely the image size is changed from the original 224 × 224 pixels to 128 × 128 pixels and 64 × 64 pixels. The following is an example of a layer diagram on the VGG-16 architecture.

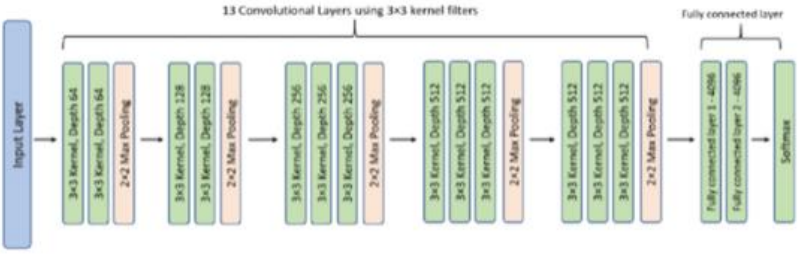

Figure 2. VGG-16 Architecture [7]

Based on Figure 2, the VGG-16 architecture has the first layer, the image input layer, which is used in the model training and testing process. Then the next layer is the convolution layer which amounts to 13 *convolutional layers. The* first and second convolutional layers consist of 64 feature kernel *filters* and the filter size is 3x3. After passing through this layer, the resulting output is passed to the *max pooling* layer. Then the third and fourth convolution layers consist of 128 feature *kernel filters* with the same *filter* size of 3x3, whose output is also passed to the max *pooling* layer. Then enter the fifth, sixth, and seventh convolution layers consisting of 256 feature *kernel filters* with a 3x3 filter size whose output also passes through the max *pooling layer* as well. Then continued with the eighth, ninth, and tenth convolution layers with 512 feature kernel filters and a filter size of 3x3, then through the *max pooling layer* again. And the eleventh to thirteenth layers also use 512 feature *kernel filters* and the output goes through the max *pooling layer*. And after that, it goes to the *Fully Connected Layer* which successively reduces the feature dimensions to the desired number of classes for classification. The *Fully Connected* Layer is followed by the ReLU activation function, except the last layer which uses the *softmax* activation function to generate the class probability distribution of the input. Finally, the *output* layer generates the class of the input image based on the probability distribution generated by the *softmax* layer and the previous layers.

2.2. Data Collection

The dataset used namely hand shapes from *American Sign Language* (ASL) taken from public data on Kaggle. [16]. The dataset contains test data and training data, both training data and testing data contain 28 classes, namely 1 image per letter plus del, nothing, and space images in the testing data. The training data contains the letters A to Z plus space, nothing, and del with 3000 images per letter. There is a total of 84,000 images of training data and 28 images of testing data. Here's a look at the classes stored in the folder. From the data obtained, here the author does not use all of the data, only some of it. The following will also include examples of datasets per letter. In this study, 28 classes consisting of letters A to Z plus *Space* and *Nothing* as shown in Figure 3. The total dataset is 14000 image data which is divided into 2 categories, namely training data and validation data with a division of 90% training data and 10% validation data. The total training data used is 12600 images while the total validation data is 1400 images which are visualized in Figure 4.

Figure 3. The 28 ASL Classification Classes

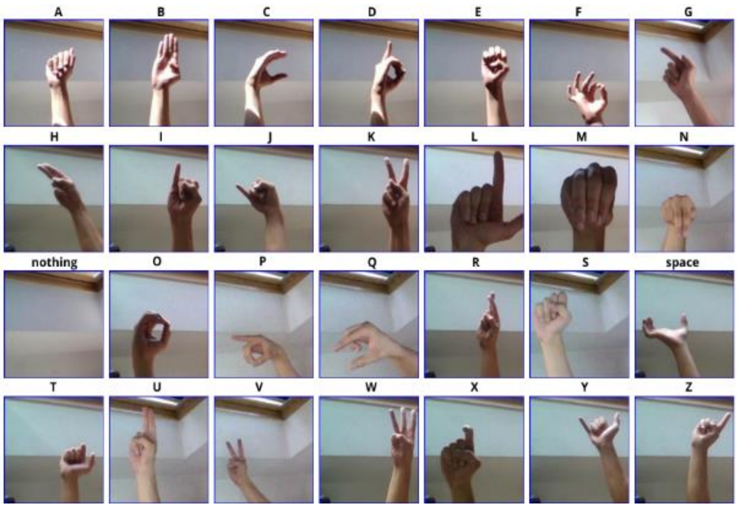

Figure 4. Visualization dataset for each letter

2.3. CNN Modeling (VGG-16)

In this section, the VGG-16 Convolutional Neural Network (CNN) model is built. VGG-16 is a CNN architecture that consists of 16 convolutional layers and *is fully connected*. This model is well known for its efficiency in learning complex image features. The *optimizer* used in this research is the Adam *optimizer*. The layers used in this research are as follows:

1. Base Model (VGG-16). That is, the VGG-16 model is loaded without a fully connected layer on top. This includes all the convolution and max pooling layers of the VGG-16 pre-trained for the image recognition task.
2. Flatten(). It is a layer used to convert the output into one dimension so that it can be used for input to the Dense layer. This is done because the Dense layer (Fully Connected Layer) which will be used next requires input in the form of a one-dimensional vector.
3. Dense(256, activation='relu'). This is a Dense layer (fully Connected Layer) with 256 units. This means that every neuron in this layer is connected to every neuron in the previous layer. Then the activation function 'relu' (Rectified Linear Unit) is used in each neuron in this layer. ReLU is a commonly used activation function for hidden layers in artificial neural networks due to its non-linear nature and computational efficiency.
4. Dense(128, activation='relu'). This layer has 128 units with the activation function 'relu'. This is the second fully connected layer after the previous Dense layer. The number of units (neurons) in this layer is less than the previous layer of 256 units to speed up computation and strengthen model generalization.
5. Dense(28, activation='softmax'). This is the last Dense layer used for classification. The number of units is 28, corresponding to the classes of the dataset being conducted in this study. The

number of classes to be predicted are A to Z, Nothing, and Space. The activation function 'softmax' is used at this stage. Softmax generates a probability distribution that describes the prediction probability for each class. The predictions from the model will be for each class, and the class with the highest probability will be selected as the final result.

2.4. CNN Model Training (VGG-16)

The VGG-16 model was trained using the previous dataset. Training is done by specifying parameters such as epoch, batch size, and optimization method. During training, the model will learn to recognize patterns and features in hand images that represent *American Sign Language*.

2.5. Result analysis

After the training is complete, the model is evaluated using the validation set. The results of this evaluation are used to analyze how good the VGG-16 model is at classifying *American Sign Language*.

3. RESULT AND DISCUSSION

This research was conducted with several main stages, namely data collection, model training, and outcome evaluation. Each stage will be explained in more detail as follows:

1) Data Collection

The data used in this research is image data about American Sign Language obtained from public datasets from Kaggle. Of the entire dataset, the data used is only 28 classes consisting of the letters A through Z plus the Space and Nothing classes, each of which has 500 image data. Then each image will be resized to 160x160 pixels before further processing.

2) Model Training

The next stage is model training using VGG-16 architecture and Adam optimization with a learning rate of 0.001 which is trained for 10 epochs. The per-epoch accuracy results can be seen in table 2 below. A graph of the training results is also included in Figure 5 after Table 2.

3) Evaluation of Results

After the training is complete, the next process is testing the model using validation data. From this process, the overall accuracy reached 98% and 89.07% on the validation data. These results show that the model has excellent performance in recognizing American Sign Language.

Table 2. Accuracy of each Epoch

Epoch	Loss	Accuracy	Validation Loss	Validation Accuracy
Epoch 1	1.9444	0.4297	0.9848	0.6962
Epoch 2	0.1668	0.9449	0.9140	0.8118
Epoch 3	0.0826	0.9725	0.8146	0.8467
Epoch 4	0.0731	0.9788	0.5434	0.8837
Epoch 5	0.0444	0.9873	1.2458	0.7900
Epoch 6	0.0519	0.9852	0.8984	0.8263
Epoch 7	0.0500	0.9853	0.5804	0.8939
Epoch 8	0.0172	0.9948	0.9190	0.8605
Epoch 9	0.0332	0.9907	0.9400	0.8605
Epoch 10	0.0377	0.9899	0.7352	0.8917

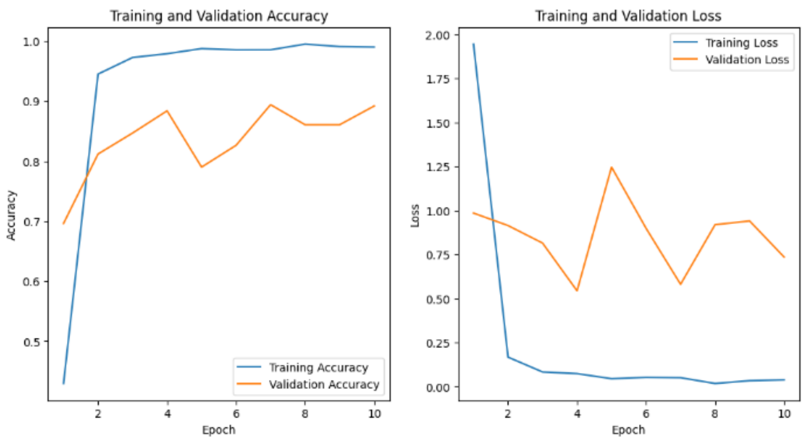

Figure 5. Graph of model training results

Table 3. Comparison of Accuracy of Previous Sign Language Research

No	Types of Sign Language	Methods	Optimizer	Number of Classes	Accuracy
1	Indonesian Sign Language System [5]	VGG-16	Adam	26	91,18%
2	Indonesian Sign Language System [5]	VGG-16	SGD	26	84,96%
3	Indonesian Sign Language System [17]	CNN	Unknown	15	88%
4	American Sign Language [2]	CNN	Unknown	26	82,1%
5	American Sign Language [6]	CNN	Adam	26	52%
6	American Sign Language (proposed research)	VGG-16	Adam	28	98%

Table 3 shows the implementation of VGG-16 with Adam's *optimizer* is very good among other methods regardless of the type of sign language studied. Based on the research results listed, in the first study conducted by [5] about the classification of Indonesian Sign Language Systems with the VGG-16 method resulted in an accuracy of 91.18% using the Adam *optimizer*. This research uses a total of 26 classes in its classification research. In the same study, the VGG-16 algorithm was also applied with the same dataset and number of classes, but with a different type of optimizer, namely using the SGD optimizer. In using the SGD optimizer and the VGG-16 algorithm, the classification also produces high overall accuracy by touching the 84.96% accuracy rate. In a later study conducted by [17], image classification on the Indonesian Sign

Language System produces a fairly good accuracy of 88% but the optimizer used is not known. In this study, the method used is the CNN method whose model is integrated into an Android application developed with the Flutter framework. With this accuracy, this research uses 15 classes in its accuracy testing. Then in the next research conducted by [2], research on *American Sign Language* with the *Convolutional Neural Network* (CNN) method produced a fairly good accuracy of 82.1%. The number of classes used is the same as in previous research, totaling 26 classes consisting of. However, in this study, it is not known what *optimizer* was used. Furthermore, the research conducted by [6] *American Sign Language* classification produces an accuracy of 52% by using the *Convolutional Neural Network* (CNN) method and Adam's *optimizer*. In this study, the number of classes used was 26 classes. Whereas in the current research, classification using *American Sign Language* with the VGG-16 method and Adam's *optimizer* has the best accuracy value, namely with a value of 98%. The number of classes used is 28 classes consisting of letters A through Z plus space and nothing. As for the evaluation results of validation data, the accuracy is quite good with a value of 89.07%.

4. DISCUSSION

The VGG-16 architecture in the CNN algorithm has also been applied in previous studies that also discuss image classification with different themes. Some examples of research used the CNN algorithm with VGG-16 architecture and has proven to produce excellent accuracy will be shown in Table 4.

Table 4. Research Using CNN Architecture VGG-16

No	Classification Image Type	Methods	Optimizer	Number of Classes	Accuracy
1	Moth [18]	VGG-16	Adam	50	95%
2	Lampung Script Handwriting [19]	VGG-16	Adam	20	91%
3	Skin Cancer [20]	VGG-16	SGD	4	99,70%
4	Fundus [21]	VGG-16	GDM	2	92,31%
5	X-ray image Lungs [22]	VGG-16	Adam	4	91,45%
6	Classification Of Batik Motif [10]	VGG-16	Adam	5	91,23%

The first research with the VGG-16 architecture is [18] where this research discusses the classification of moth images totaling 50 classes. The VGG-16 method in this study was accompanied by the Adam optimizer to produce very good accuracy reaching 95 Further research [19]which examines the classification of handwritten images of Lampung script also uses the VGG-16 method. In this study, the classes used amounted to 20 classes according to the letters of Lampung script, namely Ka, Ga, Nga, Pa, Ba, Ma, Ta, Da, Na, Ca, Ja, Nya, Ya, A, La, Ra, Sa, Wa, Ha, Gha. This research uses Adam's optimizer for optimization and produces very good accuracy as well with a value of 91%. The VGG-16 method in CNN is also used in classification research [20] to classify skin cancer with the SGD optimizer. This research uses 4 classes of skin cancer namely melanoma, squamous cell carcinoma, dermatofibroma, and nevus pigmentosus. Using the VGG-16 method, this research produces very high accuracy touching 99.70%. Research [21] also used the VGG-16 architecture on CNN in fundus classification research. In this study, fundus classification used the GDM optimizer with a total of 2 classes consisting of normal classes and neovascularization classes. With the VGG-16 method, fundus classification resulted in an accuracy of 92.31%. The last research example that proves the accuracy of the VGG-16 method is research [22] which performs classification on Lung X-ray images. There are 4 classes used in this study, namely normal X-ray images, x-ray images of COVID-19 sufferers, x-ray images for viral pneumonia, and X-ray images for lung opacity. In this study using Adam's optimizer managed to get an accuracy of 91.45%.

Based on examples of previous VGG-16 research [18], [19], [20], [21], [22] It can be concluded that this method can produce very good accuracy in image classification. What distinguishes this algorithm from other image classification algorithms is the deeper network structure and more convolution layers. The limitations of the VGG-16 algorithm are its many parameters and the potential for overfitting on small datasets. [23], [24], [25], [26]. But behind it all, VGG-16 has the advantage of its ability to extract quality features from images with excellent performance. [23], [24], [26].

Similar research that also discusses the classification of sign language has also been carried out by previous research that has been described in the previous chapter. from these research studies, researchers in the field of American sign language because not much research has been done about it.

5. CONCLUSION

This research is a study on the classification of *American Sign Language*. In this study, the dataset used amounted to 500 images per letter with a total of 14000 data with 28 classes consisting of letters A to Z plus space and nothing. This dataset is divided into 2 datasets, namely training data and validation data with a ratio of 90% for training data and 10% for validation data. CNN method with VGG-16 architecture is applied to this *American Sign Language* classification research. The selection of the VGG-16 architecture is based on the existence of previous image classification studies that used this architecture and got very good accuracy, regardless of what the study was classifying. In this American Sign Language research using optimization, namely the Adam *optimizer*. Based on our experiment, it is concluded that the use of the CNN method with VGG-16 architecture and Adam optimizer produces excellent accuracy with the overall accuracy value reaching 98%. The accuracy of the validation data evaluation reached 89.07%. From this research, it is also concluded that the use of the CNN algorithm with the VGG-16 architecture in classification produces very good accuracy as evidenced by previous research that performs classification, regardless of what the classification of previous research is about. Future development can be done using several ways, one of which is using different datasets and then improving the model by changing the layers used and also performing different preprocessing stages so that it is expected to produce higher accuracy values and lower loss values.

BIBLIOGRAPHY

[1] L. Arisandi and B. Satya, “Sistem Klarifikasi Bahasa Isyarat Indonesia (Bisindo) Dengan Menggunakan Algoritma Convolutional Neural Network,” *Jurnal Sistem Cerdas*, vol. 5, no. 3, 2022, doi: 10.37396/jsc.v5i3.262.

[2] F. D. Siswanto, C. C. Lestari, and E. Tanuwijaya, “Klasifikasi Bahasa Isyarat Amerika menggunakan Convolutional Neural Network,” *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 10, no. 1, p. 139, Jan. 2022, doi: 10.26418/justin.v10i1.47184.

[3] A. Fendiawati and M. E. Al Rivan, “Klasifikasi American Sign Language Dengan Metode VGG-19,” *MDP Student Conference*, vol. 2, no. 1, 2023, doi: 10.35957/mdp-sc.v2i1.4466.

[4] M. R. Alviando, M. E. Al Rivan, and Y. Yoannita, “KLASIFIKASI AMERICAN SIGN LANGUAGE MENGGUNAKAN FITUR SCALE INVARIANT FEATURE TRANSFORM DAN JARINGAN SARAF TIRUAN,” *Jurnal Algoritme*, vol. 1, no. 1, 2020, doi: 10.35957/algoritme.v1i1.403.

[5] M. E. Al Rivan and S. Hartoyo, “Klasifikasi Isyarat Bahasa Indonesia Menggunakan Metode Convolutional Neural Network,” *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 8, no. 2, Aug. 2022, doi: 10.28932/jutisi.v8i2.4863.

[6] K. Kersen and W. Widhiarso, “Penerapan Metode Convolutional Neural Network dalam Klasifikasi Bahasa Isyarat,” *MDP Student Conference*, vol. 2, no. 1, 2023, doi: 10.35957/mdp-sc.v2i1.4221.

[7] R. Rismiyati and A. Luthfiarta, “VGG16 Transfer Learning Architecture for Salak Fruit Quality Classification,” *Telematika*, vol. 18, no. 1, 2021, doi: 10.31315/telematika.v18i1.4025.

[8] Qorry Aina Fitroh and Shofwatul ’Uyun, “Deep Transfer Learning untuk Meningkatkan Akurasi Klasifikasi pada Citra Dermoskopi Kanker Kulit,” *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 12, no. 2, 2023, doi: 10.22146/jnteti.v12i2.6502.

[9] M. F. Naufal and S. F. Kusuma, “Penerapan Deep Learning Menggunakan VGG-16 untuk Klasifikasi Citra Glioma,” *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 3, no. 4, 2022, doi: 10.30865/json.v3i4.4122.

[10] R. F. Alya, M. Wibowo, and P. Paradise, “CLASSIFICATION OF BATIK MOTIF USING TRANSFER LEARNING ON CONVOLUTIONAL NEURAL NETWORK (CNN),” *Jurnal Teknik Informatika (Jutif)*, vol. 4, no. 1, 2023, doi: 10.52436/1.jutif.2023.4.1.564.

[11] A. Anggraini and H. Zakaria, “Penerapan Metode Deep Learning Pada Aplikasi Pembelajaran Menggunakan Sistem Isyarat Bahasa Indonesia Menggunakan Convolutional Neural Network (Studi Kasus : SLB-BC Mahardika Depok),” *JURIHUM : Jurnal Inovasi dan Humaniora*, vol. 1, no. 4, 2023, [Online]. Available: https://jurnalmahasiswa.com/index.php/jurihum

[12] I. Amri, “IMPLEMENTASI ALGORITMA CONVOLUTIONAL NEURAL NETWORK UNTUK MENERJEMAHKAN BAHASA ISYARAT,” *Jurnal Multidisiplin Saintek Volume* , vol. 2, no. 9, pp. 70–87, 2024, [Online]. Available: https://ejournal.warunayama.org/kohesi

[13] R. J. Gunawan, B. Irawan, and C. Setianingsih, “PENGENALAN EKSPRESI

WAJAH BERBASIS CONVOLUTIONAL NEURAL NETWORK DENGAN MODEL ARSITEKTUR VGG16," *e-Proceeding of Engineering*, vol. 8, no. 5, 2021.

[14] Ety Sutanty, Maukar, Dina Kusuma Astuti, and Handayani, "Penerapan Model Arsitektur VGG16 Untuk Klasifikasi Jenis Sampah," *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 3, no. 2, 2023, doi: 10.51454/decode.v3i2.331.

[15] C. Lubis, M.Kom., D. Sumarlie, and T. Handhayani, "PENGENALAN KUE TRADISIONAL INDONESIA MENGGUNAKAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK," *Computatio : Journal of Computer Science and Information Systems*, vol. 6, no. 2, 2023, doi: 10.24912/computatio.v6i2.21098.

[16] Akash Nagaraj, "ASL Alphabet," Kaggle.

[17] S. S. Sindarto, D. E. Ratnawati, and I. Arwani, "Klasifikasi Citra Sistem Isyarat Bahasa Indonesia (SIBI) dengan Metode Convolutional Neural Network pada Perangkat Lunak berbasis Android," *urnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 6, no. 5, 2022.

[18] R. Ricky, E. Hartati, and E. P. Widiyanto, "Identifikasi Ngengat menggunakan Metode Convolutional Neural Network," *MDP Student Conference*, vol. 2, no. 1, 2023, doi: 10.35957/mdp-sc.v2i1.4292.

[19] R. Rikendry and A. Maharil, "PERBANDINGAN ARSITEKTUR VGG16 DAN RESNET50 UNTUK REKOGNISI TULISAN TANGAN AKSARA LAMPUNG," *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 3, no. 2, 2022, doi: 10.33365/jatika.v3i2.2030.

[20] R. AGUSTINA, R. MAGDALENA, and N. K. C. PRATIWI, "Klasifikasi Kanker Kulit menggunakan Metode Convolutional Neural Network dengan Arsitektur VGG-16," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 10, no. 2, 2022, doi: 10.26760/elkomika.v10i2.446.

[21] W. Setiawan, "PERBANDINGAN ARSITEKTUR CONVOLUTIONAL NEURAL NETWORK UNTUK KLASIFIKASI FUNDUS," *Jurnal Simantec*, vol. 7, no. 2, 2020, doi: 10.21107/simantec.v7i2.6551.

[22] T. Berliani, E. Rahardja, and L. Septiana, "Perbandingan Kemampuan Klasifikasi Citra X-ray Paru-paru menggunakan Transfer Learning ResNet-50 dan VGG-16," *Journal of Medicine and Health*, vol. 5, no. 2, 2023, doi: 10.28932/jmh.v5i2.6116.

[23] E. Tanuwijaya and A. Roseanne, "Modifikasi Arsitektur VGG16 untuk Klasifikasi Citra Digital Rempah-Rempah Indonesia," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 21, no. 1, 2021, doi: 10.30812/matrik.v21i1.1492.

[24] N. L. W. Rahayu, N. Gunantara, and M. Sudarma, "Klasifikasi Jajanan Khas Bali Untuk Preservasi Pengetahuan Kuliner Lokal Menggunakan Arsitektur VGG-16," *SINTECH : Science And Information Technology Journal*, vol. 7, no. 1, pp. 1–14, 2024, [Online]. Available: https://doi.org/10.31598

[25] M. F. Naufal and S. F. Kusuma, "Pendeteksi Citra Masker Wajah Menggunakan CNN dan Transfer Learning," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 8, no. 6, pp. 1293–1300, 2021, doi: 10.25126/jtiik.2021865201.

[26] A. Dwi Putro and H. Tantyoko, "Hybrid Algoritma Vgg16-Net Dengan Support Vector Machine Untuk Klasifikasi Jenis Buah dan sayuran," *JTIM : Jurnal Teknologi Informasi dan Multimedia*, vol. 5, no. 2, 2023, doi: 10.35746/jtim.v5i2.335.