IJASST
International Journal of
Applied Sciences and Smart Technologies

Volume 07
Issue 02

p-ISSN 2655-8564, e-ISSN 2685-943, pp. 433-450, 2025

Sign Language Detection Models using Resnet-34 and Augmentation Techniques

Rizki Ramdhan Hilal[1], Aradea[1], Vega Purwayoga[1*]

[1]*Department of Informatics, Faculty of Engineering, Siliwangi University, Indonesia*

[*]*Corresponding Author: vega.purwayoga@unsil.ac.id*

(Received 26-06-2025; Revised 31-07-2025; Accepted 25-08-2025)

Abstract

For deaf or hard of hearing people, sign language is a primary means of communication, but low public understanding makes social engagement difficult. Researchers now use computer vision technology and Convolutional Neural Network (CNN) to detect sign language movements. Problems such as overfitting and missing gradients still exist. Using CNN and ResNet-34 architecture, as well as image augmentation to overcome this problem, this research builds a deep learning-based sign language detection model. The Indonesian Sign Language System (SIBI) dataset was used to test the model. The test results show that the model with image augmentation trained for more than 50 epochs obtained an accuracy of 99.4%, precision of 99.5%, recall of 99.5%, and an F1 score of 99.5%. The model without image augmentation produced an accuracy of 99.4%, recall of 99.3%, F1 score of 99.3%, and precision of 99.4%. ResNet-34 architecture overcomes the problem of missing gradients, while image augmentation avoids overfitting and improves model accuracy.

Keywords: Augmentation, Convolutional Neural Network, Overfitting, ResNet-34, Vanishing Gradient

1 Introduction

Humans utilize language as a tool to communicate with one another. Language can take many different forms, including signs, symbols, codes, and noises that are given meaning after being converted into human language according to predetermined rules [1], [2]. Using hand movements that follow the PUEBI Guidelines, people with hearing and speech disabilities can communicate through sign language [3].

The issue with using sign language as a communication tool is that not everyone fully understands this system of communication because there is a dearth of knowledge and resources about learning sign language, including books, courses, teachers, and other resources that can be a barrier for those who wish to take sign language classes [4]. Using

This work is licensed under a Creative Commons Attribution 4.0 International License

deep learning technology to create a photo image identification model that can assist in sign language translation is one way to address this issue. The aid of deep learning technology a model that is capable of self-learning computational procedures [5], to help those who are unable to communicate through sign language, this technology is intended to recognize hand motions and translate them into a language that they can understand sign language [4]. In earlier studies carried out by [6], Convolutional Neural Network (CNN) were used to create a sign language categorization system, and the accuracy rate was 99.82%.

Increasing the depth of the network in deep convolutional neural networks does not always lead to improved training accuracy, there are instances when training accuracy decreases [7], [8], [9], [10]. This is the vanishing gradients problem, which is a common impediment in CNN. This is because not all networks are simple to optimize, which leads to network degradation. Using residual networks, or ResNet, to add identity mapping will lessen degradation in deeper networks ResNet [11]. Previous research on the application of the ResNet architecture by [9], [12], [13], [14], [15] has successfully overcome the vanishing gradient problem in the CNN algorithm.

The possibility of overfitting arises from CNN models' memorization of specific, non-generalizable details of training images, which is another issue frequently associated with CNN algorithms digeneralisasi [16], [17], [18], [19], [20]. Therefore, by adjusting the dimensional transformation of the photos, image augmentation techniques were applied to the training samples in this study to increase the variety of images [21]. In earlier studies by [22], [23], [24], [25] on the application of augmentation techniques on datasets, has effectively addressed CNN models overfitting issue.

The primary research gaps in this study are related to the CNN algorithm's potential for vanishing gradient and overfitting in the final model, as indicated by the research references pertaining to the development of digital image detection models using this algorithm that were mentioned in the previous paragraph. Therefore, the goal of this research is to use the ResNet-34 architecture to create a detection model that can solve the disappearing gradients problem [26] because it may be applied to validation data and has a low error rate value, to produce the best accuracy outcomes, and enhancing the dataset with augmentation methods to prevent the model from becoming overfit along with incorporating dataset augmentation methods to prevent the model from becoming overfit [21], [22], [23].

The contribution of this research is to develop a ResNet-34 model combined with the application of augmentation techniques on the dataset to overcome the problems of overfitting and vanishing gradients in the model in the process of detecting the Indonesian Sign Language System (SIBI). Compared to other studies, which are limited in their use of model architecture and dependent on limited datasets, this study provides evaluation and improvement using deeper residual networks.

2 Material and Methods

The steps or procedures utilized in research are referred to as the research stages [27]. To guarantee that every step of the research is conducted in an orderly fashion, this stage offers direction and arranges every step from start to finish. The phases of the research that will be done are based on earlier studies that [19], the research stages is shown in Fig. 1.

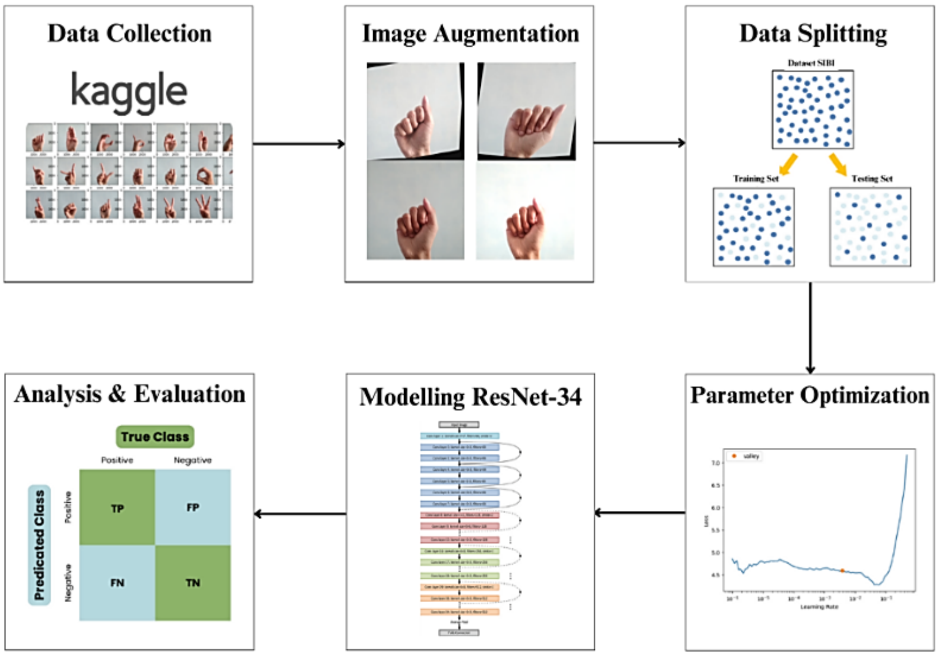


Figure 1. Research Stages

2.1 Data Collection

The collected dataset consists of digital photos of Indonesian sign language systems that were published to Kaggle in 2022 by Alvin Bintang. The dataset consists of 5280 images in total, grouped into 24 alphabetic classes (A–Y, except J and Z). There are 220 Sistem Isyarat Bahasa Indonesia (SIBI) photos per class.

2.2 Data Augmentation

The original dataset used in this study consisted of 5280 images, with 220 images in each class. To increase the variety of the dataset, augmentation techniques were applied. The augmentation techniques are based on earlier studies by [25] and [28] that employed rotation, shear horizontal and vertical, grayscale, saturation, brightness, and exposure. The arrangement of the augmentation values is shown in Table 1.

Based on the augmentation process in Table 1, the number of datasets increased to 7582 with the distribution of each class increasing by 96 images, so that the total number of datasets in each class was 316 images after the augmentation process. The augmentation technique was applied evenly to each class with the aim of maintaining class balance and avoiding class imbalance issues.

Table 1. Augmentation Technique

No.	Techniques	Description
1	Rotation	The rotation technique used is -10° and +10°.
2	Shear horizontal & vertical	The horizontal and vertical shear techniques used are ±10° horizontal and ±10° vertical.
3	Grayscale	Grayscale technique used is 15% of the total dataset.
4	Saturation	Saturation technique used is -30% and +30%.
5	Brightness	The brightness technique used is -15% and +15%.
6	Exposure	The exposure technique used is -15% and +15%.

2.3 Data Splitting

Following the preprocessing phase, datasets are separated into training and test subsets. The split_indices function from the FastAI library is used to divide the dataset. The technique is based on earlier research by [29] utilizing a ratio of 80% for training data and 20% for test data.

2.4 Modelling ResNet-34

ResNet-34 architecture, which has several convolution layers and residue layers, is used for modeling in Fig. 2. The method of entering a 224 by 224pixel picture is the first step. With an output of 128 x 128 pixels, the input image is passed through 64 filters into the first convolution layer in the second stage. Furthermore, there are four residual layers with the following output sizes 28 x 28 pixels with 4 residual blocks using 128 filters, 14 x 14 pixels with 6 residual blocks using 256 filters, and 7 x 7 pixels with 3 residual blocks using 512 filters. The size of the residual layers is 56 x 56 pixels with 3 residual blocks using 64 filters. After the residual layer, an average pooling layer reduces the output size to 1 x 1 pixels, followed by a fully connected layer with 256 and 128 neurons. The final stage is the softmax classifier layer which produces a classification with accuracy (E = 0.99). Each residual layer has skip connections that allow direct information flow, helping to overcome the vanishing gradient problem [12], [30].

2.5 Model Evaluation

Model evaluation is crucial for identifying the optimal model combination. It involves using matrices, considering factors like accuracy and the confusion matrix. The confusion matrix serves as a visual evaluation tool in machine learning [31].

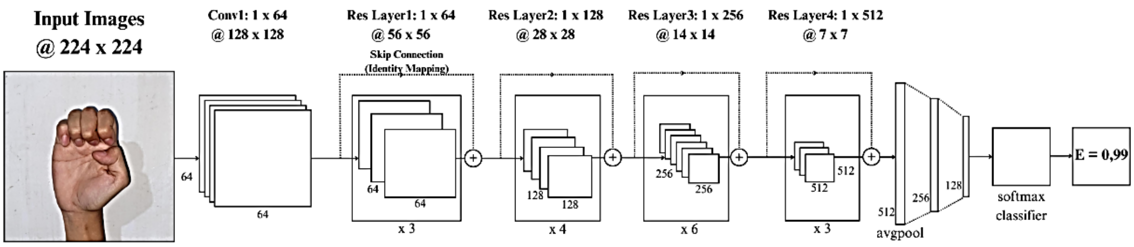


Figure 2. ResNet-34 Architecture

Its columns represent predicted class outcomes, while its rows depict actual class outcomes, facilitating the examination of all potential cases in classification problems [28]. Various metrics such as precision, recall, and F1 score are utilized within the confusion matrix. This matrix encompasses 4 key terms [32]:

a. True Positive (TP): correctly classified positive data.
b. True Negative (TN): the number of correctly classified negative data.
c. False Positive (FP): the number of negative data classified as positive.
d. False Negative (FN): the number of positive data classified as negative.

Accuracy and the confusion matrix can be formulated as shown in Equation (1)-(4) [31].

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{TP + TN}}{\boldsymbol{TP + TN + FP + FN}} \tag{1}$$

Accuracy is the ratio of correct predictions to the total data.

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

Precision is the ratio of true positive predictions to the total number of positive predictions.

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

Recall is the ratio of true positive predictions to the total number of actual positive instances.

$$F1\ Score = \frac{2\ x\ Recall\ x\ Presisi}{Recall + Presisi} \tag{4}$$

The F1 score is the weighted average of precision and recall.

3 Results and Discussions

3.1 Data Augmentation

This method generates a final dataset of 7582 datasets by using 12 picture data from each class. The augmentation techniques are expected to test the model's robustness against diverse data. An example of the application of augmentation techniques can be seen in Fig. 3.

(a)

(b)

Figure 3. (a) Rotation augmentation technique, (b) Rotation augmentation technique.

3.2 Data Splitting

Data utilized in the model training phase is referred to as training data or train data. Up to 80% of the entire dataset was randomly selected for the training data utilized in this study, yielding a training dataset of 6056 data in total. Data used in the models testing phase is called test data or testing data. With a final test dataset consisting of 1526 data, the test data used in this study accounted for up to 20% of the entire dataset.

3.3 Modelling ResNet-34

Before entering the modeling stage, the model is subjected to hyperparameter tuning by considering the most optimal learning rate. Analyzing the model for vanishing gradient and overfitting is done by comparing the visualization when tuning the model by comparing the values of train loss, valid loss and error rate.

This research uses the one-cycle policy optimization method developed by Leslie Smith in 2018. By modifying the learning rate to hasten convergence and avoid overfitting, this technique seeks to train the model effectively [33]. The best weight decay value is first 1e-6 (0.000001) and eventually 1e1 (10). Fig 4 illustrates how to use the learning rate finder approach to get the ideal learning rate. The process of finding the optimal learning rate using the learning rate finder method is shown in Fig. 4.

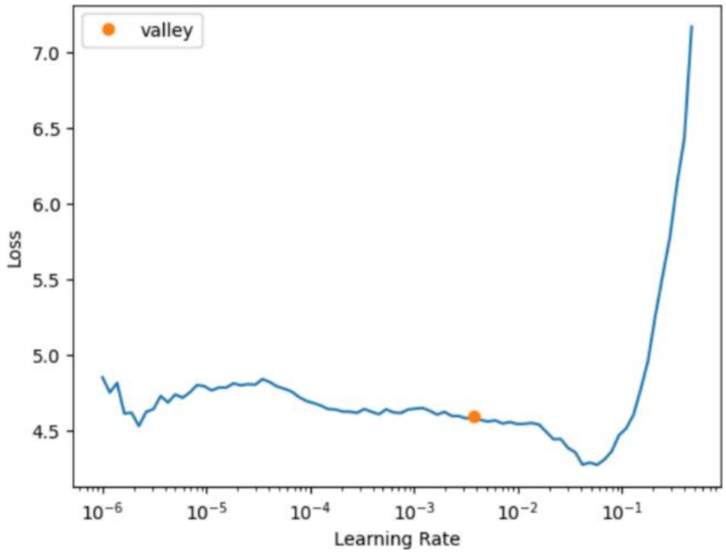
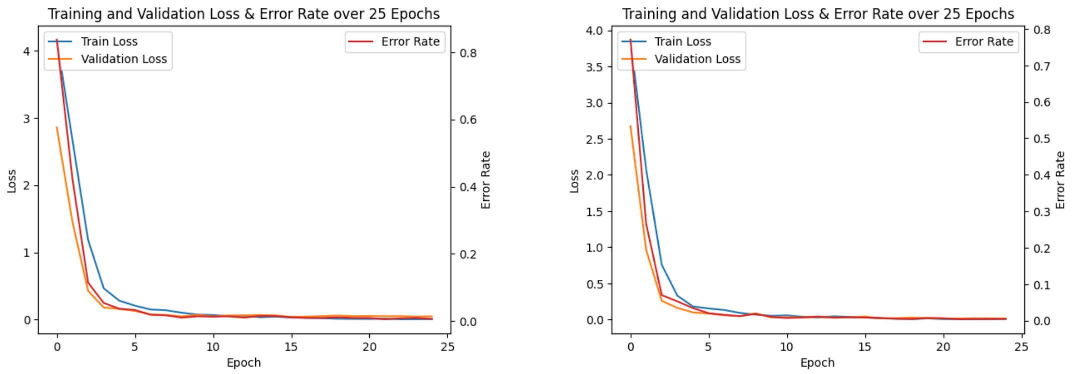


Figure 4. The Most Optimal Learning Rate Value

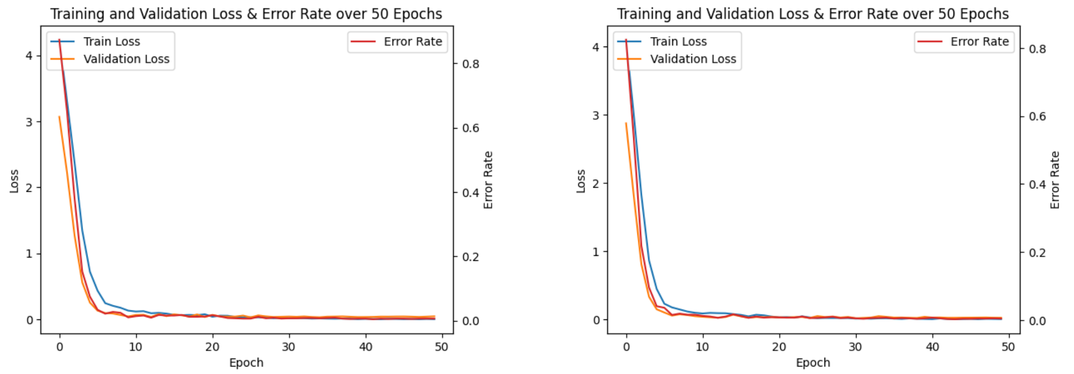
The Learning Rate Finder produced the curve shown in Fig. 3, where loss is shown on the y-axis and learning rate is plotted on the x-axis. The "valley" point on the loss curve denotes the ideal learning rate and is the lowest point. The optimal point shown as "valley" in the figure, is used to determine the lr_max value. Based on Fig. 3. Lr_max value of about 1.13×10-3 was chosen, as it shows a sharp increase in loss after the optimal point, indicating it is the best before the loss increases again.

Visualizations of train loss, valid loss and error rate at 25, 50 and 100 epochs are shown in Fig. 5 shows the training using 25 epochs, (a) without augmentation (b) with augmentation, Fig. 6 shows the training using 50 epochs, (a) without augmentation (b) with augmentation, Fig. 7 shows the training using 100 epochs, (a) without augmentation (b) with augmentation.



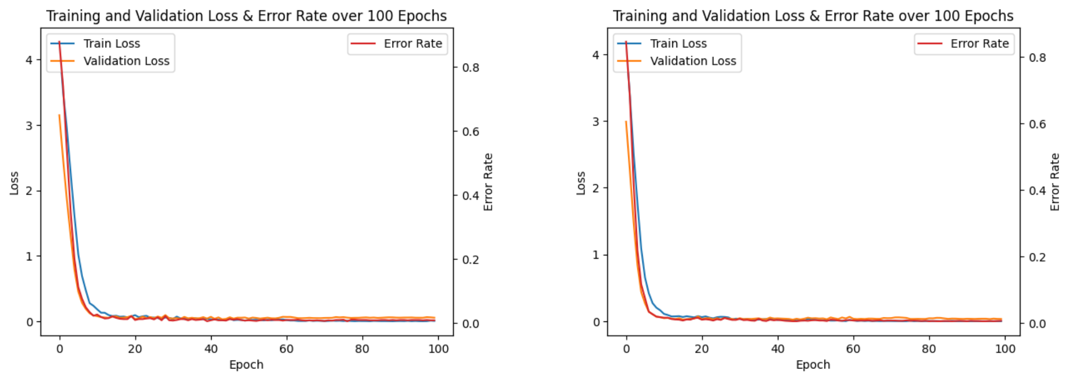
(a) (b)

Figure 5. Training and validation loss and error rate at 25 epoch for (a) without augmentation and (b) with augmentation



(a) (b)

Figure 6. Training and validation loss and error rate at 50 epoch for (a) without augmentation and (b) with augmentation



(a) (b)

Figure 7. Training and validation loss and error rate at 50 epoch for (a) without augmentation and (b) with augmentation

Comparison of visualization train loss, valid loss, and error rate at 25, 50, and 100 epochs is shown in Table 2. From the results of the comparison value in Table 2. It can be concluded that the model using augmentation techniques at 50 epochs is the best performing model by looking at the results of train loss and valid loss which tend to be balanced, indicating that the model can generalize new data well and the model also does not occur overfitting.

3.4 Model Evaluation

To determine the ideal model combination, model evaluation is required. A matrix can be used to evaluate a model by taking into account many factors, including accuracy and confusion matrix findings. A visual assessment tool for machine learning is the confusion matrix [31]. To calculate all potential cases of classification difficulties, the confusion matrix's columns reflect the predicted class results, while its rows represent the actual class results [34]. The calculation results and Comparison of precision, recall, f1-score, and accuracy is shown using confusion matrix are shown in Fig. 8.

Table 2. Comparison of Visualization Result Values

Variable	Without augmen. (25 epoch)	With augmen. (25 epoch)	Without augmen. (50 epoch)	With augmen. (50 epoch)	Without augmen. (100 epoch)	With augmen. (100 epoch)
Train loss	0.008626	0.007812	0.002607	0.002599	0.008523	0.006596
Valid loss	0.048414	0.014932	0.045147	0.021599	0.054004	0.033828
Error rate	0.006629	0.004617	0.006629	0.004617	0.005120	0.000175

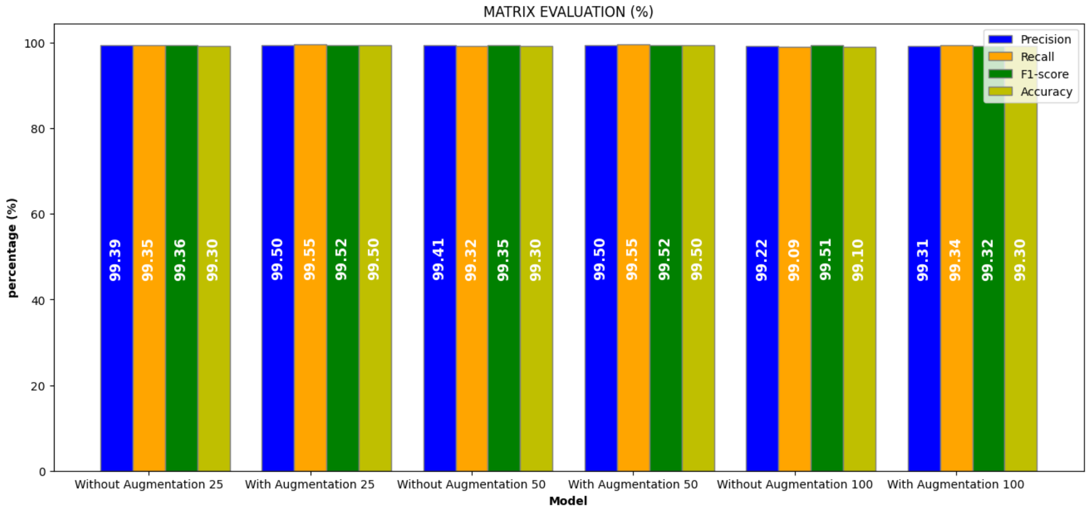
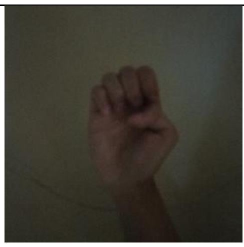


Figure 8. Visualization of Confusion Matrix

3.5 Error Analysis

Error analysis was conducted to identify limitations in the classification model. This analysis aimed to identify factors causing the model to fail in making predictions, particularly for letters of the alphabet that are similar in shape. Examples of images that were incorrectly classified by the model are shown in Table 3.

Table 3. Misclassified Alphabet Images

No.	Techniques	Actual Label	Predicted Label
1		E	S
2		N	S

3		C	X

Table 3 shows cases where the model failed to classify image objects. Most classification errors occurred because identical alphabet letters in the images indicated that the model tended to make mistakes when distinguishing between similar object structures. Minimal lighting on objects also affected the generalisation of the model's classification results.

The results of this analysis indicate that the model still struggles to generalise under low-light conditions, with objects that have high noise levels, and characters with similar shapes. To address these shortcomings, the model can be improved by adding more representative training data for alphabet classes with identical shapes and enhancing the image quality of the dataset.

3.6 Benchmarking with Previous Studies

The model evaluation was conducted by comparing the performance of the developed ResNet-34 model with the results of other relevant research models. The model evaluation is shown in Table 4.

Table 4. Benchmarking Model Results

Model / Study	Dataset Type	Architecture	Accuracy	Augmentation	Year
Arisandi et al. 2022 [6]	BISINDO	CNN 5-layer	93.00%	No	2022
Niswati et al. [11]	Cervical cancer image	ResNet-50	91.00%	No	2022
Ridhovan et al. 2022 [31]	Leaf disease	ResNet-152V2	95.00%	Yes	2022

This study	SIBI (24 classes)	ResNet-34	**99.50%**	Yes	2025

Based on the model comparison results in Table 4, the developed model shows competitive accuracy compared to previous studies. These comparison results show that the use of augmentation techniques on the dataset contributes significantly to improving model performance.

4 Conclusions

The best model performance CNN was found in the 50 epoch process using augmentation techniques with results precision 99.5%, recall 99.5%, F1-score 99.5%, and accuracy 99.5%. From these results it can be concluded that the ResNet-34 architecture used by the CNN algorithm effectively prevents vanishing gradient, and image augmentation approaches effectively prevent overfitting and increase the accuracy of the final model. The system could be a communication aid for the general public, especially those with speech and hearing impairments, provided it undergoes more real-world testing.

Acknowledgements

The author would like to thanks Allah SWT, Siliwangi University, family, friends and lectures gave me provided guidance and motivation in carrying out this research.

References

[1] Saleha and M. R. Yuwita, A. [Nama Penulis], “Semiotic Analysis of Dead End Traffic Sign Symbols,” *MAHADAYA: Jurnal Ilmu Komunikasi*, vol. 3, no. 1, pp. 65-72], [2023]. [Online]. Available: https://ojs.unikom.ac.id/index.php/mahadaya/article/view/7886

[2] F. S. Pandiangan and M. Rosadi, “Analisis Dialek Dalam Bentuk Bahasa Percakapan Dalam Film ‘Imperfect’ Karya Meira Anastasia,” *Journal of*

Educational Research and Humaniora (JERH), vol. 1, no. September, pp. 47–58, 2023.

[3] Nasha Hikmatia A.E. and M. I. Zul, "Aplikasi Penerjemah Bahasa Isyarat Indonesia menjadi Suara berbasis Android menggunakan Tensorflow," *Jurnal Komputer Terapan*, vol. 7, no. 1, pp. 74–83, 2021, doi: 10.35143/jkt.v7i1.4629.

[4] I. Sari, Fivrenodi, E. Altiarika, and Sarwindah, "Sistem Pengembangan Bahasa Isyarat Untuk Berkomunikasi dengan Penyandang Disabilitas (Tunarungu)," *Journal of Information Technology and society*, vol. 1, no. 1, pp. 20–25, 2023, doi: 10.35438/jits.v1i1.21.

[5] Nofal Anam, "Sistem Deteksi Simbol Pada Sibi (Sistem Isyarat Bahasa Indonesia) Menggunakan Mediapipe Dan ResNet-50," 2022.

[6] L. Arisandi and B. Satya, "Sistem Klarifikasi Bahasa Isyarat Indonesia (Bisindo) Dengan Menggunakan Algoritma Convolutional Neural Network," *Jurnal Sistem Cerdas*, vol. 5, no. 3, pp. 135–146, 2022, doi: 10.37396/jsc.v5i3.262.

[7] G. Latif, D. A. Alghmgham, R. Maheswar, J. Alghazo, F. Sibai, and M. H. Aly, "Deep learning in Transportation: Optimized driven deep residual networks for Arabic traffic sign recognition," *Alexandria Engineering Journal*, vol. 80, no. July 2022, pp. 134–143, 2023, doi: 10.1016/j.aej.2023.08.047.

[8] S. Mekruksavanich, N. Hnoohom, and A. Jitpattanakul, "A Hybrid Deep Residual Network for Efficient Transitional Activity Recognition Based on Wearable Sensors," *Applied Sciences (Switzerland)*, vol. 12, no. 10, 2022, doi: 10.3390/app12104988.

[9] P. A. Pattanaik, M. Mittal, M. Z. Khan, and S. N. Panda, "Malaria detection using deep residual networks with mobile microscopy," *Journal of King Saud University*

- *Computer and Information Sciences*, vol. 34, no. 5, pp. 1700–1705, 2022, doi: 10.1016/j.jksuci.2020.07.003.

[10] B. Tasci, M. R. Acharya, M. Baygin, S. Dogan, T. Tuncer, and S. B. Belhaouari, "InCR: Inception and concatenation residual block-based deep learning network for damaged building detection using remote sensing images," *International Journal of Applied Earth Observation and Geoinformation*, vol. 123, no. August, p. 103483, 2023, doi: 10.1016/j.jag.2023.103483.

[11] Z. Niswati, R. Hardatin, M. N. Muslimah, and S. N. Hasanah, "Perbandingan Arsitektur ResNet50 dan ResNet101 dalam Klasifikasi Kanker Serviks pada Citra Pap Smear," *Faktor Exacta*, vol. 14, no. 3, p. 160, 2021, doi: 10.30998/faktorexacta.v14i3.10010.

[12] H. Imaduddin, F. Y. A'la, A. Fatmawati, and B. A. Hermansyah, "Comparison of transfer learning method for COVID-19 detection using convolution neural network," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 2, pp. 1091–1099, Apr. 2022, doi: 10.11591/eei.v11i2.3525.

[13] X. Ma, W. Chen, and Y. Xu, "ERCP-Net: a channel extension residual structure and adaptive channel attention mechanism for plant leaf disease classification network," *Sci Rep*, vol. 14, no. 1, pp. 1–14, 2024, doi: 10.1038/s41598-024-54287-3.

[14] Z. Bin Niu, S. Y. Jia, and H. H. Xu, "Automated graptolite identification at high taxonomic resolution using residual networks," *iScience*, vol. 27, no. 1, p. 108549, 2024, doi: 10.1016/j.isci.2023.108549.

[15] D. Sarwinda, R. H. Paradisa, A. Bustamam, and P. Anggia, "Deep Learning in Image Classification using Residual Network (ResNet) Variants for Detection of

Colorectal Cancer," *Procedia Comput Sci*, vol. 179, no. 2019, pp. 423–431, 2021, doi: 10.1016/j.procs.2021.01.025.

[16] L. Ali and S. A. C. Bukhari, "An Approach Based on Mutually Informed Neural Networks to Optimize the Generalization Capabilities of Decision Support Systems Developed for Heart Failure Prediction," *IRBM*, vol. 42, no. 5, pp. 345–352, 2021, doi: https://doi.org/10.1016/j.irbm.2020.04.003.

[17] M. A. A. Fawwaz, K. N. Ramadhani, and F. Sthevani, "Klasifikasi Ras pada hewan peliharaan menggunakan Algoritma Convolutional Neural Network (CNN)," vol. 8, no. 1, pp. 715–730, 2020.

[18] Rima Dias Ramadhani, A. Nur Aziz Thohari, C. Kartiko, A. Junaidi, T. Ginanjar Laksana, and N. Alim Setya Nugraha, "Optimasi Akurasi Metode Convolutional Neural Network untuk Identifikasi Jenis Sampah," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 2, pp. 312–318, Apr. 2021, doi: 10.29207/resti.v5i2.2754.

[19] J. Sanjaya and M. Ayub, "Augmentasi Data Pengenalan Citra Mobil Menggunakan Pendekatan Random Crop , Rotate , dan Mixup," vol. 6, pp. 311–323, 2020.

[20] Y. Vita Via, I. Yuniar Purbasari, and A. Putra Pratama, "Analisa Algoritma Convolution Neural Network (Cnn) Pada Klasifikasi Genre Musik Berdasar Durasi Waktu," *SCAN Jurnal Teknologi dan Informasi*, vol. 17, no. 1, pp. 35–41, 2022, [Online]. Available: http://ejournal.upnjatim.ac.id/index.php/scan/article/view/3251/2003

[21] R. Z. Fadillah, A. Irawan, M. Susanty, and I. Artikel, "Data Augmentasi Untuk Mengatasi Keterbatasan Data Pada Model Penerjemah Bahasa Isyarat Indonesia (BISINDO)," *Jurnal Informatika*, vol. 8, no. 2, pp. 208–214, 2021, [Online]. Available: https://ejournal.bsi.ac.id/ejurnal/index.php/ji/article/view/10768

[22] N. E. Khalifa, M. Loey, and S. Mirjalili, "A comprehensive survey of recent trends in deep learning for digital images augmentation," *Artif Intell Rev*, vol. 55, no. 3, pp. 2351–2377, 2022, doi: 10.1007/s10462-021-10066-4.

[23] W. M. Pradnya D and A. P. Kusumaningtyas, "Analisis Pengaruh Data Augmentasi Pada Klasifikasi Bumbu Dapur Menggunakan Convolutional Neural Network," *Jurnal Media Informatika Budidarma*, vol. 6, no. 4, p. 2022, 2022, doi: 10.30865/mib.v6i4.4201.

[24] D. Putri Ayuni, Jasril, M. Irsyad, F. Yanto, and S. Sanjaya, "Augmentasi Data Pada Implementasi Convolutional Neural Network Arsitektur Efficientnet-B3 Untuk Klasifikasi Penyakit Daun Padi," *ZONAsi: Jurnal Sistem Informasi*, vol. 5, no. 2, pp. 239–249, 2023, doi: 10.31849/zn.v5i2.13874.

[25] T. B. Sasongko, H. Haryoko, and A. Amrullah, "Analisis Efek Augmentasi Dataset dan Fine Tune pada Algoritma Pre-Trained Convolutional Neural Network (CNN)," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 10, no. 4, pp. 763–768, 2023, doi: 10.25126/jtiik.20241046583.

[26] Hendri Candra Mayana and Desmarita Leni, "Deteksi Kerusakan Ban Mobil Menggunakan Convolutional Neural Network dengan Arsitektur ResNet-34," *Jurnal Surya Teknika*, vol. 10, no. 2, pp. 842–851, 2023, doi: 10.37859/jst.v10i2.6336.

[27] N. I. Sanusi, S. Ramadhani, and M. Irsyad, "Analisa Gambar X-Ray Mammography dengan Convolution Neural Network pada Deep Learning dengan Arsitektur Resnet," *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 4, p. 604, 2023, doi: 10.30865/json.v4i4.6365.

[28] M. F. Gunardi, "Implementasi Augmentasi Citra pada Suatu Dataset," *Jurnal Informatika*, vol. 9, no. 1, pp. 1–5, 2023.

[29] W. Maulana Baihaqi, C. Raras, A. Widiawati, D. P. Sabila, and A. Wati, “Analisis Gambar Sel Darah Berbasis Convolution Neural Network untuk Mendiagnosis Penyakit Demam Berdarah Convolution Neural Network-Based Image Analysis of Blood Cells to Diagnose Dengue Fever,” *Cogito Smart Journal* |, vol. 7, no. 1, 2021.

[30] R. Cendekia Vandara, S. A. Wibowo, and K. Usman, “Performance Analysis of Face Alignment On 3-Dimensional (3D) Face Reconstruction Using Modified Position Map Regression Network.”Thesis, Telkom University, 2021.

[31] A. Ridhovan and A. Suharso, “Penerapan Metode Residual Network (Resnet) Dalam Klasifikasi Penyakit Pada Daun Gandum,” *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 7, no. 1, pp. 58–65, 2022, doi: 10.29100/jipi.v7i1.2410.

[32] I. Ariawan *et al.*, “Extraction of Morphometric Features the shape of mangrove leaves based on digital images and classification using the Support Vector Machine,” *Karbala International Journal of Modern Science*, vol. 10, no. 2, May 2024, doi: 10.33640/2405-609X.3349.

[33] L. N. Smith, “A disciplined approach to neural network hyper-parameters: Part 1 -- learning rate, batch size, momentum, and weight decay,” Mar. 2018, [Online]. Available: http://arxiv.org/abs/1803.09820

[34] J. Xu, Y. Zhang, and D. Miao, “Three-way confusion matrix for classification: A measure driven view,” *Inf Sci (N Y)*, vol. 507, pp. 772–794, 2020, doi: https://doi.org/10.1016/j.ins.2019.06.064.