# Detection of Indonesian Sign Language System using Convolutional Neural Network (CNN) with Nadam Optimizer

Alamsyah Alamsyah[1*] and Dinda Ayu Anggraeni [2]

[1,2] Department of Computer Science,
Universitas Negeri Semarang, Indonesia
[1] alamsyah@mail.unnes.ac.id

**Abstract.** This study enhances sign language detection by utilizing a Convolutional Neural Network (CNN) modified with the Nadam optimizer, resulting in improved accuracy in recognizing sign language images. The research begins by downloading the Indonesian Sign Language dataset from previous studies, which is then split into 80% for training and 20% for testing. Pre-processing is applied to ensure the dataset is compatible with the CNN algorithm. The dataset is then trained using a CNN model optimized with Nadam. After training, the model is tested, achieving an accuracy of 96.88% in detecting images from the Indonesian Sign Language system. These results show improved accuracy compared to previous studies. It is hoped that the findings of this study will benefit the deaf community, particularly in the learning process at special education schools in Indonesia.

**Keywords:** Indonesian Sign Language, CNN, Nadam Optimizer, Deaf, Speech impaired.

## 1 Introduction

Communication is a fundamental human need, and for individuals with hearing impairments, sign language serves as an essential tool [1]. Unlike spoken language, which relies on sound and verbal cues, sign language is purely visual. Deaf individuals communicate through gestures using their hands, arms, facial expressions, and other body parts like the shoulders and eyebrows [2]. However, due to the significant differences between spoken and sign languages, integrating the deaf community into society remains a challenge [3], [4]. While deaf individuals rely on their visual abilities, their capacity to learn written language, which is a visual representation of spoken language, may differ from that of hearing individuals [5].

In sign language, a single gesture can have multiple meanings depending on factors such as hand shape, movement, and the location of the gesture (e.g., near the chin


N. A. S. Abdullah et al. (eds.), *Proceedings of the International Conference on Innovation & Entrepreneurship in Computing, Engineering & Science Education (InvENT 2024)*, Advances in Computer Science Research 117,
https://doi.org/10.2991/978-94-6463-589-8_32

or the eyes) [6]. Additional elements, like posture, facial expressions, rhythm, and the speed of hand movements, also contribute to variations in meaning.

Each country has its own sign language [7]. For example, American Sign Language (ASL) is used in the United States, while British Sign Language (BSL) is used in Great Britain [8]. In Indonesia, the official sign language is the Indonesian Sign Language System (ISLS), which has spurred the development of ISLS recognition technologies to aid communication within the deaf community [9].

The detection of Indonesian Sign Language (ISLS) using algorithms such as Convolutional Neural Networks (CNN) has emerged as a crucial area of research aimed at improving communication and accessibility for the deaf community. CNN, a type of neural network architecture, is widely recognized for its effectiveness in pattern recognition and object classification [10]. When applied to ISLS detection, CNN can identify hand gestures, facial expressions, and body movements that correspond to ISLS symbols. By automatically learning complex features from images, CNN allows for more accurate recognition of ISLS symbols [11]. However, previous studies using CNN for ISLS detection have encountered challenges such as suboptimal model performance and the use of non-representative datasets. Poorly optimized CNN models may result in low accuracy, and non-representative datasets may fail to capture the variability in hand movements or the contextual usage of ISLS. Additionally, the development of a real-time ISLS alphabet recognition system using CNN remains an ongoing area of research [12].

This research aims to address the limitations of ISLS detection by employing a modified CNN approach with the Nadam optimizer. These limitations include difficulties in accurately classifying signs and recognizing visually similar signs. Nadam, which combines the Adam and Nesterov Accelerated Gradient algorithms, is expected to enhance CNN performance in ISLS detection. This study seeks to improve ISLS detection accuracy compared to previous work by leveraging Nadam to make the model more efficient in feature extraction from ISLS images and more adaptive during the training process [13].

Moreover, this research has significant implications for improving accessibility and communication for the deaf community in Indonesia. By increasing the accuracy of ISLS detection, the technology developed in this study could become a more reliable and useful tool for supporting daily communication for deaf individuals. Beyond its technical contributions, this research has a meaningful social impact by advancing technology that empowers individuals with special needs, ultimately enhancing their quality of life. Therefore, this research has the potential to bring wide-reaching positive benefits to society, especially for those with special needs.

# 2 Methods

The method used in this research starts by downloading the ISLS dataset from prior studies, which includes both training and validation data. The dataset is then used to train a CNN model, optimized using the Nadam optimizer. A loss function is integrated to evaluate the model's performance in detecting sign language. During the training process, graphs tracking the model's accuracy in learning from the training and validation data are generated and analyzed.

## 2.1 Data Collection and Preprocessing

The data used in this study comes from the "ISLS Alphabet" dataset, available on the Kaggle platform. This dataset contains 2,600 images of the Indonesian Sign Language (ISLS) alphabet, which serve as input for training and testing the alphabet recognition model.

Data preprocessing is crucial to prepare the raw data for model building [14]. First, the image intensity values are rescaled to a range of 0 to 1, ensuring consistency in pixel values across the dataset. Next, the data undergoes augmentation, using techniques such as zooming (with a range of 0.1) and horizontal flipping. These techniques enhance data variability, allowing the model to learn from more diverse perspectives. Finally, the dataset is divided into two parts: 80% for training and 20% for validation. This split allows for independent evaluation of the model's performance. The results of the preprocessing steps are illustrated in Figure 1.

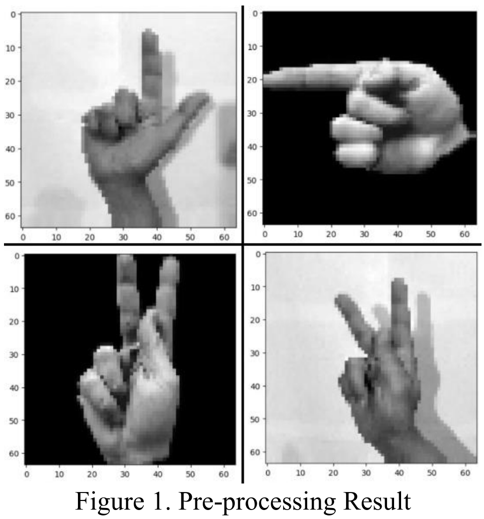

Figure 1. Pre-processing Result

## 2.2 The CNN Model with Nadam Optimizer

The CNN model is designed with an architecture consisting of several types of layers:

1. Convolutional Layers: These layers extract important features from the input images by applying convolution operations.
2. Max Pooling Layers: These layers reduce the spatial dimensions of the image by selecting the maximum value from each defined region, thus down sampling the data.
3. Dropout Layers: These layers help prevent overfitting by randomly deactivating some neuron units during training, encouraging the network to learn more general features.
4. Dense Layers: These fully connected layers aim to classify the data based on the features extracted by the previous layers, connecting each neuron in the previous layers to each neuron in the dense layer.

After constructing the CNN model, it is compiled using the Nadam optimizer. The optimizer is the algorithm used to minimize the loss function by adjusting the weights in each layer of the model. It helps the model find the optimal combination of weights during training. The Nadam optimizer is a variant of the Adam optimizer that incorporates elements of Nesterov momentum. Adam itself is one of the most popular optimizers because it combines the benefits of two other optimizers: RMSProp and Stochastic Gradient Descent with Momentum (SGDM). Nadam can be applied during model compilation as shown in Figure 2.

```
from keras.optimizers import Nadam
model.compile(optimizer=Nadam(learning_rate=0.001),
              loss='categorical_crossentropy',
              metrics=['accuracy'])
```

Figure 2. CNN with Nadam Optimizer

### 2.3 Model Training and Evaluation

Once compiled, the CNN model is trained using the training dataset. The Nadam optimizer is utilized to minimize the loss function across 50 training epochs. During this process, the model's weights and parameters are fine-tuned, enabling it to learn patterns from the training data and generalize to new, unseen data [15].

Following the training phase, the model's performance is assessed by comparing its predictions against the actual labels in the validation dataset [16]. The primary evaluation metric used is accuracy, which calculates the proportion of correct predictions out of the total predictions made.

## 3 Results and Discussion

This study successfully developed and trained a Convolutional Neural Network (CNN) model using the Nadam optimizer to detect signs from the Indonesian Sign Language System (ISLS). The model demonstrated strong performance, achieving a testing accuracy of 96.88%. Over the course of 50 epochs, the training and testing processes showed a clear increase in accuracy and a significant decrease in loss for both the training and testing datasets.

The consistent improvement in accuracy throughout the training and testing phases, along with the stable decrease in loss, indicates that the model does not suffer from overfitting and achieves good convergence. The model accurately recognized most signs, although certain sign classes remain challenging, likely due to visual similarities between some signs or variations in the way the signs are performed.

The combination of CNN and the Nadam optimizer has proven effective in detecting and classifying ISLS signs. The CNN architecture, which includes convolutional layers, max pooling, and dropout, efficiently extracts important features from the sign images. Additionally, the ReLU activation function aids in learning complex nonlinear patterns from the image data [17].

The accuracy of the CNN model with the Nadam optimizer is compared to other methods, as shown in Table 1. The table presents a performance comparison between the proposed model and several previous studies.

Table 1 compares the accuracy of the proposed model with other studies

| References | Model | Accuracy |
|---|---|---|
| **Atitallah et al., 2022** [18] | CNN | 95.94% |
| **Montiel et al., 2022** [19] | SVM | 96.6% |
| **Dima & Ahmed, 2021**[20] | YOLOv5 | 95% |
| **Sincan et al., 2020** [21] | Baseline Methods | 96.11% |
| **Liao et al., 2019** [22] | BLSTM-3D Residual Networks | 89.8% |
| **Proposed Model** | CNN with Nadam Optimizer | 96.88% |

Table 1 presents a comparison of accuracy results between the proposed model and those from other studies. The improvements achieved by the Indonesian Sign Language System (ISLS) detection model in this research surpass previous studies. By using a combination of the CNN method and the Nadam optimizer, this model achieved higher accuracy in image recognition.

Although the model performs well in classifying ISLS signs, there is still potential for improvement. For instance, using a larger and more diverse dataset or applying more advanced data augmentation techniques could enhance performance. Additionally, future research could explore techniques like Transfer Learning or implement deeper model architectures to improve model accuracy [23-25]. This research provides a strong foundation for further development in sign language recognition, which could significantly enhance accessibility and communication for the deaf community, particularly in the teaching processes of special education schools in Indonesia.

Finally, implementing the system on embedded hardware with various methods is an important step towards practical applications [26-28]. By deploying the system on embedded devices, real-time ISLS recognition applications can be developed, allowing the deaf community to use this technology in their daily lives. Further research in this area will ensure that the ISLS detection system is not only effective in controlled environments but also suitable for real-world scenarios.

# 4 Conclusion

The use of a Convolutional Neural Network (CNN) optimized with Nadam has significantly improved the detection of Indonesian Sign Language (ISLS). The CNN model trained with the Nadam optimizer achieved an accuracy of 96.88%. Accuracy consistently improved throughout the training and testing processes, indicating good model convergence. While the model accurately classified most ISLS signs, challenges remain in recognizing certain sign classes. Nevertheless, this study provides a solid foundation for further advancements in sign language recognition, potentially enhancing communication and accessibility for the deaf community, particularly in special education schools in Indonesia.

# References

1. M. Boondamnoen, K. Thongsri, T. Sahabantoegnsin, and K. Woraratpanya, “Exploring LSTM and CNN Architectures for Sign Language Translation,” in 2023 15th International Conference on Information Technology and Electrical Engineering, ICITEE 2023, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 198–203. doi: 10.1109/ICITEE59582.2023.10317660.
2. C. Singh, D. Sharma, A. P. Dubey, and N. Tyagi, “Sign Language Detection Using CNN-YOLOv8l,” in 2023 International Conference on Advances in Computation, Communication and Information Technology, ICAICCIT 2023, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 12–17. doi: 10.1109/ICAICCIT60255.2023.10465792
3. M. Zhang, S. Yang, and M. Zhao, “Deep Learning-Based Standard Sign Language Discrimination,” IEEE Access, vol. 11, pp. 125822–125834, 2023, doi: 10.1109/ACCESS.2023.3330863.
4. A. Kumar, S. Utekar, and D. Viji, “Hand Sign Gesture Recognition and Translator Using Convolution Neural Networks (CNN) and k-Nearest Neighbors (k-NN),” in International Conference on Distributed Computing and Optimization Techniques, ICDCOT 2024, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/ICDCOT61034.2024.10515667.
5. T. Janbandhu, S. Laddhani, R. Agrawal, C. Dhule, N. Chavhan, and A. Khandekar, “Sign Language Recognition Using CNN,” in 2023 International Conference on Communication,

Security and Artificial Intelligence, ICCSAI 2023, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 175–180. doi: 10.1109/ICCSAI59793.2023.10421183.

6. A. Deshpande, A. Shriwas, V. Deshmukh, and S. Kale, "Sign Language Recognition System using CNN," in Proceedings of the International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics, ICIITCEE 2023, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 906–911. doi: 10.1109/IITCEE57236.2023.10091051.
7. F. Shah, M. S. Shah, W. Akram, A. Manzoor, R. O. Mahmoud, and D. S. Abdelminaam, "Sign Language Recognition Using Multiple Kernel Learning: A Case Study of Pakistan Sign Language,"
8. P. C. Pruthvi and J. Geetha, "Convolution Neural Network for Predicting Alphabet Sign Language and Comparative Performance Analysis of CNN, KNN, and SVM Algorithms," in 2022 IEEE 3rd Global Conference for Advancement in Technology, GCAT 2022, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/GCAT55367.2022.9972174.
9. R. E. Caraka *et al.*, "Empowering deaf communication: a novel LSTM model for recognizing Indonesian sign language," *Univers Access Inf Soc*, Mar. 2024, doi: 10.1007/s10209-024-01095-1.
10. W. Aly, S. Aly, and S. Almotairi, "User-independent american sign language alphabet recognition based on depth image and PCANet features," *IEEE Access*, vol. 7, pp. 123138–123150, 2019, doi: 10.1109/ACCESS.2019.2938829.
11. B. V. Chowdary, A. P. Thota, A. Sreeja, K. N. Reddy, and K. S. Chandana, "Sign Language Detection and Recognition using CNN," in *International Conference on Sustainable Computing and Smart Systems, ICSCSS 2023 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 553–557. doi: 10.1109/ICSCSS57650.2023.10169225.
12. A. N. Sihananto, E. M. Safitri, Y. Maulana, F. Fakhruddin, and M. E. Yudistira, "Indonesian Sign Language Image Detection Using Convolutional Neural Network (CNN) Method," *Inspiration: Jurnal Teknologi Informasi dan Komunikasi*, vol. 13, no. 1, pp. 13–21, May 2023, doi: 10.35585/inspir.v13i1.37.
13. N. Nurrahma, R. Yusuf, and A. S. Prihatmanto, "Indonesian Sign Language Fingerspelling Recognition using Vision-based Machine Learning," in *2021 International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA)*, IEEE, Dec. 2021, pp. 28–33. doi: 10.1109/ICICyTA53712.2021.9689176.
14. V. Çetin and O. Yıldız, "A comprehensive review on data preprocessing techniques in data analysis," *Pamukkale University Journal of Engineering Sciences*, vol. 28, no. 2, pp. 299–312, 2022, doi: 10.5505/pajes.2021.62687.
15. R. Moradi, R. Berangi, and B. Minaei, "A survey of regularization strategies for deep models," *Artif Intell Rev*, vol. 53, no. 6, pp. 3947–3986, Aug. 2020, doi: 10.1007/s10462-019-09784-7.
16. B. T. Pham *et al.*, "Performance Evaluation of Machine Learning Methods for Forest Fire Modeling and Prediction," *Symmetry (Basel)*, vol. 12, no. 6, p. 1022, Jun. 2020, doi: 10.3390/sym12061022.
17. Y. Yu, K. Adu, N. Tashi, P. Anokye, X. Wang, and M. A. Ayidzoe, "RMAF: Relu-Memristor-Like Activation Function for Deep Learning," *IEEE Access*, vol. 8, pp. 72727–72741, 2020, doi: 10.1109/ACCESS.2020.2987829.
18. B. Ben Atitallah *et al.*, "Hand Sign Recognition System Based on EIT Imaging and Robust CNN Classification," *IEEE Sens J*, vol. 22, no. 2, pp. 1729–1737, Jan. 2022, doi: 10.1109/JSEN.2021.3130982.

19. E. Pereira-Montiel *et al.*, "Automatic sign language recognition based on accelerometry and surface electromyography signals: A study for Colombian sign language," *Biomed Signal Process Control*, vol. 71, Jan. 2022, doi: 10.1016/j.bspc.2021.103201.
20. T. F. Dima and M. E. Ahmed, "Using YOLOv5 Algorithm to Detect and Recognize American Sign Language," in *2021 International Conference on Information Technology, ICIT 2021 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., Jul. 2021, pp. 603–607. doi: 10.1109/ICIT52682.2021.9491672.
21. O. M. Sincan and H. Y. Keles, "AUTSL: A large scale multi-modal Turkish sign language dataset and baseline methods," *IEEE Access*, vol. 8, pp. 181340–181355, 2020, doi: 10.1109/ACCESS.2020.3028072.
22. Y. Liao, P. Xiong, W. Min, W. Min, and J. Lu, "Dynamic Sign Language Recognition Based on Video Sequence with BLSTM-3D Residual Networks," *IEEE Access*, vol. 7, pp. 38044–38054, 2019, doi: 10.1109/ACCESS.2019.2904749.
23. S. Tulasi Krishna, "Deep Learning and Transfer Learning Approaches for Image Classification." [Online]. Available: https://www.researchgate.net/publication/333666150
24. A. Aljabar and S. Suharjito, "BISINDO (Bahasa Isyarat Indonesia) Sign Language Recognition Using CNN and LSTM," *Advances in Science, Technology and Engineering Systems Journal*, vol. 5, no. 5, pp. 282–287, 2020, doi: 10.25046/aj050535.
25. M. Arthur Limantara and D. Tristianto, "SIBI Alphabet Detection System Based on Convolutional Neural Network (CNN) Method as Learning Media," *Internet of Things and Artificial Intelligence Journal*, vol. 4, no. 1, pp. 143–161, Mar. 2024, doi: 10.31763/iota.v4i1.716.
26. R. Setiawan, Y. Yunita, F. F. Rahman, and H. Fahmi, "BISINDO (Bahasa Isyarat Indonesia) Sign Language Recognition Using Deep Learning," IT FOR SOCIETY, vol. 09, no. 01.
27. Prasetiyo, B., Alamsyah, Hakim, M.F.A., Jumanto, Adi, M.H., "Differential augmentation data for vehicle classification using convolutional neural network," AIP Conference Proceedings, 2023, 2614, 040001, doi: 10.1063/5.0126720.
28. Alamsyah, Prasetiyo, B., Hakim, M.F.A., Pradana, F.D., "The improvement of COVID-19 prediction accuracy using optimal parameters in reccurent neural network model," AIP Conference Proceedings, 2023, 2614, 040019, doi: 10.1063/5.0125767