ILKOM Jurnal Ilmiah Vol. 17, No. 2, August 2025, pp. 98-106
Accredited 2nd by KEMDIKTISAINTEK No. 10/C/C3/DT.05.00/2025; E-ISSN 2548-7779 | P-ISSN 2087-1716

Research Article **Open Access (CC–BY-SA)**

SqueezeNet Image Embedding and Support Vector Machine for Recognizing Hand Gestures in Indonesian Sign Language System

Megan Shahra Islami [a,1]; Eko Rudiawan Jamzuri [a,2,*]
[a] *Department of Electrical Engineering, Politeknik Negeri Batam, Jl. Ahmad Yani, Kel. Teluk Tering, Kec. Batam Kota, Batam 29461, Indonesia*
[1] *meganshahra2828@gmail.com;* [2] *ekorudiawan@polibatam.ac.id;*
* *Corresponding author*

Article history: Received November 21, 2024; Revised April 22, 2025; Accepted July 08, 2025; Available online August 14, 2025

Abstract

This research proposes a hand gesture recognition method for the Sistem Isyarat Bahasa Indonesia (SIBI) sign language, integrating SqueezeNet for image feature extraction and Support Vector Machine (SVM) for classification. The study focuses on 24 static gestures representing alphabetic letters, excluding J and Z due to their motion-based representation. The dataset consists of 5280 RGB images (227×227 pixels), with 220 samples per gesture, obtained from a public Kaggle source. SqueezeNet, a lightweight CNN architecture, is used to generate 1000-dimensional feature vectors, which are then classified using an SVM with an RBF kernel (C = 1.0) to effectively handle non-linear decision boundaries. A 10-fold cross-validation was applied without data augmentation to evaluate baseline performance. The proposed method achieved 99.51% classification accuracy, with an average precision of 94.04%, recall of 94.02%, and F1-score of 94.02%. Certain gestures, such as G, H, and Q, achieved near-perfect recognition, while others, like V, presented greater classification challenges with a recall of 80.5%. Compared to existing models such as MobileNet (98% accuracy) and VGG16 (86% accuracy) on the same dataset, the SqueezeNet–SVM combination provides competitive or superior accuracy with significantly reduced computational requirements. These results highlight the method's potential for real-time integration into mobile or embedded sign language translation applications, bridging communication gaps between the deaf and hearing communities. Future work will focus on improving performance for difficult gestures, applying data augmentation to enhance generalization, and developing a prototype mobile application for real-world testing in relevant environments.

Keywords: Hand Gesture Recognition; SIBI; SqueezeNet; Image Embedding; Support Vector Machine.

Introduction

Verbal communication through speaking and hearing is the primary way individuals without hearing impairment interact. However, verbal communication is a big challenge for deaf people or individuals with hearing loss. Deaf people have difficulty understanding spoken speech, so they rely more on sign language as a medium to interact [1]. Sign language is a visual-spatial language that uses non-verbal communication through body movements, facial expressions, and gestures to convey meaning [2].

In Indonesia, the *Sistem Isyarat Bahasa Indonesia* (SIBI) and *Bahasa Isyarat Indonesia* (BISINDO) are the two sign languages commonly used by the deaf community [3]. Both utilize hand gestures to visualize the letters, words, or phrases they want to convey. The difference is that BISINDO uses both hands to demonstrate the sign, while SIBI only uses one hand [4]. With this sign language, deaf people can better understand the meaning conveyed by the interlocutor through the cues demonstrated by the hand gestures [5]. Although SIBI and BISINDO are considered adequate for use among the deaf community to communicate, public understanding of these two sign languages is still limited. This condition contributes to a communication gap between the deaf and hearing communities.

Various technologies have been developed to address the communication gap between society and the deaf community. For example, [6] developed a 3D animation application that visualizes hand gestures and lip movements according to the word being translated. Ref. [7] developed a specialized SIBI dictionary to assist students with hearing impairments in learning SIBI. In addition, [8] developed a sign language translator application to facilitate the public understanding of the meaning of gestures conveyed by deaf people. The communication gap can be minimized with the help of this translator application.

http://dx.doi.org/10.33096/ilkom.v17i2.2476.98-106

Sign language translator requires hand gesture detection and classification algorithms to convert an image or video into a word or sentence [9]. Convolutional Neural Network (CNN) is an algorithm commonly used for this purpose. Research [10] has developed CNN to detect hand gestures in video conversations and achieved an accuracy rate of 89%. In addition to using CNN, transfer learning on CNN also allows for hand gesture classification with better training efficiency [11]. Research [12] has implemented transfer learning to detect SIBI hand signals. The transfer learning implemented on the MobileNet and VGG16 CNN architectures achieved accuracy rates of 98% and 86%, respectively. In addition to CNN and transfer learning, feature extraction approaches combining machine learning-based classification algorithms have also been developed. For example, research [13] extracts hand landmark features using MediaPipe and classifies them using deep learning. MediaPipe is a framework that facilitates real-time processing of visual data, including the detection of 21 coordinate points of landmarks on the hand [14]. These points are used as input to the deep learning algorithm to predict the hand gesture being demonstrated.

Although CNN and transfer learning are effective in recognizing SIBI gestures, there are still some limitations to these methods. One of the limitations is the need to train the CNN on how to detect the hand gesture dataset. Training on CNN requires the use of high computational resources, even though it is done by transfer learning [15]. In addition to these limitations, previous research has not conducted a thorough evaluation of the resulting classification performance. The evaluation that has been done is only limited to measuring accuracy without considering other metrics such as precision, recall, and F1 score.

In contrast, this study proposes a lightweight approach combining SqueezeNet for feature extraction and Support Vector Machine (SVM) for classification. SqueezeNet enables compact image embedding with significantly fewer parameters, while SVM offers fast and efficient classification with minimal computational overhead. This method eliminates the need for deep network retraining and maintains competitive performance. Compared to state-of-the-art models such as MobileNet and VGG16, the proposed method demonstrates superior performance in classification metrics, highlighting its novelty and practical contribution to the development of accessible sign language translation systems.

A more detailed explanation related to this research is given below. Section 2 describes the method used to recognize SIBI sign language. Section 3 delivers the results and analysis of the research that has been done. The conclusion of this research is explained in Section 4.

Method

This research begins with the data collection and then continues with the feature extraction and classification. Finally, the results are evaluated to determine the overall performance of the proposed method. The detailed process in this research is described below.

A. Data Collection

The dataset used in this research is obtained from a public dataset sourced on the Kaggle website (https://www.kaggle.com/datasets/alvinbintang/sibi-dataset). This dataset consists of 5280 SIBI hand gesture images symbolizing 24 letters, where 220 hand gesture images represent each letter. It should be noted that the letters J and Z are not included in the dataset as these two letters are represented with hand gesture patterns, which are difficult to visualize with static images. **Figure 1** shows examples of hand gestures that symbolize the 24 letters of the alphabet in SIBI sign language.

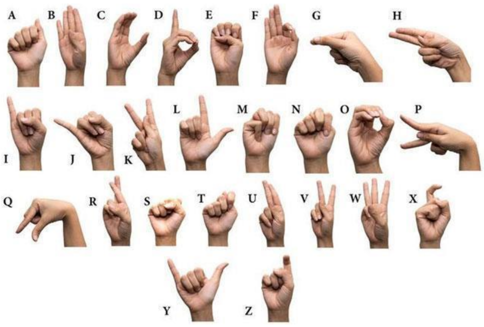

Figure 1. Hand gestures that symbolize the 24 letters of the alphabet in SIBI sign language

K-fold cross-validation with k = 10 was employed to evaluate the performance of the proposed model. This approach ensures a robust and generalized evaluation by partitioning the dataset into five equal subsets: in each iteration, four subsets are used for training and one for testing, rotating until all subsets are used as the test set once. The average performance across all folds is reported. No data augmentation techniques were applied during the experiments, as the focus was to evaluate the baseline performance of the model on the original dataset distribution..

B. Feature Extraction Using SqueezeNet Image Embedding

Feature extraction is a process for identifying and extracting significant visual information from an image, making it possible to analyze and interpret the data in a more efficient form [16]. In the computer vision context, feature extraction serves to identify distinctive features in an image, such as edges, textures, or patterns that represent objects in an image.

Features in hand gestures can be represented using various methods. One of the simple features that can represent hand gestures is the contour of the palm. Ref. [17] uses this palm contour feature and further processes it using an ensemble-based CNN to recognize the gesture. Besides contours, other features can be recognized with specific algorithms. For example, [18] uses a multiscale feature learning network to extract features in hand gestures. The feature extraction algorithm is claimed to significantly improve the accuracy of hand gesture recognition compared to traditional methods.

In this research, we extract features from images using SqueezeNet image embedding. SqueezeNet is a lightweight CNN architecture designed to reduce the number of parameters in each layer while still maintaining its accuracy [19]. This parameter reduction in SqueezeNet makes it possible to implement this architecture on limited resources computing devices, such as embedded computers or mobile phones [20].

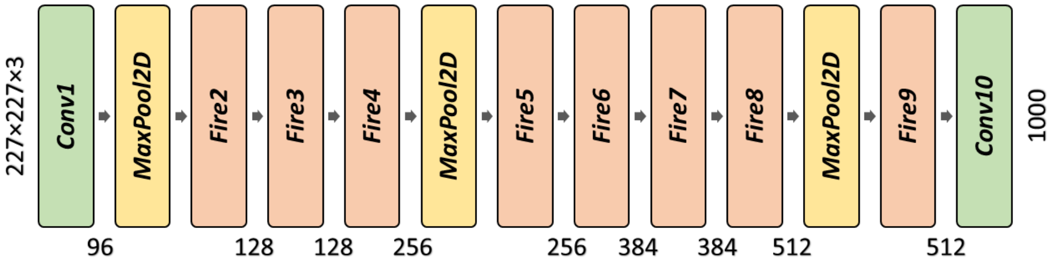

Figure 2. SqueezeNet image embedding architecture for extracting feature on hand gesture images

Figure 2 shows the overall architecture of SqueezeNet image embedding. SqueezeNet will process 227×227 hand gesture images with RGB image type through the initial convolution layer conv1. After going through the conv1 layer, the output data will be pooled to select the output with the highest value. The process is passed through several fire modules. The fire module is a squeeze layer with 1×1 convolution followed by expansion layers with 1×1 and 3×3 convolutions. This fire module aims to reduce the parameters of SqueezeNet while maintaining efficient feature extraction [21].

After going through a series of fire modules, the data is processed until the last layer, conv10, which is the 1×1 convolution layer. This layer is the layer that produces the output of the image embedding process. The result of this output is a 1000-dimensional feature vector, which SVM then uses to classify SIBI hand gestures.

C. Classification Using Support Vector Machine (SVM)

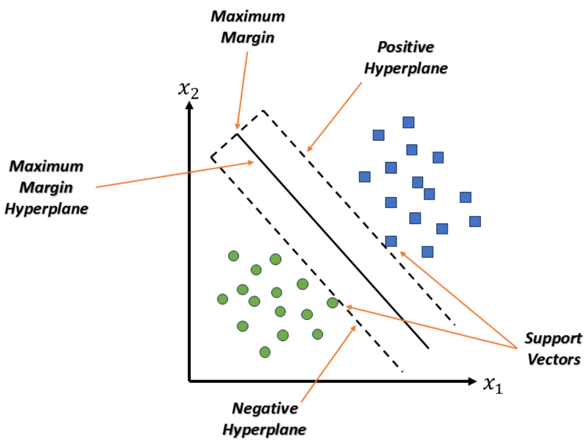

Figure 3. SVM for classification

SVM is one type of classification algorithm that predicts outcomes by constructing a hyperplane that maximizes the margin between classes [22]. SVMs are widely used for classification problems of features extracted from images. For example, [23] used SVM for disease classification in rice plants with features that have been extracted using ResNet50. As a result, this SVM classification achieved an accuracy rate of 99%.

Figure 3 illustrates the SVM used to classify a data sample consisting of two classes. Blue and green colored dots represent the data samples. As illustrated in **Figure 3**, a hyperplane is used to separate the data samples of the two classes. The distance between the hyperplane and the closest data point from each class is called the maximum margin, and this closest data point is called the support vector. The support vector is critical data in determining the position of the hyperplane and the decision boundary between classes. SVM requires a training process on these data samples with the aim of maximizing the margin between the support vectors of the two classes.

SVM has several parameter settings that can affect the classification results. One such parameter is cost (C), which serves as a trade-off between maximizing margin and minimizing classification error [24]. A higher C value penalizes misclassification more, so the algorithm has the potential for overfitting. Conversely, a lower C value can increase the tolerance to error but may result in underfitting [25].

In addition to setting parameters, SVM also has a kernel that functions to handle data that is notable to be separated linearly. The kernel is a function whose purpose is to transform data into a higher dimensional space to find the optimal separation hyperplane [26]. Some common kernel types used in SVM include linear kernel, polynomial kernel, and Radial Basis Function (RBF) kernel [27]. The use of RBF kernel in SVM allows the algorithm to capture non-linear patterns by mapping the data to a higher dimensional space [28].

This research uses SVM with cost parameter C = 1.0 and RBF kernel type to classify the output feature vector from SqueezeNet image embedding. The use of the RBF kernel is expected to classify data with non-linear patterns more accurately.

D. Evaluation Method

In this study, we evaluate the performance of the proposed method using several evaluation metrics. First, we evaluate the classification results using the confusion matrix. This matrix is used to benchmark the classification performance, which summarizes the correct and incorrect predictions and can describe true positive (TP), false positive (FP), true negative (TN), and false negative (FN) [29]. True positive refers to the amount of positive data that was correctly predicted, and true negative is the amount of negative data that was also correctly predicted. Meanwhile, a false positive, also known as Type I error, is the amount of negative data that is mistakenly predicted as positive. Meanwhile, a false negative, or Type II error, is the amount of positive data that was mistakenly predicted as negative. These four categories allow for a more in-depth evaluation of how the proposed method predicts each class in the data sample. In addition, these four categories can help identify and understand the types of errors that occur.

The second evaluation is done by calculating several metrics commonly used in classification problems. The four standard metrics for classification are classification accuracy (CA), precision, recall, and F1 score [30]. The values of these four metrics are calculated using the formulas presented in (1), (2), (3), and (4). The use of these metrics allows for a more comprehensive assessment of a classification algorithm's performance in classifying data accurately and consistently.

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

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

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

$$F1\ score = \frac{2 \times Precision \times Recall}{Precision \times Recall} \tag{4}$$

Results and Discussion

In this section, we will discuss the test results that have been conducted on SIBI hand gesture classification with SqueezeNet image embedding and Support Vector Machine (SVM). First, we will explain the results of the confusion matrix obtained, along with an analysis of the results. Secondly, we will present the classification metrics results to test the performance of the proposed method in this study. More details regarding the results of this research will be provided in the following subsections.

A. Confusion Matrix Result

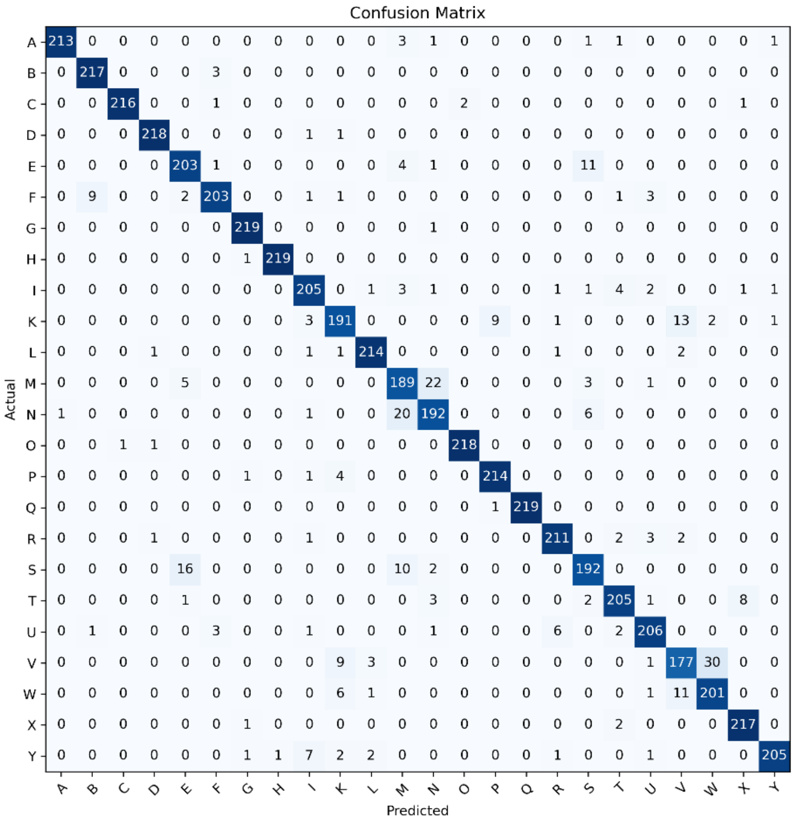

Figure 4. Confusion matrix of SqueezeNet image embedding and SVM classification on SIBI hand gestures

Figure 4 displays the confusion matrix results obtained from testing the classification of SIBI gestures using SqueezeNet image embedding and SVM. This confusion matrix presents a comparison of the predicted results with the actual data. Each row in the confusion matrix represents the actual class, while each column shows the resulting prediction.

Based on the result shown in Figure 4, gestures symbolizing the letters H, G, and Q show very high classification performance. A total of 219 hand gesture images were correctly predicted by the proposed method. However, there are still some misclassifications. For example, gesture G is misclassified as the letter N, gesture H is misclassified as the letter G, and gesture Q is misclassified as the letter P. In addition, there are some additional errors where gestures H, P, X, and Y are detected as letter G. This phenomenon indicates that the proposed method still experiences confusion in distinguishing gestures that may be visually similar.

Meanwhile, gestures symbolizing the letter V showed low classification performance. There were several significant misclassifications. For example, the gestures K, L, R, and W were predicted several times as the letter V. In total, only 177 data were correctly predicted as the letter V. In total, only 177 data were correctly predicted as the letter V. This result shows that the V gesture is complicated to predict compared to other gestures. This result indicates the need for improvement in the proposed method, especially on this difficult-to-predict gesture.

B. Classification Metrics Performance

Table 1 shows the evaluation results of SqueezeNet image embedding and SVM in classifying 24 SIBI gestures. This evaluation is done on several classification metrics, namely, classification accuracy (CA), precision, recall, and F1 score. Each metric provides an overview of the performance of the proposed method in correctly classifying each gesture.

Overall, SqueezeNet image embedding and SVM showed excellent performance in classifying SIBI gestures, with an average CA score of 99.51%, precision of 94.04%, recall of 94.02%, and F1 score of 94.02%. A high F1 score value indicates a good balance between precision and recall. Meanwhile, a high precision value indicates the ability of the proposed method to minimize false positives, while a high recall reflects its ability to reduce false negatives. Some gestures, such as H, Q, O, and D, show almost perfect performance, with CA and F1 scores reaching 100%. This condition indicates that the proposed method can recognize gestures that symbolize these letters very well.

Table 1. Performance evaluation of SqueezeNet image embedding and SVM for SIBI hand gesture classification

Gesture	CA (%)	Precision (%)	Recall (%)	F1 score (%)
A	99.8	99.5	96.8	98.2
B	99.8	95.6	98.6	97.1
C	99.9	99.5	98.2	98.9
D	99.9	98.6	99.1	98.9
E	99.2	89.4	92.3	90.8
F	99.5	96.2	92.3	94.2
G	99.9	98.2	99.5	98.9
H	100.0	99.5	99.5	99.5
I	99.4	92.3	93.2	92.8
K	99.0	88.8	86.8	87.8
L	99.8	96.8	97.3	97.1
M	98.7	82.5	85.9	84.2
N	98.9	85.7	87.3	86.5
O	99.9	99.1	99.1	99.1
P	99.7	95.5	97.3	96.4
Q	100.0	100.0	99.5	99.8
R	99.6	95.5	95.9	95.7
S	99.0	88.9	87.3	88.1
T	99.5	94.5	93.2	93.8
U	99.5	94.1	93.6	93.8
V	98.7	86.3	80.5	83.3
W	99.0	86.3	91.4	88.7
X	99.8	95.6	98.6	97.1
Y	99.7	98.6	93.2	95.8

Although the result has excellent performance, it still has potential for improvement, especially in distinguishing certain gestures. There are some gestures with low prediction results. For example, gesture V achieved a recall of only 80.5% and F1 Score of 83.3%. This result shows that the proposed method has difficulty in accurately identifying this gesture.

Compared to existing methods, such as MobileNet and VGG16 applied to the same SIBI dataset [12], which achieved overall accuracies of 98% and 86%, respectively, the proposed method demonstrates a favorable balance of performance and computational efficiency. Unlike deep CNN models that require extensive training and resources, the SqueezeNet–SVM approach achieves comparable or better results with significantly lower complexity, making it suitable for practical deployment.

Although the model performs well on the current dataset, its generalization to other datasets or real-world conditions remains a critical consideration. Since this study did not apply data augmentation and used only the original dataset, there is potential for overfitting the specific lighting, background, and hand shapes present in the training data. This could limit performance when deployed in uncontrolled environments.

Moreover, the proposed method also offers significant promise for real-world applications. Given its minimal computational resources, it can be adapted for integration with smartphone cameras or edge devices. However, real-world deployment would require further optimization to handle low-light environments, motion blur, and real-time inference constraints. Future implementation should include testing under such conditions to ensure practical reliability.

Conclusion

This research has proposed a combination of SqueezeNet image embedding and Support Vector Machine (SVM) for hand gesture recognition in SIBI sign language. The experimental results describe that the proposed method can achieve a high level of accuracy, with a CA value of 99.51%. In addition, measurements on other classification metrics such as precision, recall, and F1 score reached a satisfactory figure, which is around 94%. The proposed methods managed to recognize most of the hand gestures very well, especially for gestures that symbolize the letters G, H, and Q, with almost perfect classification results. However, recognition results are still less than optimal, especially for the V gesture. This result indicates that the proposed method is quite effective but still needs improvement in handling certain gestures that have visual similarities or a higher level of difficulty in recognition. Furthermore, the proposed method offers a promising opportunity for integration into real-time applications. With minimal computational cost, the model is well-suited for deployment on mobile platforms or embedded systems.

This study open opportunities for further development. First, future research can consider a data augmentation to increase the number of datasets in the training phase so that the generalization of the algorithm can increase. Second, it is necessary to consider the use of other types of CNN architecture that can improve the classification metrics. Third, future development needs to focus on integrating the model into a prototype mobile application. Real-world testing also needs to be conducted in relevant environments, such as schools for the deaf or public service institutions. These field trials will assess the system's robustness under varied conditions, including low-light settings, diverse backgrounds, and user variability. Through these future efforts, the proposed method can be further enhanced to support the development of an accessible and practical sign language translation system that helps bridge communication gaps between the deaf community and the public.

Acknowledgment

We would live to thank the Department of Electrical Engineering, Politeknik Negeri Batam, for providing resources and equipment support for this research.

References

[1] Q. M. Areeb, Maryam, M. Nadeem, R. Alroobaea, and F. Anwer, "Helping Hearing-Impaired in Emergency Situations: A Deep Learning-Based Approach," *IEEE Access*, vol. 10, pp. 8502–8517, 2022, doi: 10.1109/ACCESS.2022.3142918.

[2] T. N. Fitria, "The Use of Sign Language as a Media for Delivering Information on National Television News Broadcasts," *ELP (Journal of English Language Pedagogy)*, vol. 9, no. 1, Art. no. 1, Jan. 2024, doi: 10.36665/elp.v9i1.764.

[3] 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, Jun. 2023, doi: 10.35438/jits.v1i1.21.

[4] I. D. A. Rachmawati, R. Yunanda, M. F. Hidayat, and P. Wicaksono, "Deep Transfer Learning for Sign Language Image Classification: A Bisindo Dataset Study," *Engineering, MAthematics and Computer Science Journal (EMACS)*, vol. 5, no. 3, Art. no. 3, Sep. 2023, doi: 10.21512/emacsjournal.v5i3.10621.

[5] 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.

[6] E. Rakun, S. Muzahidin, I. S. A. Darmana, and W. Setiaji, "SIBI (Sign System Indonesian Language) Text-to-3D Animation Translation Mobile Application," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 13, no. 9, Art. no. 9, Dec. 2022, doi: 10.14569/IJACSA.2022.0130950.

[7] S. Syamsuddin, T. Pristiwaluyo, W. A. Saleh, and Z. Zulfitrah, "Development of Indonesian Sign Language System (SIBI) Dictionary Application for Students With Special Needs," *AL-ISHLAH: Jurnal Pendidikan*, vol. 15, no. 3, Art. no. 3, Sep. 2023, doi: 10.35445/alishlah.v15i3.3529.

[8] M. Yunus and Y. Anwar, “Aplikasi Penerjemah Bahasa Isyarat Indonesia ke Dalam Huruf Abjad,” *Jurnal Sintaks Logika*, vol. 2, no. 1, Art. no. 1, Jun. 2022, doi: 10.31850/jsilog.v2i1.1726.

[9] M. A. Yasir and Y. H. Ali, "Review on Real Time Background Extraction: Models, Applications, Environments, Challenges and Evaluation Approaches," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 17, no. 02, Art. no. 02, Feb. 2021, doi: 10.3991/ijoe.v17i02.18013.

[10] I. I. Arifah, F. N. Fajri, and G. Q. O. Pratamasunu, "Deteksi Tangan Otomatis pada Video Percakapan Bahasa Isyarat Indonesia Menggunakan Metode YOLO dan CNN," *Journal of Applied Informatics and Computing*, vol. 6, no. 2, Art. no. 2, Nov. 2022, doi: 10.30871/jaic.v6i2.4694.

[11] D.-Q. Yang *et al.*, "A Systematic Study on Transfer Learning: Automatically Identifying Empty Camera Trap Images Using Deep Convolutional Neural Networks," *Ecological Informatics*, vol. 80, p. 102527, May 2024, doi: 10.1016/j.ecoinf.2024.102527.

[12] M. B. Subkhi, M. Y. Trinurais, R. K. A. Wibowo, and B. R. Prakosa, "Deteksi Bahasa Isyarat Berdasarkan SIBI (Sistem Bahasa Isyarat) Menggunakan Transfer Learning," *Seminar Nasional Teknologi & Sains*, vol. 3, no. 1, Art. no. 1, Jan. 2024, doi: 10.29407/stains.v3i1.4347.

[13] M. Maryamah, M. A. Pratama, M. R. Erfit, N. M. Farhani, and I. A. Hartono, “Klasifikasi Abjad SIBI (Sistem Bahasa Isyarat Indonesia) Menggunakan Mediapipe dengan Metode Deep Learning,” *PROSIDING SEMINAR NASIONAL SAINS DATA*, vol. 3, no. 1, Art. no. 1, Nov. 2023, doi: 10.33005/senada.v3i1.102.

[14] W. Rhee *et al.*, "Unconstrained Lightweight Control Interface for Robot-Assisted Minimally Invasive Surgery Using MediaPipe Framework and Head-Mounted Display," *Virtual Reality*, vol. 28, no. 2, p. 114, May 2024, doi: 10.1007/s10055-024-00986-1.

[15] H. Ahn *et al.*, "BioEdge: Accelerating Object Detection in Bioimages with Edge-Based Distributed Inference," *Electronics*, vol. 12, no. 21, Art. no. 21, Jan. 2023, doi: 10.3390/electronics12214544.

[16] B. Liu *et al.*, "An Energy-Efficient SIFT Based Feature Extraction Accelerator for High Frame-Rate Video Applications," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 69, no. 12, pp. 4930–4943, Dec. 2022, doi: 10.1109/TCSI.2022.3199475.

[17] A. Sen, T. K. Mishra, and R. Dash, "A Novel Hand Gesture Detection and Recognition System Based on Ensemble-Based Convolutional Neural Network," *Multimed Tools Appl*, vol. 81, no. 28, pp. 40043–40066, Nov. 2022, doi: 10.1007/s11042-022-11909-0.

[18] H. Liang, L. Fei, S. Zhao, J. Wen, S. Teng, and Y. Xu, "Mask-Guided Multiscale Feature Aggregation Network for Hand Gesture Recognition," *Pattern Recognition*, vol. 145, p. 109901, Jan. 2024, doi: 10.1016/j.patcog.2023.109901.

[19] J. Zhang, H. Zhu, P. Wang, and X. Ling, "ATT Squeeze U-Net: A Lightweight Network for Forest Fire Detection and Recognition," *IEEE Access*, vol. 9, pp. 10858–10870, 2021, doi: 10.1109/ACCESS.2021.3050628.

[20] H. J. Lee, I. Ullah, W. Wan, Y. Gao, and Z. Fang, "Real-Time Vehicle Make and Model Recognition With the Residual SqueezeNet Architecture," *Sensors*, vol. 19, no. 5, Art. no. 5, Jan. 2019, doi: 10.3390/s19050982.

[21] N. Yamanakkanavar, J. Y. Choi, and B. Lee, "SM-SegNet: A Lightweight Squeeze M-SegNet for Tissue Segmentation in Brain MRI Scans," *Sensors*, vol. 22, no. 14, Art. no. 14, Jan. 2022, doi: 10.3390/s22145148.

[22] Y. Wang, W. Liao, H. Shen, Z. Jiang, and J. Zhou, "Some Notes on the Basic Concepts of Support Vector Machines," *Journal of Computational Science*, vol. 82, p. 102390, Oct. 2024, doi: 10.1016/j.jocs.2024.102390.

[23] douaa S. Alwan and M. H. Naji, “Rice Diseases Classification by Residual Network 50 (ResNet50) and Support Vector Machine (SVM) Modeling,” *Journal of Kufa for Mathematics and Computer*, vol. 10, no. 1, pp. 96–101, Mar. 2023, doi: 10.31642/JoKMC/2018/100114.

[24] C. Shi, J. Zhu, M. Xu, X. Wu, and Y. Peng, "An Approach of Spectra Standardization and Qualitative Identification for Biomedical Materials Based on Terahertz Spectroscopy," *Scientific Programming*, vol. 2020, no. 1, p. 8841565, 2020, doi: 10.1155/2020/8841565.

[25] Y. Dong, Z. Xia, and Z. Xia, "A Two-Level Approach to Choose the Cost Parameter in Support Vector Machines," *Expert Systems with Applications*, vol. 34, no. 2, pp. 1366–1370, Feb. 2008, doi: 10.1016/j.eswa.2007.01.004.

[26] B. Hicdurmaz, N. Calik, and S. Ustebay, "Gauss-like Logarithmic Kernel Function to Improve the Performance of Kernel Machines on the Small Datasets," *Pattern Recognition Letters*, vol. 179, pp. 178–184, Mar. 2024, doi: 10.1016/j.patrec.2024.01.014.

[27] N. Elshamy, A. AbouElenen, and S. Elmougy, "Automatic Detection of Software Defects based on Machine Learning," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 14, no. 3, Art. no. 3, 42/30 2023, doi: 10.14569/IJACSA.2023.0140340.

[28] A. Koné, A. Es-Sabar, and M.-T. Do, "Application of Machine Learning Models to the Analysis of Skid Resistance Data," *Lubricants*, vol. 11, no. 8, Art. no. 8, Aug. 2023, doi: 10.3390/lubricants11080328.

[29] I. Düntsch and G. Gediga, "Indices for Rough Set Approximation and the Application to Confusion Matrices," *International Journal of Approximate Reasoning*, vol. 118, pp. 155–172, Mar. 2020, doi: 10.1016/j.ijar.2019.12.008.

[30] W. Sun, B. Fu, and Z. Zhang, "Maize Nitrogen Grading Estimation Method Based on UAV Images and an Improved Shufflenet Network," *Agronomy*, vol. 13, no. 8, Art. no. 8, Aug. 2023, doi: 10.3390/agronomy13081974.