

Dynamic Adaptation in Deep Learning for Enhanced Hand Gesture Recognition

Abdirahman Osman Hashi

SIMAD University, Department of Computing, Mogadishu, Somalia | Department of Computer Science, Faculty of Computing, University Teknologi Malaysia, 81310 Skudai, Johor, Malaysia
wadani12727@gmail.com (corresponding author)

Siti Zaiton Mohd Hashim

Department of Software Engineering, Faculty of Computing, University Teknologi Malaysia, 81310 Skudai, Johor, Malaysia
sitizaiton@utm.my

Azurah Bte Asamah

Department of Computer Science, Faculty of Computing, University Teknologi Malaysia, 81310 Skudai, Johor, Malaysia
azurah@utm.my

Received: 28 April 2024 | Revised: 26 May 2024 | Accepted: 27 May 2024

Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | DOI: <https://doi.org/10.48084/etasr.7670>

ABSTRACT

The field of Human-Computer Interaction (HCI) is progressing quickly with the incorporation of gesture recognition, which requires advanced systems capable of comprehending intricate human movements. This study introduces a new Dynamic Adaptation Convolutional Neural Network (DACNN) that can adjust to different human hand shapes, orientations, and sizes. This allows for more accurate identification of hand gestures over a wide range of variations. The proposed model includes a thorough process of collecting and preparing data from the Sign Language MNIST dataset. This is followed by a strong data augmentation procedure that provides a wide variety of realistic variations. The architecture utilizes sophisticated convolutional layers to leverage the capabilities of deep learning to extract and synthesize essential gesture features. A rigorous training procedure, supplemented with a ReduceLROnPlateau callback, was used to assure the model's generalization and efficiency. The experimental findings provide remarkable results, showing a substantial accuracy of 99% in categorizing a wide range of hand movements. This study makes a significant contribution to the field of hand gesture recognition by introducing morphological operations, thus enriching input data quality and expanding the model's applicability in diverse HCI environments.

Keywords-hand gesture recognition; human-computer interaction; deep learning; neural network architecture; real-time gesture analysis; morphological data processing; adaptive learning systems

I. INTRODUCTION

With advances in artificial intelligence and its increasing applications, a deeper understanding of human behavior can enhance the interaction between machines and humans [1]. A popular topic in computer vision is the study of how people behave and how to make computers or robots behave like them [2]. Computer systems can analyze visual data to assess human behavior, determine their requirements, and then respond appropriately. Hand Gesture Recognition (HGR) is an efficient method to improve communication between people and robots [3]. HGR is often used in human communication as a visual means of expressing thoughts through coordinated hand movements and serves as the primary communication medium for people with speech and hearing impairments [4]. According to the World Health Organization, approximately 5% of the

global population have moderate to severe hearing impairments and rely on their local sign languages for communication [5].

HGR has significant communicative value, especially for people who experience moderate to severe hearing impairments. For these people, local sign languages are the mainstay of communication. Thus, the ability to recognize and interpret hand gestures accurately is not only a technical achievement but also a facilitator of inclusion, bridging communication gaps in society [6]. HCI is now a crucial part of daily activities. The use of hand gestures in HCI has garnered significant interest due to its user-friendly approach to interaction with technology. Hand gestures are used to convey ideas and thoughts through movements of the palm and fingers. Individuals should interact with machines using hand movements without requiring additional input devices. For

effective communication between humans and robots, machines must be able to identify hand gestures. Hence, HGR has become a significant study field in modern times, serving many purposes such as natural user interfaces [7].

Gesture recognition with computer technology has great potential as it can serve as a translator for sign languages. This method could successfully close the communication gap between both societies, providing significant benefits in promoting engagement and comprehension between those who use sign language and those who do not [8]. Sign language hand movements can be classified into two types: static and dynamic gestures. Static gestures involve placing hands and fingers in space without movement, while dynamic gestures require continual hand movements throughout time. HGR in sign language translation can be recognized by two methods: vision-based and sensor-based recognition [9]. Sensor-based techniques require the signer to wear electronic equipment such as data gloves, accelerometers, or bands. These components detect changes in movement and transmit the data to a computer for further analysis. Although research has shown promising recognition results for this technology, its cost and impracticality make it unsuitable for common HCI interfaces for consumers [10]. Meanwhile, vision-based HGR plays a pivotal role in improving HCI by allowing users to communicate with machines more intuitively and naturally [11]. This technology has attracted significant attention from researchers due to its potential applications in various fields, including sign language interpretation, virtual reality, augmented reality, and assistive technologies. Using Deep Learning (DL) techniques, researchers have been able to revolutionize the field of gesture recognition, making significant advances in the accuracy and efficiency of recognition systems. DL algorithms have enabled the development of sophisticated models capable of interpreting complex hand gestures in real time, thus improving the overall user experience in HCI applications [12].

The use of DL in sensor-based HGR has transformed traditional paradigms, allowing for more robust and accurate recognition systems [13]. In [14], real-time HGR was proposed using the YOLOv3 model. In [15], a multi-model sensor-based system was proposed for HGR. By analyzing and interpreting hand movements captured through cameras or sensors, DL models can extract meaningful features and patterns, enabling machines to effectively understand and respond to hand gestures. This technology has paved the way for the development of innovative HCI systems that can interpret a wide range of gestures with high precision, facilitating seamless interaction between users and machines [16].

In recent years, the surge of interest in vision-based HGR has been largely fueled by remarkable breakthroughs in DL. Many studies have explored a spectrum of DL architectures, pushing the boundaries of how machines interpret human gestures. For instance, in [17], a DL approach was proposed for age-based gesture classification for Indian sign language. The core of these investigations lies in the quest to provide computers with a level of perceptual intelligence that mirrors human intuition [18]. This quest is fundamental for the evolution of interfaces that require minimal effort and are

intuitive for users, leading to a seamless HCI experience. Convolutional Neural Networks (CNNs) have emerged as one of the most powerful tools for image analysis and are used in the HGR domain. CNNs excel at hierarchical feature extraction, which is crucial for recognizing complex patterns in visual data. By learning spatial hierarchies of features, CNNs can robustly handle the variability in hand gesture appearance due to differences in hand shapes, sizes, and orientations between individuals. This adaptive feature extraction has proven essential for the development of gesture recognition models that can operate with high accuracy in diverse user groups [19].

Furthermore, the dynamism of hand gestures necessitates the use of models that can understand temporal dependencies. Recurrent Neural Networks (RNNs) are adept at processing sequences of data, making them ideal for analyzing the temporal progression of hand gestures. When it comes to continuous gesture recognition or interpreting gestures that consist of a series of movements, RNNs can provide context by considering the sequence of frames, leading to more accurate recognition of gestures over time [20]. Incorporating attention mechanisms into neural network architectures marks a significant advance in the field. Attention mechanisms enable models to focus on the most relevant parts of the input data, similar to how humans pay attention to certain aspects of a visual scene while ignoring others. This selective focus has been instrumental in improving the performance of gesture recognition systems, especially in cluttered or dynamic environments where irrelevant movements can easily confound less sophisticated models [21].

Therefore, the application of Machine Learning (ML) and DL models reveals the potential of these technologies in learning complex gesture patterns. Hybrid models and ensemble classifiers, which combine multiple techniques to improve recognition rates, are innovative approaches being explored in the field [22]. Additionally, the use of hand motion segmentation and real-time detection techniques further enhances the ability to recognize gestures accurately and efficiently [20]. Furthermore, the exploration of multimodal HGR, which combines information from different sources such as skeletal data, depth data, and RGB images, presents a holistic approach to understanding and interpreting human gestures. This approach not only increases the accuracy of gesture recognition but also enhances the system's ability to recognize gestures in a wide range of scenarios and environments [23]. Therefore, it is important to have a hand gesture model that can dynamically adapt to the morphological features of hand gestures.

This study investigates advances in vision-based HGR for HCI using DL methods and provides insights into the potential to enhance gesture recognition systems and shape the future of HCI. The contributions of this work are:

- Presents a novel neural network architecture that emphasizes dynamic adaptability to various hand gestures, distinguishing itself from previous studies that focus predominantly on static HGR.

- Uses cutting-edge DL methods for both feature extraction and classification tasks, using layers and procedures that are particularly designed to address the intricacies of HGR.

II. METHODOLOGY

Vision-based HGR is a complex field that is leading the way in innovative GCI, especially in the area of DL. The intricate and diverse nature of human hand motions poses a significant obstacle that requires an advanced approach to utilize neural networks. The proposed architecture is designed to address this problem using a custom DL model that can be adjusted to the various morphological features of hands. This approach combines data preprocessing, augmentation, and a sophisticated CNN architecture to improve flexibility and accuracy in recognition. Figure 1 shows the general flowchart of the method followed.

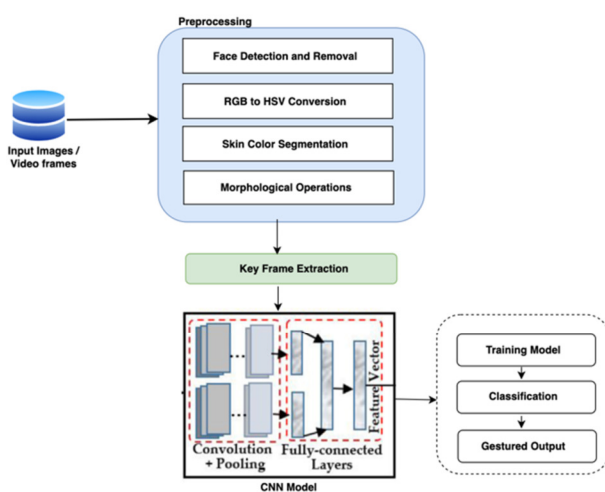


Fig. 1. Methodology.

A. Data Collection and Preprocessing

The first phase was to collect and prepare the data for training the DL model. The Sign Language MNIST dataset was used [24]. This dataset was preprocessed by removing labels for separate handling and normalizing the pixel values to a range of 0 to 1 for computational efficiency. The dataset comprises images of hand gestures, each representing a different letter of the sign language alphabet, which were reshaped from 1-D vectors to 28×28 pixel images to match the input requirements of the CNN. The preprocessing phase was extensive and crucial for the subsequent recognition process. It started with face detection and removal to ensure that only hand gestures are analyzed, eliminating potential distractions for the model. Following this, the images were converted from RGB to the HSV color space, a transformation that often leads to better segmentation results due to its closer alignment with the human perception of colors. After conversion, skin color segmentation was applied to isolate the hand from the background, which is particularly beneficial for focusing the model's learning on the gestures themselves. Finally, morphological operations such as erosion and dilation were used to enhance the quality of the segmented hand image, removing noise and filling gaps in the detected hand region.

B. Data Augmentation

Data augmentation aimed to simulate the kind of variability that a real-world application would encounter. The ImageDataGenerator not only performs the standard transformations but is fine-tuned to reflect the nuances of hand gestures. It was ensured that the augmentations did not stray too far from realistic scenarios, as overly distorted images could harm the model's learning process. This phase was specifically designed to address the core objective of the study, which is to achieve a model that dynamically adapts to diverse morphologies. By simulating a wide variety of angles and hand positions, the model was trained to recognize the essence of each gesture, independent of the hand's shape, size, or orientation.

C. Model Architecture

The feature extraction process is the heart of the model architecture and is carried out by the convolutional layers. These layers dissect the input images into features that the network will use to distinguish one gesture from another. The network's ability to extract and learn from these features is paramount to the overall adaptability and accuracy of the HGR system. Once the features are extracted, the network's fully connected layers interpret them, effectively synthesizing the information into a comprehensive understanding that leads to classification. Following the convolutional layers, max-pooling layers were implemented to reduce spatial dimensions, thus reducing the computational load and potentially overfitting. Dropout layers were also strategically included to randomly ignore a subset of neurons during training, forcing the network to learn redundant representations of the data. The final part is a densely connected layer that interprets the features extracted by the convolutional layers, culminating in a softmax activation function that outputs a probability distribution over the 24 classes of hand gestures. The softmax layer then assigns a probability to each gesture class, ensuring that the output is a clear and decisive prediction of the input gesture.

D. Training and Validation

The learning process was fine-tuned using the ReduceLROnPlateau callback to make training adaptable by reducing the learning rate when the validation accuracy plateaued. This optimization prevents learning stagnation and promotes the discovery of optimal weights within the network. Validation was implemented as a continuous check on the model's ability to generalize. Using a validation set separate from the training data was aimed at monitoring and ensuring that the model was learning to recognize gestures and not just memorizing the training dataset.

E. Performance Evaluation

The performance evaluation phase was multi-faceted, starting with a classification report that provided a detailed account of the model's precision, recall, and F1-score for each class. This evaluation is vital for understanding the model's efficacy and identifying any specific gestures that may require further attention or data augmentation. The confusion matrix offers a visual and quantitative analysis of the model performance, highlighting instances of misclassification. This matrix is a powerful tool for pinpointing where the model

excels and where it may confuse one gesture with another. Such insights are crucial for ongoing model refinement and achieving adaptability to diverse hand morphologies.

III. RESULTS AND DISCUSSION

Figure 2 shows the results of converting images from RGB to HSV. Skin color segmentation was instrumental, as it refined the focus on gestural nuances by isolating hands from varying backgrounds. This preprocessing not only increased computational efficiency but also improved learning accuracy by eliminating unnecessary visual distractions and facilitating the detection of morphological operations.

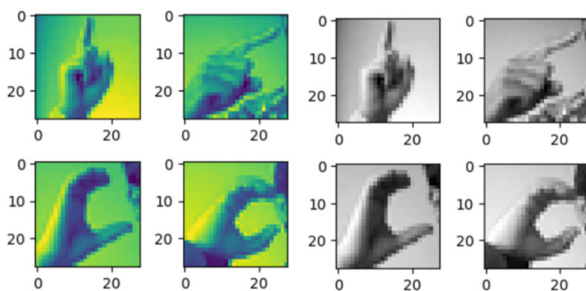


Fig. 2. RGB to HSV images.

A. Results

Figure 3 details the architecture of the neural network model. The initial convolutional layer (conv2d), with its multiple filters, plays a critical role in detecting various features in the input images, such as edges and textures. The subsequent batch normalization layer is crucial in stabilizing the learning process by normalizing the output of the previous convolutional layer, leading to faster convergence and improved overall performance. The use of max-pooling (max_pooling2d) reduced the spatial dimensions of the feature maps, effectively summarizing the presence of features while making the detection process invariant to small translations of the input. This operation also helps to reduce the computational load. Dropout layers act as a form of regularization to prevent overfitting, encouraging the network to learn more robust features that are not reliant on a small number of neurons.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 28, 28, 75)	750
batch_normalization (Batch Normalization)	(None, 28, 28, 75)	300
max_pooling2d (MaxPooling2D)	(None, 14, 14, 75)	0
conv2d_1 (Conv2D)	(None, 14, 14, 50)	33800
dropout (Dropout)	(None, 14, 14, 50)	0
batch_normalization_1 (Batch Normalization)	(None, 14, 14, 50)	200
max_pooling2d_1 (MaxPooling2D)	(None, 7, 7, 50)	0
conv2d_2 (Conv2D)	(None, 7, 7, 25)	11275
batch_normalization_2 (Batch Normalization)	(None, 7, 7, 25)	100
max_pooling2d_2 (MaxPooling2D)	(None, 4, 4, 25)	0
flatten (Flatten)	(None, 400)	0
dense (Dense)	(None, 512)	205312
dropout_1 (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 24)	12312

Fig. 3. Model architecture.

Figure 4 shows the training progression over the epochs, demonstrating the model's learning efficiency. The gradual decrease in loss and consistent improvement in both training and validation accuracy highlight the model's capability to learn and generalize from the dataset. The model's validation accuracy aligns closely with the training accuracy, suggesting that the model is not overfitting but rather capturing generalizable patterns. The plateauing of accuracy and subsequent reduction in learning rate by the ReduceLRonPlateau callback indicates an effective strategy to further improve the model's weights when learning improvement stagnates. The final epoch shows that the model achieves impressive accuracy, a clear indication of the model's high proficiency in recognizing and classifying hand gestures.

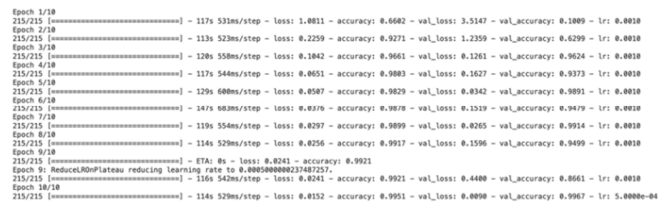


Fig. 4. Training loss per epoch.

The reduction in learning rate after the ninth epoch suggests that the model's performance began to plateau, and the intervention helped to fine-tune the learning process, which is observed in the improved post-adjustment accuracy. The model achieved an impressive final accuracy rate, indicating a high level of precision in recognizing and classifying hand gestures, as shown in Figure 5. The accuracy plot on the left exhibits a rapid ascent to high accuracy levels during the initial epochs, followed by a stable plateau, with the training accuracy (green line) hovering near perfection at 0.99 and the validation accuracy (red line) closely mirroring this trend. This demonstrates the model's substantial capability to learn and generalize from the dataset without overfitting, as there is minimal divergence between the training and validation accuracy scores. The loss plot on the right underscores the model's efficient learning curve, with a steep decline in the initial epochs reflecting a rapid reduction in the error rate of the model's predictions. Following this swift descent, both the training and testing loss exhibit a steady state with minor fluctuations, indicating a stable convergence.

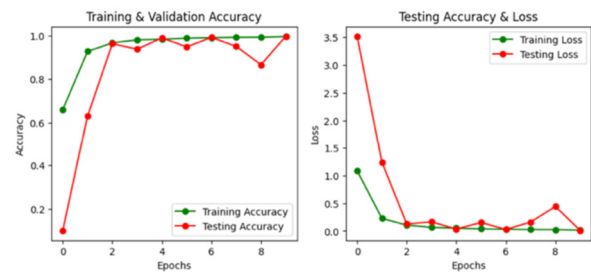


Fig. 5. Training and testing accuracy and loss.

Achieving a 99% accuracy score is a notable indication of the model's performance, highlighting its strength and the success of the implemented architecture and training methods.

The model's near-flawless accuracy demonstrates its remarkable ability to identify and categorize the hand gestures presented, which aligns with the initial objective of developing a versatile HGR system. Figure 6 presents the correct prediction indices in the form of a confusion matrix, revealing a nearly perfect classification with a high density of true positives along the matrix's diagonal. This visualization provides compelling evidence of the model's proficiency in differentiating between gesture classes with minimal confusion. The accuracy score of 0.99% suggests that the vast majority of gestures were correctly classified, with only a minimal number scattered across the off-diagonal cells. This near-perfect distribution demonstrates the model's exceptional ability not only to grasp the complex patterns within the hand gesture data but also to generalize well to new, unseen data, as indicated by the validation results.

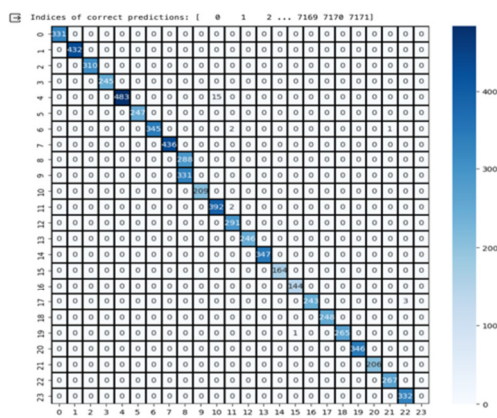


Fig. 6. Confusion matrix.

B. Gesture Output

Figure 7 demonstrates the model's ability to discern and accurately classify hand gestures. The agreement between the predicted and actual classes highlights the effectiveness of the proposed architecture. For instance, in class 6, the model correctly identifies the gesture with a high degree of certainty, indicating that the nuanced features of this class have been effectively captured during training. Similarly, the results for class 5 showcase the model's ability to recognize and interpret gestures with subtle differences from other classes, a crucial requirement in the practical application of gesture-based communication. The accuracy is further exemplified in classes 9 and 0. Here, the model not only demonstrates its strength in capturing the static positioning of the hands but also its proficiency in understanding the dynamic nature of gesture transitions. The system responds to both static and dynamic gestures, enabling a more fluid and natural interaction within the HCI realm. These visual outputs are not just a mere display of classification correctness but also an affirmation of the underlying sophisticated data processing and learning mechanisms. The seamless alignment between predicted and actual classes across a variety of gestures underscores the model's robust generalization capabilities and reinforces the model's potential for deployment in real-world scenarios where the interpretation of a diverse array of hand gestures is paramount.

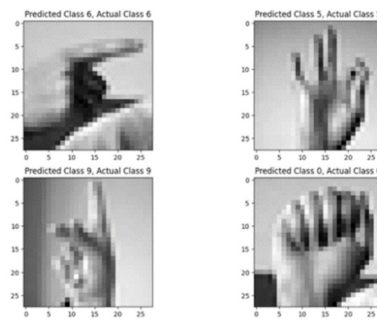


Fig. 7. Output of hand gestures.

C. Comparative Study

When it comes to the advancing field of HGR using neural networks, it is crucial to evaluate the effectiveness of new models by comparing them with previous solutions. This proposed architecture, hereafter referred to as the Dynamic Adaptation Convolutional Neural Network (DACNN), was compared with three other established models, focusing on the accuracy of each model as the primary metric for performance evaluation. The DACNN model was compared to:

- Static Feature Convolutional Neural Network (SFCNN) [25], which relies on static hand features without dynamic adaptation, with a reported accuracy of 0.94.
- Basic Convolutional Neural Network (BCNN) [26], a straightforward CNN model without advanced adaptations or augmentation techniques that achieved 0.90 accuracy.
- Enhanced Geometric Neural Network (EGNN) [27], which uses geometric feature enhancements for gesture recognition, with an accuracy of 0.92.

TABLE I. ACCURACY COMPARISON

#	Model Name	Reference	Accuracy
1	SFCNN	[25]	0.94%
2	BCNN	[26]	0.90%
3	EGNN	[27]	0.92%
4	DACNN	Proposed	0.99%

The DACNN model significantly outperforms its counterparts in accuracy. Although SFCNN, BCNN, and EGNN demonstrate competent gesture recognition capabilities, they lack the dynamic adaptability and advanced feature processing of DACNN. DACNN's superior performance is attributed to its innovative network design, which incorporates mechanisms to handle variability in hand gestures more effectively. These include a deeper layer structure, advanced data augmentation techniques, and the inclusion of dynamic adaptability in recognizing continuous gestures. The high accuracy of DACNN demonstrates its robustness against overfitting, a common challenge in complex models. This was achieved through a sophisticated training regimen that includes dropout layers for regularization and ReduceLRonPlateau callback to fine-tune learning rates. This comparative study presents the DACNN model as a leading-edge solution, offering the potential for extensive applications in HCI, sign language interpretation, and beyond.

IV. CONCLUSION

This study introduced the DACNN model, which marks a significant step forward in the field of HGR for HCI. DACNN demonstrated a profound ability to dynamically adapt to the diverse morphological characteristics of human hands, boasting an impressive accuracy rate of 0.99. This breakthrough is a direct result of the model's sophisticated architecture, which is designed to effectively process and learn from an extensive array of hand shapes, orientations, and sizes. The success of the DACNN model underlines the importance of comprehensive data preprocessing and innovative augmentation techniques in enhancing the adaptability and accuracy of HGR systems. The efficacy of this approach was consistently reflected in various phases of model training and validation, and comparison with other models demonstrated its superior performance and potential. The success of DACNN opens up numerous avenues for future research and applications. One immediate direction is its integration into real-time HCI systems, exploring its responsiveness and utility in dynamic environments.

REFERENCES

- [1] R. Rastgoo, K. Kiani, S. Escalera, and M. Sabokrou, "Multi-modal zero-shot dynamic hand gesture recognition," *Expert Systems with Applications*, vol. 247, Aug. 2024, Art. no. 123349, <https://doi.org/10.1016/j.eswa.2024.123349>.
- [2] L. I. Barona López, F. M. Ferri, J. Zea, Á. L. Valdivieso Caraguay, and M. E. Benalcázar, "CNN-LSTM and post-processing for EMG-based hand gesture recognition," *Intelligent Systems with Applications*, vol. 22, Jun. 2024, Art. no. 200352, <https://doi.org/10.1016/j.iswa.2024.200352>.
- [3] 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, Jan. 2024, Art. no. 109901, <https://doi.org/10.1016/j.patcog.2023.109901>.
- [4] Z. Wang *et al.*, "Optimization of inter-subject sEMG-based hand gesture recognition tasks using unsupervised domain adaptation techniques," *Biomedical Signal Processing and Control*, vol. 92, Jun. 2024, Art. no. 106086, <https://doi.org/10.1016/j.bspc.2024.106086>.
- [5] A. Bahuguna, G. Bhaumik, and M. C. Govil, "Local Extrema Min-Max Pattern: A novel descriptor for extracting compact and discrete features for hand gesture recognition," *Biomedical Signal Processing and Control*, vol. 93, Jul. 2024, Art. no. 106203, <https://doi.org/10.1016/j.bspc.2024.106203>.
- [6] P. Balaji and M. Ranjan Prusty, "Multimodal fusion hierarchical self-attention network for dynamic hand gesture recognition," *Journal of Visual Communication and Image Representation*, vol. 98, Feb. 2024, Art. no. 104019, <https://doi.org/10.1016/j.jvcir.2023.104019>.
- [7] Z. Zhang, Q. Shen, and Y. Wang, "Electromyographic hand gesture recognition using convolutional neural network with multi-attention," *Biomedical Signal Processing and Control*, vol. 91, May 2024, Art. no. 105935, <https://doi.org/10.1016/j.bspc.2023.105935>.
- [8] T. R. Gadekallu *et al.*, "Hand gesture recognition based on a Harris Hawks optimized Convolution Neural Network," *Computers and Electrical Engineering*, vol. 100, May 2022, Art. no. 107836, <https://doi.org/10.1016/j.compeleceng.2022.107836>.
- [9] T. Triwiyanto, I. P. A. Pawana, and W. Caesarendra, "Deep learning approach to improve the recognition of hand gesture with multi force variation using electromyography signal from amputees," *Medical Engineering & Physics*, vol. 125, Mar. 2024, Art. no. 104131, <https://doi.org/10.1016/j.medengphy.2024.104131>.
- [10] T. M. Tai, Y. J. Jhang, Z. W. Liao, K. C. Teng, and W. J. Hwang, "Sensor-Based Continuous Hand Gesture Recognition by Long Short-Term Memory," *IEEE Sensors Letters*, vol. 2, no. 3, pp. 1–4, Sep. 2018, <https://doi.org/10.1109/LSSENS.2018.2864963>.
- [11] D. Sarma and M. K. Bhuyan, "Methods, Databases and Recent Advancement of Vision-Based Hand Gesture Recognition for HCI Systems: A Review," *SN Computer Science*, vol. 2, no. 6, Aug. 2021, Art. no. 436, <https://doi.org/10.1007/s42979-021-00827-x>.
- [12] T. L. Dang, S. D. Tran, T. H. Nguyen, S. Kim, and N. Monet, "An improved hand gesture recognition system using keypoints and hand bounding boxes," *Array*, vol. 16, Dec. 2022, Art. no. 100251, <https://doi.org/10.1016/j.array.2022.100251>.
- [13] Q. Gao, Y. Chen, Z. Ju, and Y. Liang, "Dynamic Hand Gesture Recognition Based on 3D Hand Pose Estimation for Human–Robot Interaction," *IEEE Sensors Journal*, vol. 22, no. 18, pp. 17421–17430, Sep. 2022, <https://doi.org/10.1109/JSEN.2021.3059685>.
- [14] A. Mujahid *et al.*, "Real-Time Hand Gesture Recognition Based on Deep Learning YOLOv3 Model," *Applied Sciences*, vol. 11, no. 9, Jan. 2021, Art. no. 4164, <https://doi.org/10.3390/app11094164>.
- [15] H. G. Doan and N. T. Nguyen, "Fusion Machine Learning Strategies for Multi-modal Sensor-based Hand Gesture Recognition," *Engineering, Technology & Applied Science Research*, vol. 12, no. 3, pp. 8628–8633, Jun. 2022, <https://doi.org/10.48084/etasr.4913>.
- [16] S. Wang *et al.*, "Hand gesture recognition framework using a lie group based spatio-temporal recurrent network with multiple hand-worn motion sensors," *Information Sciences*, vol. 606, pp. 722–741, Aug. 2022, <https://doi.org/10.1016/j.ins.2022.05.085>.
- [17] R. M. Badiger, R. Yakkundimath, G. Konnurmath, and P. M. Dhulavvagol, "Deep Learning Approaches for Age-based Gesture Classification in South Indian Sign Language," *Engineering, Technology & Applied Science Research*, vol. 14, no. 2, pp. 13255–13260, Apr. 2024, <https://doi.org/10.48084/etasr.6864>.
- [18] C. Xia, A. Saito, and Y. Sugiura, "Using the virtual data-driven measurement to support the prototyping of hand gesture recognition interface with distance sensor," *Sensors and Actuators A: Physical*, vol. 338, May 2022, Art. no. 113463, <https://doi.org/10.1016/j.sna.2022.113463>.
- [19] W. Zhou and K. Chen, "A lightweight hand gesture recognition in complex backgrounds," *Displays*, vol. 74, Sep. 2022, Art. no. 102226, <https://doi.org/10.1016/j.displa.2022.102226>.
- [20] D. Chopra, N. Joshi, and I. Mathur, "A Review on Machine Translation in Indian Languages," *Engineering, Technology & Applied Science Research*, vol. 8, no. 5, pp. 3475–3478, Oct. 2018, <https://doi.org/10.48084/etasr.2288>.
- [21] S. R. Bose and V. S. Kumar, "In-situ recognition of hand gesture via Enhanced Xception based single-stage deep convolutional neural network," *Expert Systems with Applications*, vol. 193, May 2022, Art. no. 116427, <https://doi.org/10.1016/j.eswa.2021.116427>.
- [22] Z. Wang, G. Li, and L. Yang, "Dynamic Hand Gesture Recognition Based on Micro-Doppler Radar Signatures Using Hidden Gauss–Markov Models," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 2, pp. 291–295, Oct. 2021, <https://doi.org/10.1109/LGRS.2020.2974821>.
- [23] M. Al-Hammadi *et al.*, "Deep Learning-Based Approach for Sign Language Gesture Recognition With Efficient Hand Gesture Representation," *IEEE Access*, vol. 8, pp. 192527–192542, 2020, <https://doi.org/10.1109/ACCESS.2020.3032140>.
- [24] "Sign Language MNIST." [Online]. Available: <https://www.kaggle.com/datasets/datamunge/sign-language-mnist>.
- [25] L. K. S. Tolentino, R. O. S. Juan, A. C. Thio-ac, M. A. B. Pamahoy, J. R. R. Forteza, and X. J. O. Garcia, "Static Sign Language Recognition Using Deep Learning," *International Journal of Machine Learning and Computing*, vol. 9, no. 6, pp. 821–827, Dec. 2019, <https://doi.org/10.18178/ijmlc.2019.9.6.879>.
- [26] S. Chavan, X. Yu, and J. Saniie, "Convolutional Neural Network Hand Gesture Recognition for American Sign Language," in *2021 IEEE International Conference on Electro Information Technology (EIT)*, Mt. Pleasant, MI, USA, May 2021, pp. 188–192, <https://doi.org/10.1109/EIT51626.2021.9491897>.
- [27] S. Joudaki and A. Rehman, "Dynamic hand gesture recognition of sign language using geometric features learning," *International Journal of Computational Vision and Robotics*, vol. 12, no. 1, pp. 1–16, Jan. 2022, <https://doi.org/10.1504/IJCVR.2022.119239>.