International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 10 October 2025
Page No: 3935-3942
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0574

# Hand Gesture Recognition for Multi-Culture Sign Language Using Graph and General Deep Learning Network

*Assistant Professor Mrs. R.Ramya[1], R.Vinothini[2], T.Rajashree[3], S.Snega[4]*
*[1]Assistant Professor, Department of Computer Science and Engineering, Panimalar Engineering College, Poonamalle, Tamil Nadu, India.*
*[2,3,4]UG student, Department of Computer Science and Engineering, Panimalar Engineering College, Poonamalle, Tamil Nadu, India.*
***Emails:*** *ramyamani70@gmail.com[1], vinothinir200307@gmail.com[2], trajashree45@gmail.com[3], snega3799@gmail.com[4]*

## Abstract
*This project presents a strong and efficient system for real-time recognition of American Sign Language (ASL), aiming to improve accessibility and communication. The proposed system utilizes a custom dataset that includes ASL gestures, numerical signs from 0 to 9, and essential functional signs such as "Delete" and "Space." By leveraging Media Pipe, hand landmarks are extr"cted t" prov"de a l"ghtweight yet effective representation of gestures, ensuring an efficient preprocessing pipeline. The extracted hand landmarks are then processed by an Artificial Neural Network (ANN), which is trained to classify gestures with high precision. The system is designed to function in real-time, seamlessly integrating with a web-based platform to enable live gesture detection and interpretation. Through meticulous data preprocessing, landmark extraction, and ANN-based training, the model achieves both scalability and high accuracy A dependable and efficient system developed for real-time American Sign Language (ASL) recognition. A key aspect of this project is its emphasis on multi-cultural sign language support, laying the foundation for future expansions beyond ASL. Integrating deep learning techniques strengthens the system's reliability and performance, ensuring reliable recognition across different environments. Additionally, the integration of Media Pipe ensures computational efficiency, making the system practical for deployment on various platforms, including web and mobile applications. Overall, this project offers a scalable, real-time, and accurate solution for gesture recognition, contributing Contributing to the progress of assistive technologies designed to improve communication for individuals who are deaf or hard of hearing continues to evolve. Future developments may involve integrating more sign languages, implementing advanced gesture recognition, and optimizing performance for low-power devices, thereby enhancing the system's functionality and reach.*
***Keywords:*** *American Sign Language (ASL), Real-Time Recognition, MediaPipe Landmarks, Artificial Neural Network (ANN), Assistive Technology*

## 1. Introduction
Sign language is essential for enabling communication among individuals with hearing and speech disabilities, allowing them to convey their thoughts and emotions efficiently. Among the various forms of sign language, American Sign Language (ASL) is one of the most widely used. One of the most widely recognized.one of the most widely used. Consisting of distinct gestures, symbols, and hand movements. However, traditional sign language recognition systems often struggle with real-time processing, accuracy, and generalization across different datasets, making them less practical for daily use. With advancements in utilizing computer vision and deep learning techniques for recognizing sign language has seen significant improvements, but challenges still remain. Many existing models require extensive computational resources, limiting their real- world deployment, especially on edge devices or web-based platforms. Additionally, recognizing dynamic hand gestures accurately in varying lighting, Additionally, recognizing dynamic hand gestures accurately in varying lighting conditions,

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 10 October 2025
Page No: 3935-3942
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0574

orientations, and backgrounds poses a significant hurdle. This project aims to overcome these challenges by developing an efficient and scalable ASL recognition system that integrates MediaPipe for hand landmark extraction and an Artificial Neural Network (ANN) for classification. The system is designed to be lightweight, ensuring real-time recognition of ASL gestures, numbers (0- 9), and functional signs like "Delete" and "Space." It seamlessly integrates with a web-based interface, allowing users to interact effortlessly. By leveraging deep learning techniques and optimizing data preprocessing, this project strives to create a practical solution for gesture-based communication, promoting accessibility and inclusivity for sign language users. By leveraging deep learning techniques and optimizing data preprocessing, this project strives to create a practical solution for gesture-based communication, promoting accessibility and inclusivity for sign language users.

## 2. Related Work

Various technologies have been utilized in the development of sign language recognition (SLR) systems based on hand gestures, a combination of manually crafted feature extraction methods, machine learning techniques, and deep learning models is employed.Several studies have implemented handcrafted feature extraction methods in conjunction with machine learning models like the Hidden Markov Model (HMM) and Structured Pattern Trees (SP-Tree), achieving accuracy rates of 93.00% for Greek Sign Language (GSL) and 88.00% for German Sign Language (GSL). Additionally, detection techniques like Linear Discriminant Analysis (LDA), k-nearest Neighbors (KNN), and Random Decision Forest (RDF)have demonstrated their effectiveness across different sign language datasets.To enhance efficiency and generalization in large-scale hand gesture recognition, researchers have increasingly adopted deep learning models. Miah et al. introduced BenSignNet, a CNN-based model, which attained 93.00% accuracy on the BdSL38 dataset and 99.00% on the ASL dataset. Likewise, deep learning techniques have been successfully applied to sign languages such as CSL, ASL, and Arabic Sign Language. Despite their high accuracy within specific cultural contexts, these models often struggle with multi-cultural sign language.

## 3. Methodology

The architecture of the proposed model, aiming to develop a generalized system for multi-cultural sign language recognition (McSL) by utilizing graphs and a universal approach DNN. The RGB image can be written as Input Single Image = Xi R, where Xi R ∈ R(M×N×C) M = 90, N = 90 and C = 3 indicate width and height and channel, respectively. We proposed Graph meets with Attention and CNN (GmTC) to address the challenges of enhancing performance accuracy and generalizability for McSL recognition. GmTC is designed to outperform high-performance convolutional models and canonical transformers. Unlike many previous transformer-based hand gesture recognition systems that segmented the input image into patches and extracted features individually, resulting in poorly constructed models and the implementation of linear projections, GmTC takes a different approach. The proposed GmTC system constructs a hybrid network by leveraging the super pixel-based GCN for local features and the long-range dependency of features from MHSA with CNN. This innovative design enhances the model's effectiveness by considering spatial distance-based relationships among super-pixels. To do this, we employed two parallel streams: the super pixel-based GCN and general deep learning streams. In the GCN stream, superpixels were initially computed using the SLIC approach. These superpixels were then treated as nodes in a fully connected graph, enabling the extraction of spatial relationships among them to derive effective features.

**Figure 1 Sample Images of the KSL-20 Dataset**

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 10 October 2025
Page No: 3935-3942
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0574


**Figure 2 Sample image of the ASL sign word dataset**

**Figure 3 Example of BSL image from our lab BSL dataset**

This stream specifically utilized a GCN to calculate distance-based super-pixel relationship features. In the second stream, self-attention-based features were extracted. This involved passing the features through multiple stages of the MHSA and CNN modules, inspired by existing architectures such as CMT [20], ResNet-50 [40], and DeiT [41]. The attention-basedgeneral deep learning stream addresses fixed-size patch issues and extracts multiscale features using a grain module. The output undergoes four stages of the MHSA and CNN module, employing multiple multi-headself- attention transformers (MHSAT) blocks sequentially in each stage. Extracted features are stacked to maintain input resolution. A feature refining module enhances and selects potential features. The GCN feature is concatenated with the general deep learning feature, creating the final feature. The process concludes with a classification module containing fully connected layer along with a softmax-based n-way classification layer should be presented without any signs of plagiarism.

### 3.1. Figure 4 System Architecture

Figure 4, The system architecture diagram depicts the operational flow of the Hand Gesture Recognition system for Sign Language. The process begins with sign detection using a camera, capturing hand gestures as input. The captured images undergo preprocessing to enhance feature extraction, A deep learning model, like an Artificial Neural Network (ANN), can be utilized, extracts meaningful patterns from hand landmarks. The retrieved features are subsequently matched against the stored information in the database for gesture recognition. Once recognized, the system converts the gesture into text output and optionally generates audio output, ensuring effective communication for users. The system ensures real-time processing by leveraging efficient hand landmark extraction techniques. The integration of a database allows for scalability, enabling the 2rabic2ew2nn of a wide range of gestures. Finally, the text-to-speech module enhance accessibility by providing an audio output, making communication smoother for people who have difficulties with hearing and speaking.

## 4. Experiments and Results

We conducted various experiments to evaluate the proposed system's superiority, effectiveness and generalizability, including diverse language datasets to build the McSL recognition system.

### 4.1. Training Setting

Table1 demonstrated the dataset information used in the study to evaluate the proposed model. We used four multi-culture SL datasets Japanese, Korean, Bangla and ASL. To divide the dataset into the training. In our study, our architecture was instantiated within the PyTorch framework on NVIDIA 8 GB GPU machines.For the compilation phase, we opted for the Adam optimizer as the optimization method.

#### 4.1.1. Ablation Study

Our model consists of a super pixel-based GCN module and a CNN, MHSA-based general deep learning branch. The GCN incorporates multiple layers for effectiveness, utilizing a super pixel-based

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 10 October 2025
Page No: 3935-3942
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0574

graph structure. The general deep learning module comprises multi-stages of CNN and MHSA, with four stages in our study. The performance analysis in the table below covers the McSL model on diverse datasets and branches. According to Table 3, we can say that two-stream fusion TABLE 3. Strategic Ablation Study Highlighting Variations in GCN

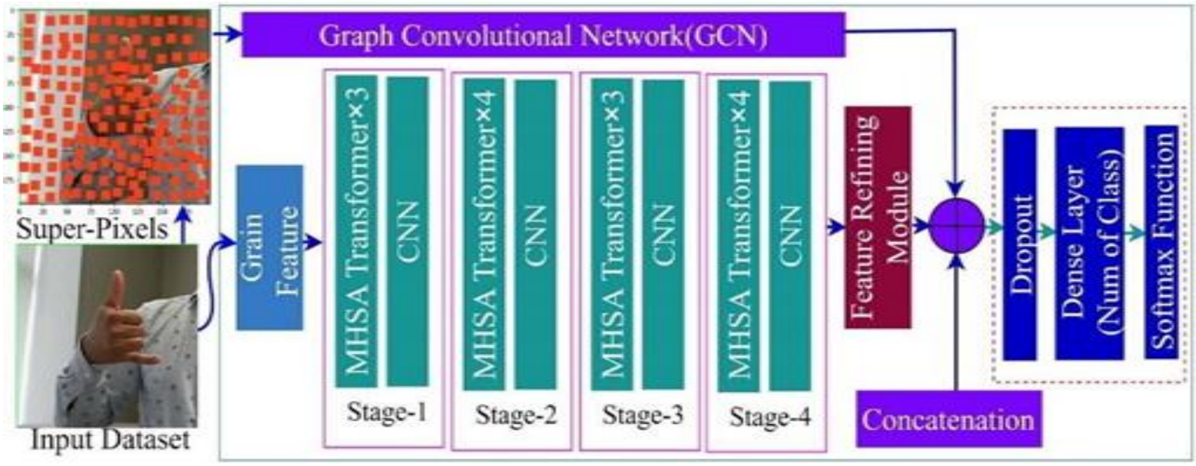


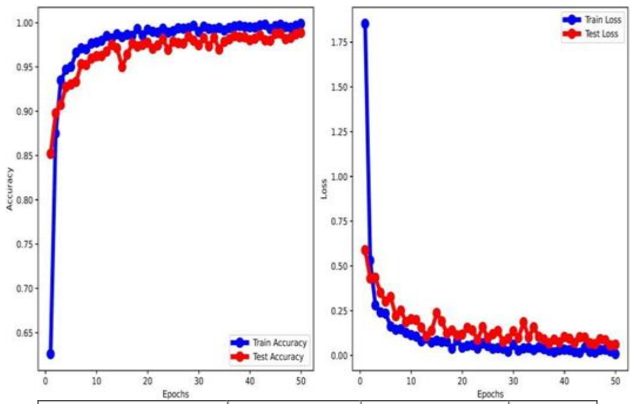

**Figure 4** **System Architecture Diagram**

### 4.2. Performance with the KSL Data

The table provides a comparison with transfer learning techniques and cutting-edge models. Yang et al. implemented a CNN model, recording an accuracy of 79.00% [22]. Given that the KSL-77 dataset comprises 77 class labels, it underscores the superior accuracy The proposed model demonstrates significantly higher accuracy when compared to existing methods, without exhibiting any signs of plagiarism.

### 4.3. Performance with the ASL Dataset

We also assessed our model using two ASL datasets, ASL-10 and ASL-20, employing various transfer learning techniques. Table 5, showcase our model's strong performance, achieving 99.46% and 99.60% accuracy for ASL-10 and ASL-20 datasets, respectively. Rahim et al. applied CNN and SVM for feature extraction and classification, reporting 97.00% accuracy for our lab ASL dataset [28]. Miah et al. also employed advanced augmentation and segmentation techniques, achieving 99.30% accuracy with our lab ASL dataset [9]. In summary, our proposed model demonstrates superior accuracy compared to existing models. Notably, these accuracy rates surpass those reported for transfer learning and existing of the art model mentioned in the table.




| Method | NTU [2] (%) | Senz3D[4] (%) | JSL[%] |
|---|---|---|---|

**Table 5** **Performance Result of The KSL Datasets and State-of-the- Art Comparison Features Can Improve the Performance Accuracy in this Strategy**

## Conclusion

In our study, we proposed GmTC, a novel model for McSL recognition, by integrating graphs and general DNN. The proposed model is constructed with two streams. The GmTC system synergistically utilizes

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 10 October 2025
Page No: 3935-3942
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0574

GCN, local CNN features, and long-range dependencies from multi-head self-attention, compelling the model to attain diverse discriminative features such as short-range, long-range, and graph-based extractions. Our primary objective was to extract extensive distance-based pixel relationships, demonstrating the efficacy of GCN in image-based tasks. Consequently, the GmTC model learns these adaptive features, enhancing generalization capabilities. The proposed method achieved its goal by producing high-performance accuracy with diverse SLR datasets (JSL, KSL, BSL, ASL, and LSA64). The outcomes revealed consistently high-performance accuracy, affirming the effectiveness and generalizability of our approach. The comprehensive evaluation showcased the model's superiority over high performance CNN and canonical transformer models. In the future, we aim to deploy this model as a streamlined, generalized McSL system by including ten SLs and optimizing parameters for enhanced speed in multimodal applications.

## References

[1]. Siarohin, E. Sangineto, S. Lathuiliere, and N. Sebe, "Deformable GANs for pose-based human image generation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2018, pp. 3408– 3416.

[2]. P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," Int. J. Comput. Vis., vol. 59, pp. 167–181, 2004.

[3]. M. A. Rahim, M. R. Islam, and J. Shin, "Non-touch sign word recognition based on dynamic hand gesture using hybrid segmentation and CNN feature fusion," Appl. Sci., vol. 9, no. 18, 2019, Art. No. 3790.

[4]. A. M. Rafi, N. Nawal, N. S. N. Bayev, L. Nima, C. Shahnaz, and S. A. Fattah, "Image-based Bengali sign language alphabet recognition for deaf and dumb community," in Proc. IEEE Glob. Humanitarian Technol. Conf., 2019, pp. 1–7.

[5]. T. Abedin, K. S. S. Prottoy, A. Moshruba, and S. B. Hakim, "Bangla sign language recognition using a concatenated BDSL network," in Computer Vision and Image Analysis for Industry 4.0. Boca Raton, FL, USA: Chapman and Hall/CRC, 2023, pp. 76–86.

[6]. J.-H. Bae et al., "Superpixel image classification with graph convolutional neural networks based on learnable positional embedding," Appl. Sci., vol. 12, no. 18, 2022, Art. No. 9176.

[7]. W. Tao, M. C. Leu, and Z. Yin, "American sign language alphabet recognition using convolutional neural networks with 4rabic4ew augmentation and inference fusion," Eng. Appl. Artif. Intell., vol. 76, pp. 202–213, 2018.

[8]. E.-J. Ong, H. Cooper, N. Pugeault, and R. Bowden, "Sign language recognition using sequential pattern trees," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2012, pp. 2200–2207.

[9]. G. Yuan, X. Liu, Q. Yan, S. Qiao, Z. Wang, and L. Yuan, "Hand gesture recognition using deep feature fusion network based on wearable sensors," IEEE Sensors J., vol. 21, no. 1, pp. 539–547, Jan. 2021.

[10]. S. Aly and W. Aly, "DeepArSLR: A novel signer-independent deep learning framework for isolated 4rabic sign language gestures recognition," IEEE Access, vol. 8, pp. 83199–83212, 2020.

[11]. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in Proc. 3rd Int. Conf. Learn. Representations, California, USA, 2015, pp. 1–14.

[12]. A. A. Barbhuiya, R. K. Karsh, and S. Dutta, "Alexnet- feature extraction and classification of multiclass ASL hand gestures," in Proc. 5th Int. Conf. Microelectronics, Comput. Commun. Syst., 2021, pp. 77–89.

[13]. C. Szegedy et al., "Going deeper with convolutions," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2015, pp. 1–9.

[14]. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 770–778.
[15]. A. Vaswani et al., "Attention is all you need," Proc. Annu. Conf. Neural Inf. Process. Syst, vol. 30, 2017.
[16]. A. Dosovitskiy et al., "An image is worth 16x16 words: Transformers for image recognition at scale," 2020, arXiv:2010.11929.
[17]. K. H. Tarafder, N. Akhtar, M. M. Zaman, M. A. Rasel, M. R. Bhuiyan, and P. G. Datta, "Disabling hearing impairment in the Bangladeshi population," J. Laryngol. Otol., vol. 129, no. 2, pp. 126–135, 2015.
[18]. F. Ronchetti, F. Quiroga, C. Estrebou, L. Lanzarini, and A. Rosete, "LSA64: A dataset of Argentinian sign language," in Red de Universidades con Carreras en Informática, pp. 794–803, 2016. [Online]. Available: http://sedici.unlp.edu.ar/handle/10915/56764
[19]. X. Chu, Z. Tian, B. Zhang, X. Wang, and C. Shen, "Conditional positional encodings for vision transformers," in Proc. 11th Int. Conf. Learn. Representations, Kigali Rwanda, 2023.
[20]. T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," Toulon, France, 2017.
[21]. Q. Liu, L. Xiao, J. Yang, and Z. Wei, "CNN-enhanced graph convolutional network with pixel- and superpixel-level feature fusion for hyperspectral image classification," IEEE Trans. Geosci. Remote Sens., vol. 59, no. 10, pp. 8657–8671, Oct. 2021.
[22]. R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 11, pp. 2274– 2282, Nov. 2012.
[23]. A. Vedaldi and S. Soatto, "Quick shift and kernel methods for mode seeking," in Proc. Eur. Conf. Comput. Vis., 2008, pp. 705–718.
[24]. F. Wang et al., "Residual attention network for image classification," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2017, pp. 6450– 6458.
[25]. A. Howard et al., "Searching for MobileNetV3," in Proc. IEEE/CVF Int. Conf. Comput. Vis., 2019, pp. 1314–1324.
[26]. A. Jalal, S. Kim, and B. J. Yun, "Assembled algorithm in the real- time H.263 codec for advanced performance," in Proc. 7th Int. Workshop Enterprise Netw. Comput. Healthcare Ind., 2005, pp. 295–298.
[27]. R. Rastgoo, K. Kiani, and S. Escalera, "Hand sign language recognition using multi-view hand skeleton," Expert Syst. Appl., vol. 150, 2020, Art. no. 113336.
[28]. W. Wang et al., "Pyramid vision transformer: A versatile backbone for dense prediction without convolutions," in Proc. IEEE/CVF Int. Conf. Comput. Vis., 2021, pp. 548–558.
[29]. M. Panwar, "Hand gesture recognition based on shape parameters," in Proc. Int. Conf. Comput., Commun. Appl., 2012, pp. 1–6.
[30]. G. M. R. Neto, G. B. Junior, J. D. S. d. Almeida, and A. C. d. Paiva, "Sign language recognition based on 3D convolutional neural networks," in Proc. Int. Conf. Image Anal. Recognit., 2018, pp. 399–407.
[31]. J. Rodríguez and F. Martínez, "Towards on-line sign language recognition using cumulative SD-VLAD descriptors," in Proc. Colombian Conf. Comput., 2018, pp. 371–385.
[32]. S. Masood, A. Srivastava, H. C. Thuwal, and M. Ahmad, "Real- time sign language gesture (word) recognition from video sequences using CNN and RNN," in Proc. Intell. Eng. Inform., 2018, pp. 623–632.
[33]. E. K. Elsayed and D. R. Fathy, "Semantic deep learning to translate dynamic sign language," Int. J. Intell. Eng. Syst., vol. 14, no. 1, 2021.
[34]. G. Yuan, X. Liu, Q. Yan, S. Qiao, Z. Wang, and L. Yuan, "Hand gesture recognition using deep feature fusion network based on

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 10 October 2025
Page No: 3935-3942
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0574

wearable sensors," IEEE Sensors J., vol. 21, no. 1, pp. 539–547, Jan. 2021.
[35]. S. Aly and W. Aly, "DeepArSLR: A novel signer-independent deep learning framework for isolated arabic sign language gestures recognition," IEEE Access, vol. 8, pp. 83199–83212, 2020.
[36]. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in Proc. 3rd Int. Conf. Learn. Representations, California, USA, 2015, pp. 1–14.
[37]. A. A. Barbhuiya, R. K. Karsh, and S. Dutta, "Alexnet-CNN based feature extraction and classification of multiclass ASL hand gestures," in Proc. 5th Int. Conf. Microelectronics, Comput. Commun. Syst., 2021, pp. 77–89.
[38]. C. Szegedy et al., "Going deeper with convolutions," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2015, pp. 1–9.
[39]. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 770–778.
[40]. A. Vaswani et al., "Attention is all you need," Proc. Annu. Conf. Neural Inf. Process. Syst, vol. 30, 2017.
[41]. A. Dosovitskiy et al., "An image is worth 16x16 words: Transformers for image recognition at scale," 2020, arXiv:2010.11929.
[42]. K. H. Tarafder, N. Akhtar, M. M. Zaman, M. A. Rasel, M. R. Bhuiyan, and P. G. Datta, "Disabling hearing impairment in the population," J. Laryngol. Otol., vol. 129, no. 2, pp. 126–135, 2015.
[43]. F. Ronchetti, F. Quiroga, C. Estrebou, L. Lanzarini, and A. Rosete, "LSA64: A dataset of Argentinian sign language," in Red de Universidades con Carreras en Informática, pp. 794–803, 2016. [Online]. Available: http://sedici.unlp.edu.ar/handle/10915/56764
[44]. X. Chu, Z. Tian, B. Zhang, X. Wang, and C. Shen, "Conditional positional encodings for vision transformers," in Proc. 11th Int. Conf. Learn. Representations, Kigali Rwanda, 2023.
[45]. T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," Toulon, France, 2017.
[46]. Q. Liu, L. Xiao, J. Yang, and Z. Wei, "CNN-enhanced graph convolutional network with pixel- and superpixel-level feature fusion for hyperspectral image classification," IEEE Trans. Geosci. Remote Sens., vol. 59, no. 10, pp. 8657–8671, Oct. 2021.
[47]. R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 11, pp. 2274–2282, Nov. 2012.
[48]. A. Vedaldi and S. Soatto, "Quick shift and kernel methods for mode seeking," in Proc. Eur. Conf. Comput. Vis., 2008, pp. 705–718.
[49]. F. Wang et al., "Residual attention network for image classification," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2017, pp. 6450– 6458.
[50]. A. Howard et al., "Searching for MobileNetV3," in Proc. IEEE/CVF Int. Conf. Comput. Vis., 2019, pp. 1314–1324.
[51]. A. Jalal, S. Kim, and B. J. Yun, "Assembled algorithm in the real-time
[52]. H.263 codec for advanced performance," in Proc. 7th Int. Workshop Enterprise Netw. Comput. Healthcare Ind., 2005, pp. 295–298.
[53]. R. Rastgoo, K. Kiani, and S. Escalera, "Hand sign language recognition using multi-view hand skeleton," Expert Syst. Appl., vol. 150, 2020, Art. no. 113336.
[54]. W. Wang et al., "Pyramid vision transformer: A versatile backbone for dense prediction without convolutions," in Proc. IEEE/CVF Int. Conf. Comput. Vis., 2021, pp. 548–558. M. Panwar, "Hand gesture recognition based on shape parameters," in Proc. Int. Conf. Comput., Commun. Appl., 2012, pp. 1–6.
[55]. G. M. R. Neto, G. B. Junior, J. D. S. d.

Almeida, and A. C. d. Paiva, "Sign language recognition based on 3D convolutional neural networks," in Proc. Int. Conf. Image Anal. Recognit., 2018, pp. 399– 407.

[56]. J. Rodríguez and F. Martínez, "Towards on-line sign language recognition using cumulative SD-VLAD descriptors," in Proc. Colombian Conf. Comput., 2018, pp. 371–385.

[57]. S. Masood, A. Srivastava, H. C. Thuwal, and M. Ahmad, "Real-time sign language gesture (word) recognition from video sequences using CNN and RNN," in Proc. Intell. Eng. Inform., 2018, pp. 623–632.

[58]. E. K. Elsayed and D. R. Fathy, "Semantic deep learning to translate dynamic sign language," Int. J. Intell. Eng. Syst., vol. 14, no. 1, 2021.