Deep Learning Approach for Sign Language Recognition using DenseNet201 with Transfer Learning

Yasir Altaf
Department of ComputerScience
Maulana Azad National Urdu University
Hyderabad, India
yasiraltafwani66@gmail.com

Abdul Wahid
Department of Computer Science and Infromation Technology
Maulana Azad National Urdu University
Hyderabad, India
wahidabdul76@yahoo.com

Mudasir Manzoor Kirmani
Department of Computer Science, Division of Social Science
FoFy SKUAST-K
Srinagar, India
mmkirmani@gmail.com

2023 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS) | 979-8-3503-9874-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/SCEECS57921.2023.10063044

***Abstract*—Transfer learning has been utilized to solve many complex real-world problems. Over the last several years, transfer learning had many applications in image and video recognition. To get higher recognition rates, deep and wider architectures of the convolutional neural networks (CNN) have been designed. In this research, we proposed a novel transfer learning-based model using a popular CNN architecture called DenseNet201 for the recognition of Indian Sign Language (ISL) hand gestures. We applied transfer learning to DenseNet201 by freezing some of its layers to retain its knowledge of generalization and fine-tuning the remaining layers for ISL dataset. Pre-trained DenseNet201 was used to extract the features of the gesture images. To classify the ISL gesture, custom layers were added to the pretrained DenseNet201 model. The proposed model helped to achieve higher accuracy of 100%.**

***Keywords*—Transfer Learning, DenseNet201, Deep Learning, Sign Language Recognition, Multilayer DenseNets, Image recognition.**

I. Introduction

The language used by deaf and mute people is called sign language (SL). It is used to propagate messages among them and also used by them to communicate with normal people. In every country, people with speech and hearing loss have their own ways of expressing ideas using sign language. For example, people living in America use American Sign Language (ASL) and the people of India use Indian Sign Language (ISL). Similarly, all other countries have their own sign language. However, due to the varied nature of sign language across countries, people face difficulties in exchanging their ideas when it is communication between two deaf-mute persons or between normal and deaf-mute persons. To bridge this communication gap, technology known as Sign Language Recognition and Translation plays an important role. The various benefits of such a technology are: (a) removing the communication barrier between the deaf-mute people, (b) assisting speech and hearing loss people in their daily life activities, such as studying, learning, travelling, marketing, banking, and health quires; (c) accelerating the teaching learning at schools meant for speech and hearing-impaired students, and (d) improving the digital literacy among deaf-mute people and also improving their lifestyle with enhanced communication and understanding. A number of automatic sign language recognition and translation systems have been developed over the last several years to meet the communication needs of deaf and hearing-impaired individuals [1], [2].This area is still an active area for research and requires improvements to implement such systems in various sectors, such as education to teach and assist speech and hearing loss students, public places such as railway stations, airports, and bus stops to guide deaf-mute people and help them by understanding their quires, providing feedback to them, and bringing them closer to the hearing world [3]–[5].With advancements in technologies researchers have adopted many approaches for sign language recognition, such as sensory-based, Vision-based and deep learning-based approaches are widely used. Traditionally, most sign language recognition and translation systems utilize contact-based approaches, wherein the signer needs some sort of special wearable device (like glove equipped with sensors) to perform gestures, and such systems did not grow in large scale for the development of digital assisting applications for deaf-mute people because of their obvious limitation with the hard-to-use circuitry of wearable devices. To simplify such systems, vision-based approaches provide new, easy-to-use, robust, and adequate methods for large-scale development of sign language-based assisting systems. Vision-based systems for sign language recognition and translation are easy to use and do not require any hard-to- use wearable devices; instead, they utilize high-quality cameras to capture the signs performed by the signer. In recent years, computer vision-based sign language recognition systems have achieved major advancements with the help of deep learning, and many companies and researchers have developed various applications and digit-assisting devices for deaf-mute people. With recent advancements of deep learning in machine learning, different convolutional neural network architectures have been employed for sign language recognition to achieve higher recognition accuracies [6].In this research, we propose a deep-learning-based method for the recognition of static gestures of numbers and letters in the ISL dataset. The proposed method uses DenseNet201 and transfer learning techniques to achieve higher recognition accuracies than most studies in the literature. The remainder of this paper is organized as follows. Section II presents a brief summary of various related works. Section III describes the proposed work, DenseNet architecture, and transfer learning using DenseNet201. Section IV presents the dataset and data processing technique used. Section V provides experimental results and evaluation metrics used for the performance analysis of the prosed model. Finally, section VI presents the conclusion and future works.

II. Related Work

For sign language recognition and sign language based assistive technology for deaf and mute people, many researchers worldwide have worked on different sign languages such as Indian [7]–[9], American [10], [11], Mexican [12], Chinese [13], [14], Brazilian [15], and Arabic [16]. And have developed various deep learning-based methods to support efficient and effective gesture recognition. Tang et al. [17] proposed a two-stage hand-gesture recognition system using a Kinect sensor and deep neural networks. The system used DNNs for automatic feature extraction and recognition of hand postures and achieved a recognition accuracy of 98.12%. Microsoft Kinect and a convolution neural network were used in [18] to develop a system for the recognition of 20 Italian gestures. This system is able to generalize different users and backgrounds data irrespective of the data that occurred during training. A validation accuracy of 91.7% was achieved by using this system. To deal with long-term motions (i.e., motions or actions that can span over a relatively long period of time), Duan et al. [19] presented a two-stream framework for gesture recognition using a convolution neural network with a consensus voting mechanism. The framework achieved better results than those its closest competitors. Chong et al. [6] presented a deep-learning-based approach for the recognition and interpretation of American signs. A wearable IMU sensory device was used to capture gestures, which were classified using the LSTM model. This model was used to recognition 28 commonly used words of the American sign language with an accuracy of up to 99.89%. To reduce the response time while recognizing hand gestures and converting them into audible sound using WebRTC, Gupta et al. [20] presented a hybrid CNN-LSTM model. This web-based model acts as a media for deaf-mute people to share their ideas with normal people. Another sign language recognition system based on the 3DCNN architecture was presented by Kraljevic et al. [21] for the use of deaf communities in their work environment. This system is able to recognize 25 different signs of Croatian sign language commonly used in home environments. With recent advancements, Wi-Fi based technologies have been adopted to enable pervasive and device-free gesture recognition [22]; such systems use off-shelf Wi-Fi devices for gesture recognition and support assisted living. Recently, deep learning-based methods [23]–[26] have been incorporated with Wi-Fi based technology to improve the performance of sign language-based assistive systems.

III. Proposed Work

in this work, we applied transfer learning to DenseNet201 for the recognition of ISL hand signs (as described in Section B). Transfer Learning is one of the most useful and widely used techniques in deep learning, which helps transfer the knowledge of the model already acquired while solving any other relevant problem to solve a new problem. Transfer learning not only saves time and computational resources but also achieves higher accuracy when solving new but relevant problems.

A. DenseNets

DenseNet is a convolutional neural network with dense connectivity between layers; that is, there are direct feed-forward connections from each layer to every other layer (as shown in Fig. 1). Therefore, the number of connections in a densenet of n layers is $n(n+1)/2$ [27]. The key benefits of DenseNets are: the resolution of vanishing gradients, deep supervision, and regularization to minimize overfitting that occurs due to the availability of small training datasets [28]. The dense connectivity in DenseNet helps to extract the deep features from an input image with the help of feature concatenation, that is, each layer receives the feature maps from its preceding layers. Mathematically, it is defined as:

$$A_i = H_i([A_0, A_1, \ldots, A_{i-1}]) \tag{1}$$

where, $[A_0, A_1, \ldots, A_{i-1}]$ represents the concatenation of the feature maps of all preceding layers of the ith layer, and H_i is a composite function of BN (branch normalization), Relu (rectified linear unit) and conv (a 3×3 convolution). The feature map concatenation in DenseNet enhances the information flow between layers and consequently improves its generalization abilities. Different variants of DenseNet have been used in several studies [29], [30]. In this study we used one of its architectures called DenseNet201 for the recognition of ISL hand signs.

B. Transfer Learning using DenseNet201

As previously mentioned, in this study, DenseNet201 was used for the recognition of ISL gestures using transfer learning. The DenseNet201 architecture (shown in Fig. 2 consists of 708 layers) is a deep, condensed, and parametrically compressed model because it contains transition layers that are used to reduce the spatial size of the input images and reduces their feature maps to a smaller number. Another

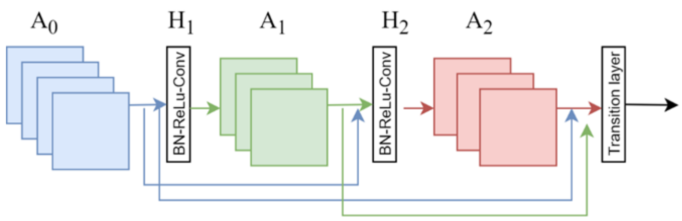

Fig. 1: A three-layer DenseNet architecture.

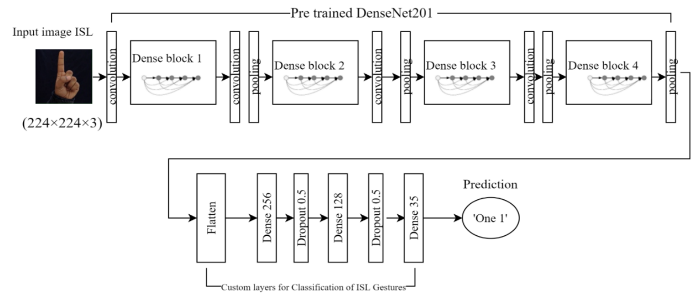

Fig. 3: Architecture of the proposed model.

advantage of DenseNet201 is that it is easy to train and achieve efficient parameter optimization through feature reuse. We used DenseNet201 to extract the features of the gesture images using its learned weights on the ImageNet dataset by freezing its layers. Finally, five custom layers, including two hidden layers with 256 and 128 neurons, two dropout layers with a value of 0.5, and one output layer with 35 neurons (equivalent to the number of classes in ISL dataset), were added, followed by softmax activation to classify the gestures. The overall architecture of our proposed model is shown in Fig 3.

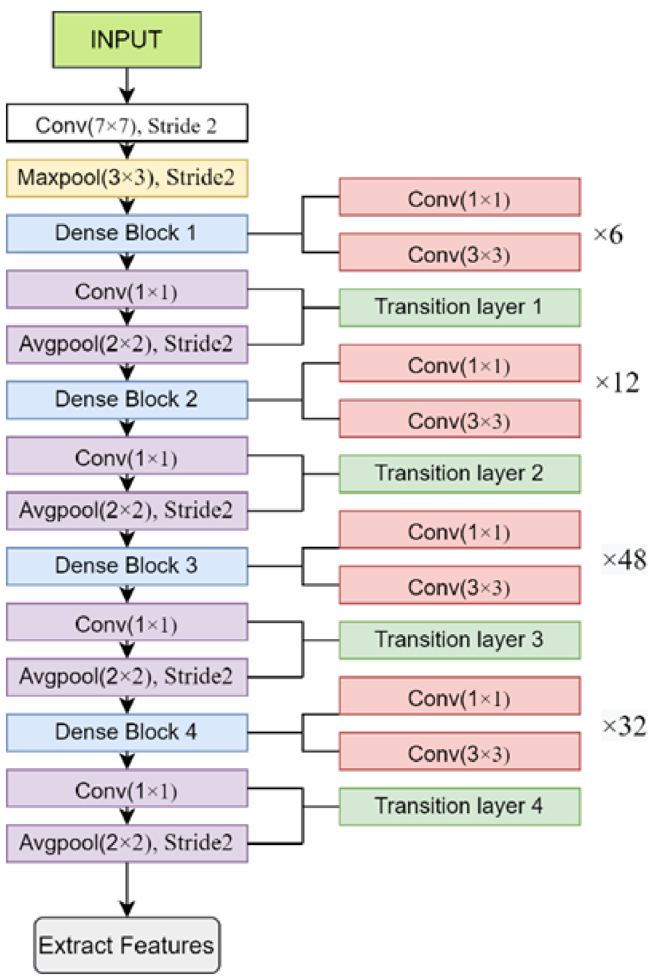

Fig. 2: DenseNet201 layered architecture.

IV. Data and Data Processing

The ISL hand-gesture image dataset available on Kaggle (https://www.kaggle.com/datasets/prathumarikeri/indian-sign-language-isl) was used in this experiment. This dataset is publicly available and consists of 42,000 RGB images of 35 hand signs including 9 numbers (1-9) and 26 letters (A-Z). Each image had a size of $128 \times 128 \times 3$, and we rescaled these images to a higher resolution of $224 \times 224 \times 3$ to increasing the performance of the model with higher resolution inspired by Thambawita et al. [31]. To validate the model, we divided the dataset into three sets: training, validation, and testing at a ratio 80:15:05 (3,360 images were used for training, 6,300 for validation, and 2,100 for testing). After rescaling the image to a higher resolution, data augmentation techniques were applied to the training set, including a shear range of 0.2, zoom up to 0.2, rotation up to 20° and horizontal filliping. A sample of the training data after the augmentation is shown in Fig. 4. All images were normalized to rescale the pixel values from the range of 0-255 to the range of 0-1 and were converted to numpy arrays for feature extraction. The description of the dataset used in this study is given in Table I.

Fig. 4: ISL hand gesture image dataset sample after augmentation.

TABLE I: Description of the data set used in this study

Dataset	Total No. of classes	No. of images in train set	No. of images in val. set	No. of images in test set	Total no. of images in data set
ISL	35	33,600	6,300	21,00	42,000

V. Results

In this research, the proposed DenseNet201 based transfer learning model was trained to recognize and classify ISL hand gestures. A pretrained DenseNet201 was used to extract the features of the gesture images. For transfer learning, we set the trainable property of the DenseNet201 layers to false, which prevented the weights (obtained by DenseNet201 from ImageNet) from being updated during training with the ISL dataset. The proposed model was trained for 10 epochs, and the validation loss was calculated for each epoch. The model is saved to the best with minimum validation loss, or the validation curve reaches saturation or a steady-state condition. Training data were provided to the model in batches with a batch size of 64. The optimization algorithm for our proposed model was Adam, with a learning rate of 0.001. The Adam optimization algorithm is used to optimize the 'gradient descent' by combining the two gradient decent methodologies namely 'momentum' and 'RMSP' algorithms. Mathematically, it is defined as:

$$W_{t+1} = W_t - \hat{m}_t \left(\frac{\alpha}{\sqrt{\hat{v}_t + \epsilon}} \right) \tag{2}$$

Where, W_t is the weights at time t, W_{t+1} is the weights at time t+1, $\hat{m}_t$ & $\hat{v}_t$ are bias- corrected weight parameters, α is the Learning rate parameter, and ϵ is a small positive constant (10^{-8}) to avoid 'division by zero' error.

The classification performance of the proposed model given in Fig.6 was accessed using the four metrics defined in Eqs. 3, 4, 5, and 6 and visualized using the confusion matrix given in Fig. 7. The training and validation performances of the proposed model were visualized by plotting their graphs, as shown in Fig. 5. The model achieved higher accuracies of 100% for both validation and test. Good values of precision, recall, and F1-score indicate that the model performance in classifying all classes accurately is satisfactory. A comparison of the proposed and existing deep learning-based sign language recognition models is presented in Table II.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{6}$$

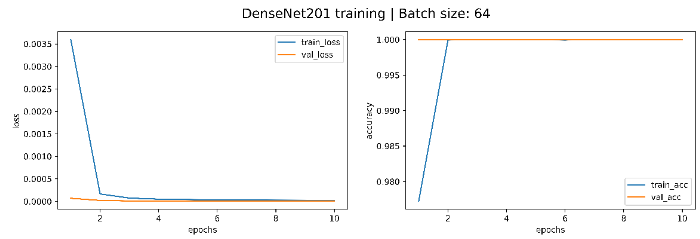

Fig. 5: Training results of the proposed model

	precision	recall	f1-score	support
0	1.00	1.00	1.00	60
1	1.00	1.00	1.00	60
2	1.00	1.00	1.00	60
3	1.00	1.00	1.00	60
4	1.00	1.00	1.00	60
5	1.00	1.00	1.00	60
6	1.00	1.00	1.00	60
7	1.00	1.00	1.00	60
8	1.00	1.00	1.00	60
9	1.00	1.00	1.00	60
10	1.00	1.00	1.00	60
11	1.00	1.00	1.00	60
12	1.00	1.00	1.00	60
13	1.00	1.00	1.00	60
14	1.00	1.00	1.00	60
15	1.00	1.00	1.00	60
16	1.00	1.00	1.00	60
17	1.00	1.00	1.00	60
18	1.00	1.00	1.00	60
19	1.00	1.00	1.00	60
20	1.00	1.00	1.00	60
21	1.00	1.00	1.00	60
22	1.00	1.00	1.00	60
23	1.00	1.00	1.00	60
24	1.00	1.00	1.00	60
25	1.00	1.00	1.00	60
26	1.00	1.00	1.00	60
27	1.00	1.00	1.00	60
28	1.00	1.00	1.00	60
29	1.00	1.00	1.00	60
30	1.00	1.00	1.00	60
31	1.00	1.00	1.00	60
32	1.00	1.00	1.00	60
33	1.00	1.00	1.00	60
34	1.00	1.00	1.00	60
accuracy			1.00	2100
macro avg	1.00	1.00	1.00	2100
weighted avg	1.00	1.00	1.00	2100

Fig. 6: Classification report of the proposed model.

TABLE II: Comparison of this study with existing related studies.

Study	Model	Dataset	Accuracy (%)
Abdul et al. [32]	Inception- BiLSTM	Arabic sign language (KSU-ArSL)	84.20
Molchanov et at. [33]	3DCNN	VIVA challenge dataset	77.05
Islam et al. [34]	VGG16	BdSL	99.92
Mistree et al. [35]	MobileNet	ISL	97.26
Wangchuk et al. [36]	CNN	BSL	97.62
Suri et al. [37]	CapsNet	Constructed	94
Katoch et al. [8]	CNN	ISL	99
Singh et al. [38]	3DCNN	ISL	88.24
Dhulipala et al. [39]	CNN	British Sign Language	97.40
Patil et al. [40]	CNN	ISL	95
Sharma et al. [9]	CNN	ISL	99.52
Nandi et al. [41]	CNN	ISL	99.76
This study	DenseNet201	ISL	100

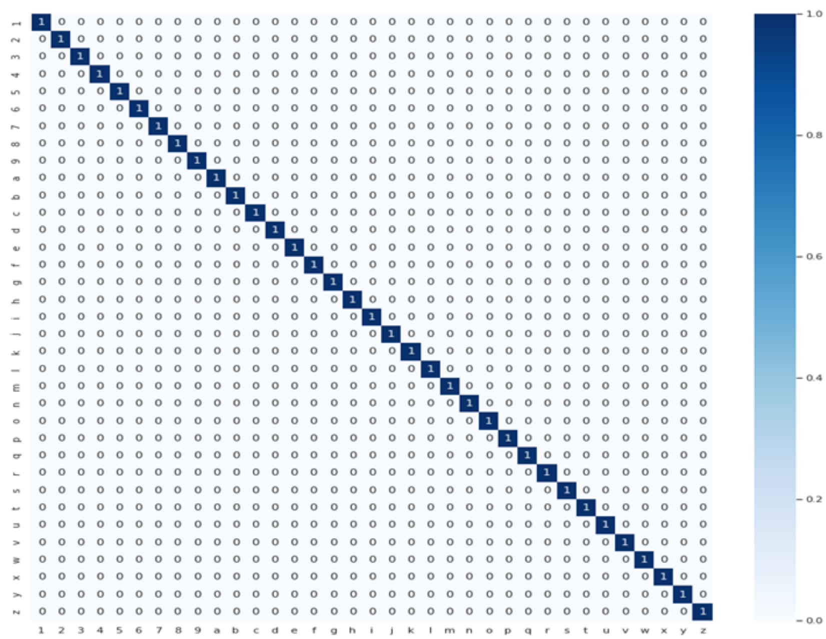

Fig. 7: Confusion Matrix of the proposed model.

VI. Conclusion and Future Work

In this paper, we are dedicated to the recognition of static hand gestures in Indian sign language with higher accuracy. We propose a novel transfer learning-based model using pre-trained DenseNet201 and a convolutional neural network (CNN). This model utilized the pre-trained DenseNet201 model for the extraction of features from the gesture images and the custom CNN layers for classifying these gesture images into their appropriate classes. This model achieved 100% accuracy for the ISL dataset. The accuracy achieved by this model is significantly higher than that of some existing sign language recognition studies. This model proved helpful in reducing loss during epochs without overfitting. Efficient and highly accurate automatic sign language recognition and translation systems are required at various places for deaf and mute communities to bridge their communication gap. This model can improve the recognition rate of such a system. In future work, we intend to extend this model to develop an efficient sign language-based learning assistive system for deaf and mute students.

References

[1] M. Sobhan, M. Z. Chowdhury, I. Ahsan, H. Mahmud and M. K. Hasan, "A Communication Aid System for Deaf and Mute using Vibrotactile and Visual Feedback," 2019 International Seminar on Application for Technology of Information and Communication (iSemantic), 2019, pp. 184-190, doi: 10.1109/ISEMANTIC.2019.8884323.

[2] K. Yousaf et al, "A Novel Technique for Speech Recognition and Visualization Based Mobile Application to Support Two-Way Communication between Deaf-Mute and Normal Peoples," Wireless Communications & Mobile Computing (Online), p. 12, 2018.

[3] E.E. Abdallah and E. Fayyoumi, "Assistive Technology for Deaf People Based on Android Platform," Procedia Computer Science, vol. 94, pp. 295-301, 2016.

[4] M.A. Azizan, I. Zulkiflee and N. Norhashim, "Development of Sign Language Translator for Speech Impairment Person," in Human-Centered Technology for a Better Tomorrow, Springer Singapore, 2022, pp. 115–131.

[5] S. Sundaramoorthy and B. Muthazhagan, "Human-Computer Interaction System for Improving Digital Literacy Among Speech- and Hearing-Impaired in Developing Countries," in Digital Literacy and Socio-Cultural Acceptance of ICT in Developing Countries, Cham, Springer International Publishing, 2021, pp. 139-156.

[6] T.W. Chong and B.J. Kim, "American Sign Language Recognition System Using Wearable Sensors with Deep Learning Approach," The Journal of the Korea institute of electronic communication sciences, vol. 15, no. 2, p. 291–298, 2020.

[7] N. Musthafa and C.G. Raji, "Real time Indian sign language recognition system," Materials Today: Proceedings, vol. 58, pp. 504-508, 2022.

[8] S. Katoch, V. Singh and U. S. Tiwary, "Indian Sign Language recognition system using SURF with SVM and CNN," Array, vol. 14, p. 100141, 2022.

[9] S. Sharma, and S. Singh, "Recognition of indian Sign Language (ISL) using Deep Learning Model.," Wireless personal communications, vol. 123, pp. 671-692, 2021.

[10] C.K.M. Lee, Kam K.H. Ng, Chun-Hsien Chen, H.C.W. Lau, S.Y. Chung, and T. Tsoi, "American sign language recognition and training method with recurrent neural network," Expert Systems with Applications, vol. 167, p. 114403, 2021.

[11] S. Sharma, and K. Kumar, "ASL-3DCNN: American sign language recognition technique using 3-D convolutional neural networks," Multimed Tools Appl, vol. 80, p. 26319–26331, 2021.

[12] K.M. Peréz, D.M.C. Esparza, J. Terven, A.M.H. Navarro, T.G. Ramírez, and A.R. Pedraza , "Automatic Recognition of Mexican Sign Language Using," Appl. Sci., vol. 12, p. 5523, 2022.

[13] Q. Xiao, M. Qin and Y. Yin, "Skeleton-based Chinese sign language recognition and generation for bidirectional communication between deaf and hearing people," Neural Networks, vol. 125, pp. 41-55, 2020.

[14] J. Zhang, W. Zhou, C. Xie, J. Pu and H. Li, "Chinese sign language recognition with adaptive HMM," in IEEE International Conference on Multimedia and Expo (ICME), 2016.

[15] S.G.M. Almeida, F.G. Guimarães, and J.A. Ramírez, "Feature extraction in Brazilian Sign Language Recognition based on phonological structure and using RGB-D sensors," Expert Systems with Applications, vol. 41, no. 16, pp. 7259-7271, 2014.

[16] M. Elpeltagy,M. Abdelwahab,M. E. Hussein,A. Shoukry,A. Shoala, and M. Galal, "Multi-modality-based Arabic sign language recognition," IET Comput. Vis., vol. 12, pp. 1031-1039, 2018.

[17] A. Tang, K. Lu, Y. Wang, J. Huang and H. Li, "A Real-time Hand Posture Recognition System Using Deep Neural," ACM Trans. Intell. Syst. Techno, vol. 9, no. 4, p. 23, 2013.

[18] L. Pigou, S. Dieleman, P. J. Kindermans and B. Schrauwen, "Sign Language Recognition Using Convolutional Neural Networks," in Computer Vision - ECCV 2014 Workshops, 2015.

[19] J. Duan, J. Wan, S. Zhou, X. Guo, and S. Z. Li, "A Unified Framework for Multi-Modal Isolated Gesture Recognition," ACM Trans. Multimedia Comput. Commun. Appl, vol. 9, no. 4, p. 17, 2017.

[20] Gupta M, Thakur N, Bansal D, Chaudhary G, Davaasambuu B, and Hua Q, "CNN-LSTM Hybrid Real-Time IoT-Based Cognitive Approaches for ISLR with WebRTC: Auditory Impaired Assistive Technology," J Healthc Eng, 2022.

[21] L. Kraljević, M. Russo, M. Pauković, and M. Šarić, "A Dynamic Gesture Recognition Interface for Smart Home Control based on Croatian Sign Language," Applied Sciences, vol. 10, no. 7, 2020.

[22] H. Farhana, T. Ahmed, H. Ahmad, K. Narasingamurthi , H. Harkat and S. K. Phang, "DF-WiSLR: Device-Free Wi-Fi-based Sign Language Recognition," Pervasive and Mobile Computing, vol. 69, p. 101289, 2020.

[23] M. G. Martín, R. S. Segundo, F. F. Martínez, and J. F. López, "Improving physical activity recognition using a new deep learning architecture and post-processing techniques," Engineering Applications of Artificial Intelligence, vol. 92, p. 103679, 2020.

[24] Y. Ma, G. Zhou, S. Wang, H. Zhao, and W. Jung, "SignFi: Sign Language Recognition," in Proc. ACM Interact. Mob. Wearable Ubiquitous Technol, 2018.

[25] Y. Xu, W. Yang, M. Chen, S. Chen and L. Huang, "Attention-Based Gait Recognition and Walking Direction Estimation in Wi-Fi Networks," IEEE Transactions on Mobile Computing, vol. 21, no. 2, pp. 465-479, 2022.

[26] H. Xue, W. Jiang, C. Miao, F. Ma, S. Wang, Y. Yuan, S. Yao, A. Zhang, and L. Su, "DeepMV: Multi-View Deep Learning for Device-Free Human Activity Recognition," in Proc. ACM Interact. Mob. Wearable, 2020.

[27] G. Huang, Z. Liu, L. Van Der Maaten and K. Q Weinberger, "Densely Connected Convolutional Networks," in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[28] M. Z. Hossain, F. Sohel, M. F. Shiratuddin, and H. Laga, "Attention-Based Image Captioning Using DenseNet Features," in Neural Information Processing, Cham, 2019.

[29] B. Aksoy , O. K. M. Salman , and Ö. Ekrem, "Detection of Turkish Sign Language Using Deep," Applid Artifial Intelligence, vol. 35, no. 12, p. 952–981, 2021.

[30] M. Kołodziej, E. Szypuła, A. Majkowski and R. Rak, "Using deep learning to recognize the sign alphabet," Przeglad Elektrotechniczny, vol. 6, p. 32, 2022.

[31] V. Thambawita, I. Strümke, S.A. Hicks, P. Halvorsen, S. Parasa and M. A. Riegler, "Impact of Image Resolution on Deep Learning Performance in Endoscopy Image Classification: An Experimental Study Using a Large Dataset of Endoscopic Images," Diagnostics-Basel, vol. 11, no. 12, p. 2183, 2021.

[32] W. Abdul, M. Alsulaiman, S.U. Amin, M. Faisal, G. Muhammad, F. R. Albogamy, M. A. Bencherif and H. Ghaleb, "Intelligent real-time Arabic sign language classification using attention-based inception and BiLSTM," Computers and Electrical Engineering, vol. 95, p. 107395, 2021.

[33] P. Molchanov, S. Gupta, K. Kim and J. Kautz, "Hand gesture recognition with 3D convolutional neural networks," in IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2015.

[34] M. M. Islam, M.R. Uddin, M. N. AKhtar and K.M. R. Alam, "Recognizing multiclass Static Sign Language words for deaf and dumb people of Bangladesh based on transfer learning techniques," Informatics in Medicine Unlocked, vol. 33, 2022.

[35] K. Mistree, D. Thakor and B. Bhatt, "Indian Alphabets and Digits Sign Recognition Using Pretrained Model," Smart Innovation, Systems and Technologies, vol. 283, 2022.

[36] K. Wangchuk, P. Riyamongkol and R. Waranusast, "Real-time Bhutanese Sign Language digits recognition system using Convolutional Neural Network," ICT Express, vol. 7, no. 2, pp. 215-220, 2021.

[37] K. Suri and R. Gupta, "Continuous sign language recognition from wearable IMUs using deep capsule networks and game theory," Computers & Electrical Engineering, vol. 78, pp. 493-503, 2019.

[38] D. Singh, "3D-CNN based Dynamic Gesture Recognition for Indian Sign Language Modeling," Procedia Computer Science, vol. 189, pp. 76-83, 2021.

[39] S. Dhulipala, F. F. Adedoyin and A. Bruno , "Sign and Human Action Detection Using Deep Learning," J. Imaging, vol. 8, no. 192, pp. 1-34, 2022.

[40] R. Pati, V. Pati, A. Bahuguna and G. Datkhile, "Indian Sign Language Recognition using Convolutional Neural Network," in ITM Web of Conferences (ICACC), 2021.

[41] U. Nandi, A. Ghorai, M. M. Singh, C. Changdar, S. Bhakta and R. K. Pal, "Indian sign language alphabet recognition system using CNN with diffGrad optimizer and stochastic pooling," Multimedia Tools and Applications, 2022.