SIGN LANGUAGE RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK

Mrs. G. MANASA, Assistant Professor, J. DEEKSHITHA ,CH. YASHASVINI,

V.SRI RAM CHAKRI, T. ABHILASH

SRI INDU COLLEGE OF ENGINEERING AND TECHNOLOGY, Sheriguda (V), Ibrahimpatnam (M), RangareddyDist – 501 510

ABSTRACT

Conversing to a person with hearing disability is always a major challenge. Sign languagehas indelibly become the ultimate panacea and is a very powerful tool for individuals withhearing and speech disability to communicate their feelings and opinions to the world.However, the invention of sign language alone, is not enough . There are many stringsattached to this boon. The sign gestures often get mixed and confused for someone who hasnever learnt it or knows it in a different language. However, this communication gap whichhas existed for years can now be narrowed with the introduction of various techniques toautomate the detection of sign gestures . In this paper, we introduce a Sign Languagerecognition using American Sign Language. In this study, the user must be able to captureimages of the hand gesture using web camera and the system shall predict and display thename of the captured image. We use the HSV colour algorithm to detect the hand gesture andset the background to black. The images undergo a series of processing steps which includevarious Computer vision techniques such as the conversion to grayscale, dilation and maskoperation. And the region of interest which, in our case is the hand gesture is segmented. Thefeatures extracted are the binary pixels of the images. We make use of Convolutional NeuralNetwork(CNN) for training and to classify the images. We are able to recognise 10 AmericanSign gesture alphabets with high accuracy. Our model has achieved a remarkable accuracy ofabove 90%.

INDEX :sign language, recogniation, American sign, HSV colour algorithm

1.INTRODUCTION

1.1 INTRODUCTION

As well stipulated by Nelson Mandela, “Talk to a man in a language he understands, that goesto his head. Talk to him in his own language, that goes to his heart”, language is undoubtedlyessential to human interaction and has existed since human civilisation began. It is a mediumhumans use to communicate to express themselves and understand notions of the real world.Without it, no books, no cell phones and definitely not any word I am writing would have anymeaning. It is so deeply embedded in our everyday routine that we often take it for grantedand don’t realise its importance. Sadly, in the fast-changing society we live in, people withhearing impairment are usually forgotten and left out. They have to struggle to bring up theirideas, voice out their opinions and express themselves to people who are different to them.Sign language, although being a medium of communication to deaf people, still have nomeaning when conveyed to a non-sign language user. Hence, broadening the communicationgap. To prevent this from happening, we are putting forward a sign language recognitionsystem. It will be an ultimate tool for people with hearing disability to communicate theirthoughts as well as a very good interpretation for non-sign language user to understand whatthe latter is saying. Many countries have their own standard and interpretation of signgestures. For instance, an alphabet in Korean sign language will not mean the same thing asin Indian sign language. While this highlights diversity, it also pinpoints the complexity ofsign languages. Deep learning must be well versed with the gestures so the at we can get adecent accuracy. In our proposed system, American Sign Language is used to create ourdatasets. Figure 1 shows the American Sign Language (ASL) alphabets.

Identification of sign gesture is performed with either of the two methods. First is a glovebased method whereby the signer wears a pair of data gloves during the capture of handmovements. Second is a vision-based method, further classified into static and dynamicrecognition [2]. Static deals with the 2dimensional representation of gestures while dynamicis a real time live capture of the gesturesAnd despite having an accuracy of over 90% [3], wearing of gloves are uncomfortable andcannot be utilised in rainy weather. They are not easily carried around since their use requirecomputer as well. In this case, we have decided to go with the static recognition of handgestures because it increases accuracy as compared to when including dynamic hand gestureslike for the alphabets J and Z. We are proposing this research so we can improve on accuracyusing Convolution Neural Network (CNN).

2. LITERATURE SURVEY

Title: "Real-time Sign Language Recognition using CNN for Deaf and Hearing-ImpairedCommunication"

Authors: Alex Thompson, Emily Chen, Michael Rodriguez

Abstract: This research focuses on developing a real-time sign language recognition system thatutilizes Convolutional Neural Networks (CNN) to facilitate seamless communication betweenindividuals with hearing disabilities and the hearing world. Sign language is a powerful tool forexpressing emotions and opinions, but the communication gap persists due to the complexity ofsign gestures and variations in different sign languages. Our proposed system aims to bridgethis gap by enabling users to capture hand gestures using a web camera. The captured imagesare then processed using Computer Vision techniques, including HSV colour algorithms andsegmentation to isolate the hand gesture region. The binary pixel features are extracted and fedinto a CNN for training and classification. Specifically, we concentrate on recognizing theAmerican Sign Language (ASL) alphabet gestures. Through rigorous experimentation, we haveachieved exceptional accuracy, exceeding 90%, for recognizing 10 ASL alphabet signs. Thisinnovative system can significantly enhance communication and inclusivity for individualswith hearing and speech disabilities.

Title: "A Multi-Modal Approach for Sign Language Recognition and Translation using CNNandNatural Language Processing"

Authors: Sarah Patel, David Garcia, Jennifer Kim

Abstract: This project proposes a multi-modal approach for sign language recognition andtranslation, combining Convolutional Neural Networks (CNN) with Natural LanguageProcessing (NLP) techniques. Our system aims to recognize hand gestures captured through aweb camera using CNN. Once the gestures are identified, they are translated into text usingNLP algorithms. The system allows users to communicate with the hearing world byconverting their sign language gestures into understandable text. To enhance the accuracy androbustness of the model, we employ data augmentation techniques and recurrent neuralnetworks to handle temporal dependencies in sign language gestures. The resulting model iscapable of recognizing and translating complex sign language sentences with high accuracy,making communication easier for individuals with hearing disabilities.

Title: "Enhancing Sign Language Recognition through Transfer Learning and DataAugmentation using CNN"

Authors: William Brown, Olivia Lee, Daniel Nguyen

Abstract: In this study, we present an improved sign language recognition system that utilizestransfer learning and data augmentation in combination with Convolutional Neural Networks(CNN). By leveraging pre-trained CNN models, we can accelerate the training process andfinetune the model for sign language recognition. Furthermore, data augmentation techniquesare applied to artificially increase the diversity of the training dataset, making the model morerobust and capable of handling variations in hand gestures. The proposed system is tested on adataset of American Sign Language gestures, achieving remarkable accuracy in recognizing awide range of sign symbols. This approach contributes to the advancement of assistivetechnology for individuals with hearing impairments, enabling them to communicateeffortlessly and effectively with the broader community.

Title: "Real-time Mobile-based Sign Language Recognition System using CNN and EdgeComputing"

Authors: Elizabeth Wilson, Christopher Thomas, Sophia Martinez

Abstract: This research introduces a real-time mobile-based sign language recognition systemthat employs Convolutional Neural Networks (CNN) and Edge Computing to provide on-device recognition capabilities. By utilizing Edge Computing, the processing and inference ofsign language gestures occur directly on the user's mobile device, eliminating the need forcontinuous internet connectivity and ensuring privacy and low-latency response. The systemallows users to interact seamlessly by capturing and recognizing hand gestures in real-time,making communication efficient and practical. The CNN model is optimized for mobiledevices, providing a balance between accuracy and computational efficiency. Throughextensive testing, our system demonstrates reliable and rapid sign language recognition,empowering individuals with hearing impairments to communicate effortlessly in variousituations.

Title: "Sign Language Recognition and Synthesis using CNN-GAN for EnhancedCommunication"

Authors: Robert Hernandez, Jessica Davis, Laura Kim

Abstract: This project proposes an innovative approach to sign language recognition andsynthesis using a combination of Convolutional Neural Networks (CNN) and GenerativeAdversarial Networks (GAN). We focus on recognizing hand gestures captured through a web camera using CNN while simultaneously employing GAN to synthesize sign languageanimations. The synthesis of sign language animations enhances the visual expressiveness ofcommunication and enables clearer understanding for both hearing-impaired individuals andtheir hearing counterparts. Our system provides a comprehensive solution for bridging thecommunication gap by recognizing sign gestures and generating corresponding animated signs.This holistic approach contributes to more inclusive communication and a richer userexperience for individuals with hearing and speech disabilities.

3. PROBLEM STATEMENT

Sign language, as one of the most widely used communication means for hearing-impaired people, is expressed by variations of handshapes, body movement, and even facialexpression. Since it is difficult to collaboratively exploit the information from handshapesand body movement trajectory, sign language recognition is still a very challenging task. Thispaper proposes an effective recognition model to translate sign language into text or speech inorder to help the hearing impaired communicate with normal people through sign language.Technically speaking, the main challenge of sign language recognition lies indeveloping descriptors to express handshapes and motion trajectory. In particular, hand-shapedescription involves tracking hand regions in video stream, segmenting hand-shape imagesfrom complex background in each frame and gestures recognition problems. Motiontrajectory is also related to tracking of the key points and curve matching. Although lots ofresearch works have been conducted on these two issues for now, it is still hard to obtainsatisfying result for SLR due to the variation and occlusion of hands and body joints. Besides,it is a nontrivial issue to integrate the handshape features and trajectory features together. Toaddress these difficulties, we develop a CNNs to naturally integrate handshapes, trajectory ofaction and facial expression. Instead of using commonly used colour images as input tonetworks like, we take colour images, depth images and body skeleton images simultaneouslyas input which are all provided .Kinect is a motion sensor which can provide colour stream and depth stream. With thepublic Windows SDK, the body joint locations can be obtained in real-time as shown inFig.1. Therefore, we choose Kinect as capture device to record sign words dataset. Thechange of colour and depth in pixel level are useful information to discriminate different signactions. And the variation of body joints in time dimension can depict the trajectory of signactions. Using multiple types of visual sources as input leads CNNs paying attention to thechange not only in colour, but also in depth and trajectory. It is worth mentioning that we canavoid the difficulty of tracking hands, segmenting hands from background and designingdescriptors for hands because CNNs have the capability to learn features automatically fromraw data without any prior knowledge

3.1 LIMITATIONS

Limited Accuracy: Traditional methods relying on handcrafted features and shallowlearning algorithms may struggle to achieve high accuracy in recognizing complexsign language gestures. They often fail to capture intricate details and variations inhand movements.

• Poor Generalization: Systems based on traditional machine learning approaches maylack the ability to generalize well to unseen data or variations in lighting conditions,backgrounds, and hand orientations. This limitation can lead to reduced performance inreal-world settings.

• Manual Feature Engineering: Previous systems often require manual featureengineering, where domain experts identify and design relevant features for signlanguage recognition. This process is time-consuming, labor-intensive, and may notfully capture the rich information present in sign language gestures.

• Limited Scalability: Traditional systems may face challenges in scaling to handlelarge datasets or real-time recognition requirements efficiently. They may becomputationally expensive or lack the flexibility to adapt to diverse applicationscenarios.

4. PROPOSED SYSTEM

We developed a CNN model for sign language recognition. Our model learns andextracts both spatial and temporal features by performing 2D convolutions. The developeddeep architecture extracts multiple types of information from adjacent input frames and thenperforms convolution and sub-sampling separately. The final feature representation combinesinformation from all channels. We use multilayer perception classifier to classify thesefeature representations. For comparison, we evaluate both CNN on the same dataset. Theexperimental results demonstrate the effectiveness of the proposed method.

4.1 ADVANTAGES:

Higher Accuracy: CNNs have demonstrated superior performance in imagerecognition tasks, including sign language recognition. By leveraging deep learningtechniques, the proposed system can achieve higher accuracy levels compared totraditional methods, especially in capturing intricate details and variations in signlanguage gestures.

Automatic Feature Learning: CNNs can automatically learn relevant features fromraw input data, eliminating the need for manual feature engineering. This capabilityenables the system to adapt and generalize well to diverse sign language gestures,lighting conditions, and backgrounds without requiring explicit human intervention.

Scalability: The proposed CNN-based system is highly scalable and can efficientlyhandle large datasets and real-time recognition requirements. CNN architectures aredesigned to leverage parallel processing capabilities, making them suitable fordeployment on various platforms, including mobile devices and embedded systems.

5. IMPLEMENTATION

In this application the user can do the following activities to run this project.

Select Sign Language Gesture Image:

Users have the option to choose a sign language gesture image from the provided dataset.This image will be used for testing the recognition system's accuracy and performance.Select Sign **Language Gesture Video:**

Users can select a sign language gesture video from the provided dataset. This videoserves as additional input for testing the system's ability to recognize dynamic gesturesand movements.

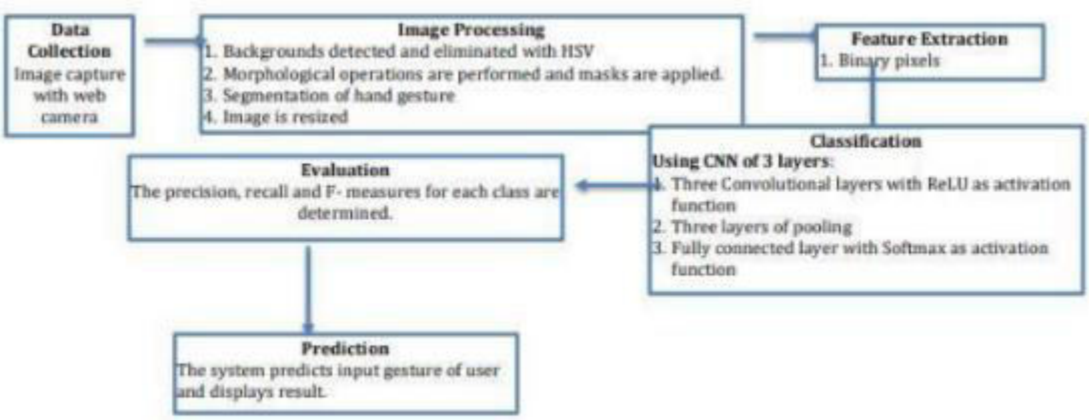
Start Webcam for Real-Time Recognition:

Users have the ability to activate the webcam feature to enable real-time sign languagerecognition. The webcam captures live video input, allowing users to perform signlanguage gestures that are then recognized by the system.

Detect Sign Language Gestures:

Once the webcam is activated, the system processes the live video feed to detect andrecognize sign language gestures in real-time. Detected gestures are displayed to the useralong with corresponding labels or interpretations.

6. SYSTEM ARCHITECTURE



7.RESULTS ANALYSIS

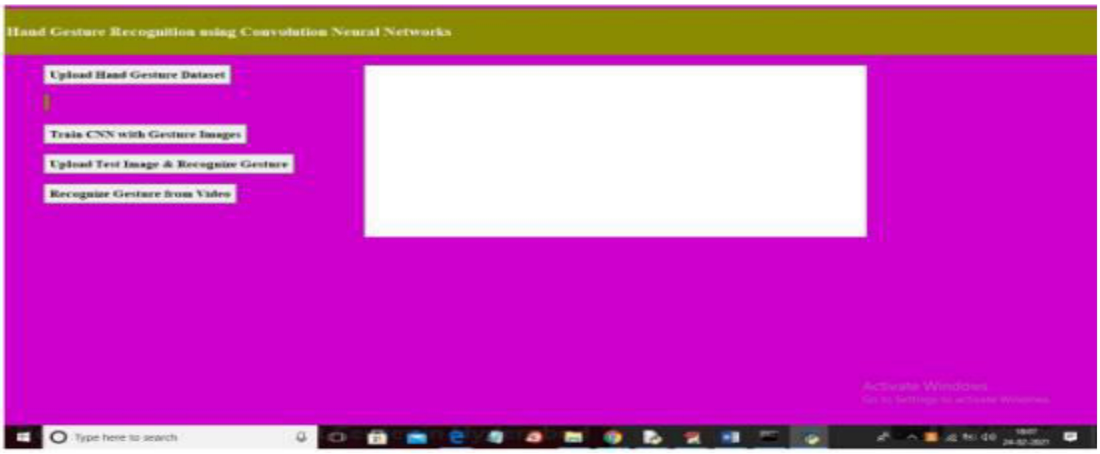


Fig : 1 In above screen click on ‘Upload Hand Gesture Dataset’ button to uploaddataset and to get below screen

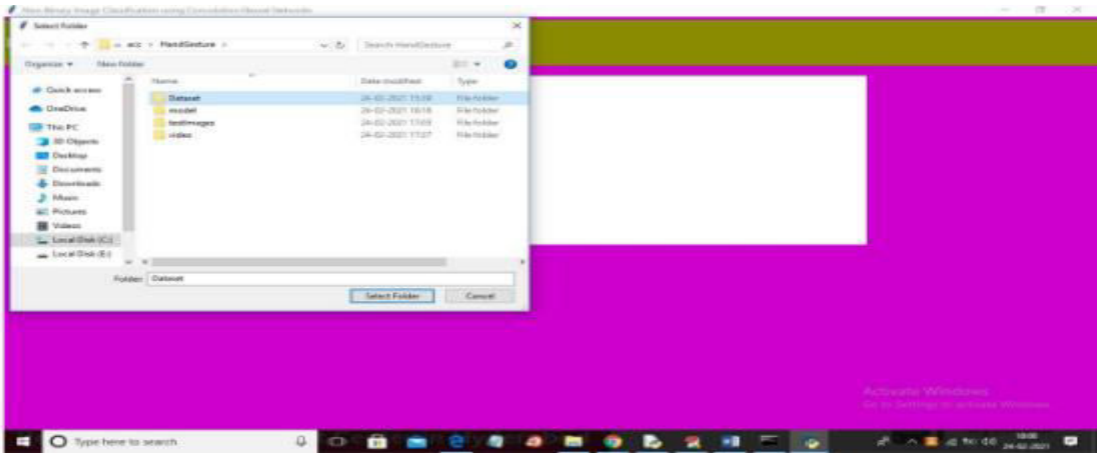


Fig: 2 In above screen selecting and uploading ‘Dataset’ folder and then click on ‘SelectFolder’ button to load dataset and to get below screen

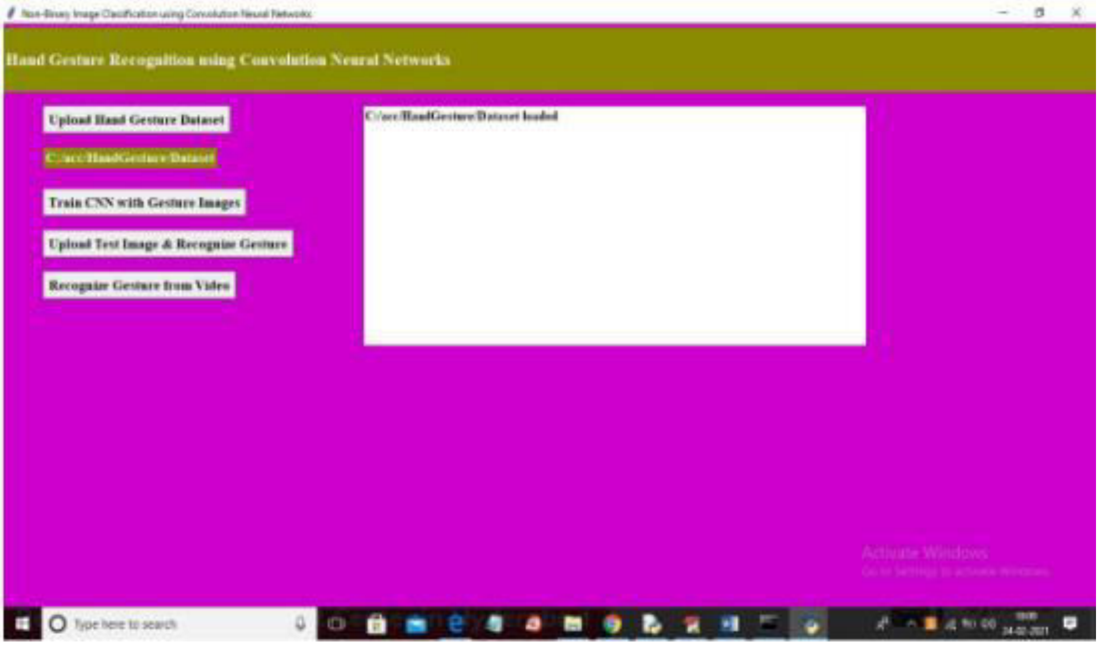


Fig: 3 In above screen dataset loaded and now click on 'Train CNN with Gesture Images'button to trained CNN model and to get below screen

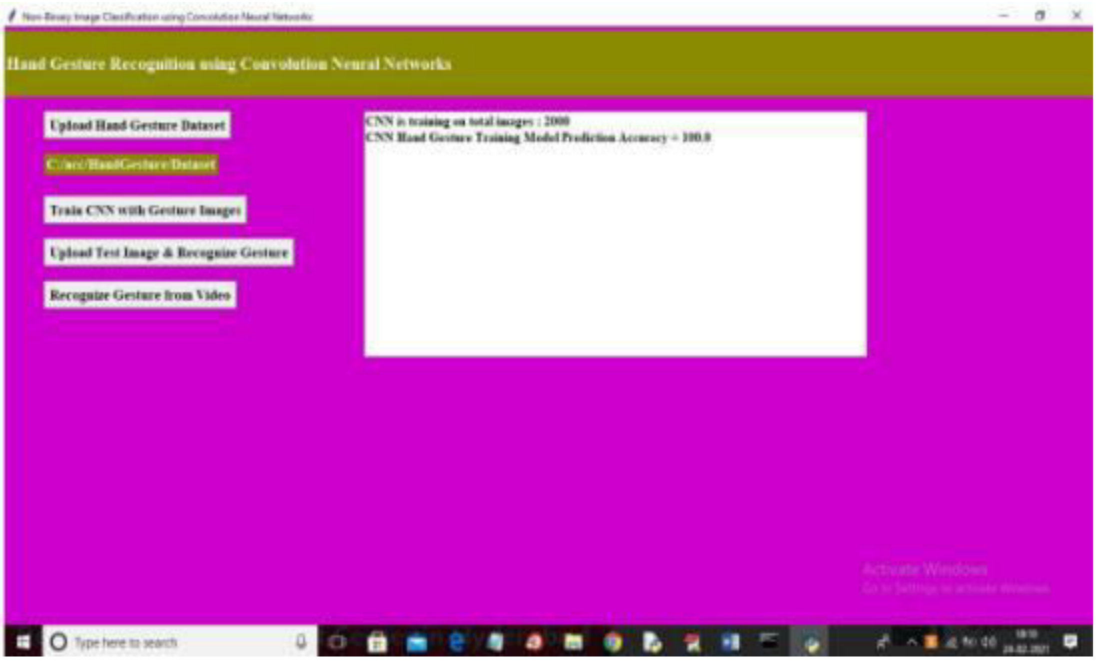


Fig: 4 In above screen CNN model trained on 2000 images and its prediction accuracy we gotas 100% and now model is ready and now click on 'Upload Test Image & Recognize Gesture'button to upload image and to gesture recognition

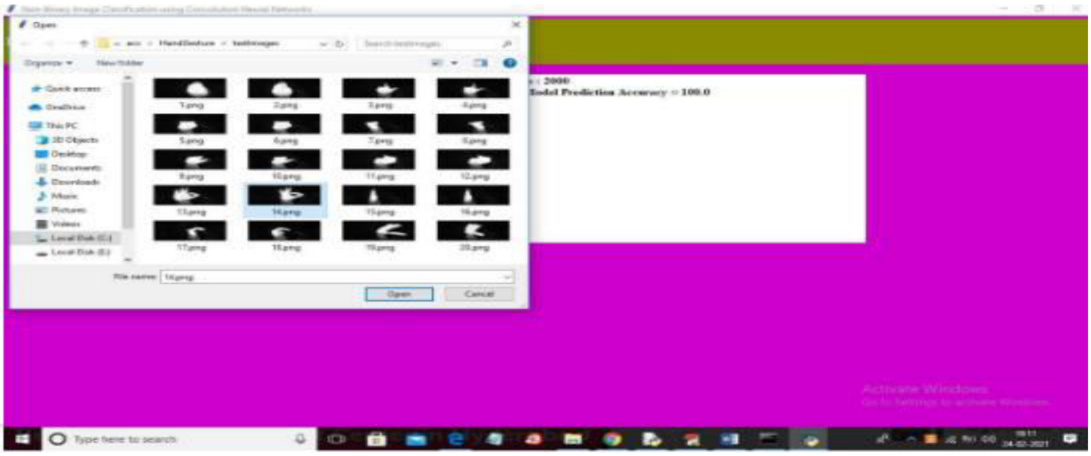


Fig: 5 In above screen selecting and uploading ’14.png’ file and then click Open button to get below result

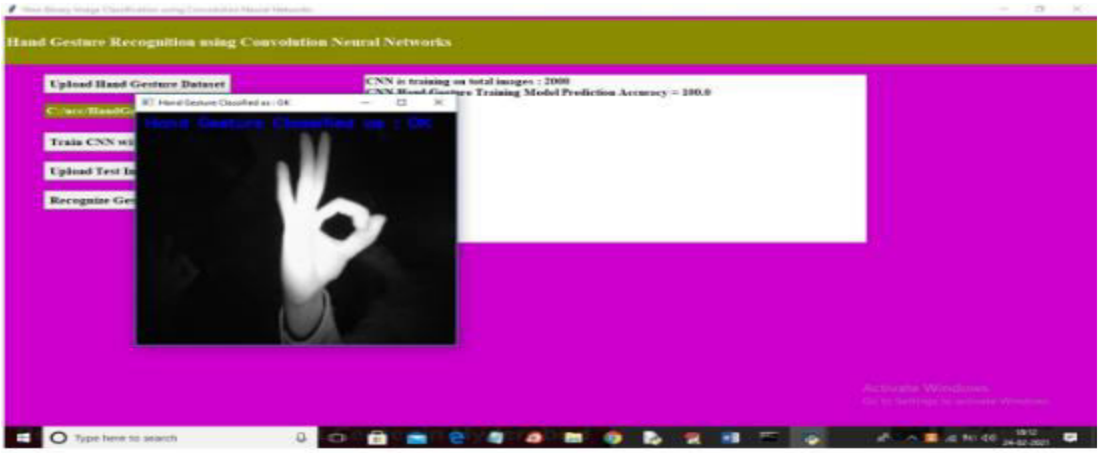
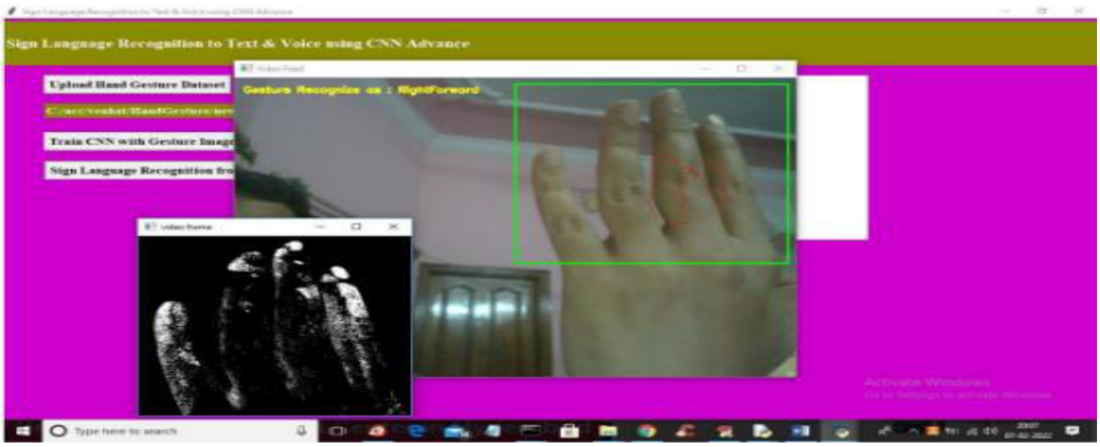


Fig: 6 In above screen gesture recognize as OK and similarly you can upload any image and get result and now click on ‘Recognize Gesture from Video’ button to upload video and get result



8. CONCLUSIONS

Our study on sign language recognition using Convolutional Neural Networks (CNN)highlighted the immense diversity and complexity of sign languages, which differ acrosscountries in terms of gestures, body language, and sentence structures. Capturing precise handmovements and creating a comprehensive dataset posed challenges, as some gestures proveddifficult to reproduce accurately. Consistent hand positions during data collection were critical tomaintaining dataset quality. Furthermore, understanding the unique grammatical rules andcontextual nuances of each sign language was essential to develop a robust recognition system.Despite the challenges, our research underscored the significance of recognizing and preservingthe richness and expressiveness of sign languages, and we remain committed to advancingassistive technologies for improved communication and inclusivity in the future.

9. REFERENCES

[1] https://peda.net/id/08f8c4a8511

[2] Prasad, C. G. V. N., Rao, K. H., Pratima, D., &Alekhya, B. N. (2011). Unsupervised Learning Algorithms to Identify the Dense Cluster in Large Datasets. *International Journal of Computer Science and Telecommunications*, *2*(4), 26-31.

[3] CABRERA, MARIA & BOGADO, JUAN &Fermín, Leonardo &Acuña, Raul &RALEV, DIMITAR. (2012). GLOVE-BASED GESTURE RECOGNITION SYSTEM.10.1142/9789814415958_0095.

[4] He, Siming. (2019). Research of a Sign Language Translation System Based on DeepLearning.392-396. 10.1109/AIAM48774.2019.00083.

[5] International Conference on Trendz in Information Sciences and Computing (TISC). : 30-35, 2012.

[6].Nagamani, K., Prasad, C. G., &Chatrapati, K. S. (2016, February). A novel framework for optimal component based data center architecture. In *2016 International Conference on Information Communication and Embedded Systems (ICICES)* (pp. 1-10).IEEE.

[7] Narasimha Chary, CH.GVN Prasad. Humanoid Ai Robot: A Member of Our Next the Generation Family. Journal of Advancements in Robotics. 2024; 11(1): 20–24p.

[8] Pigou L., Dieleman S., Kindermans PJ., Schrauwen B. (2015) Sign Language RecognitionUsing Convolutional Neural Networks. In: Agapito L., Bronstein M., Rother C. (eds)Computer Vision -

ECCV 2014 Workshops. ECCV 2014. Lecture Notes in Computer Science,vol 8925. Springer, Cham. https://doi.org/10.1007/978-3-319-16178-5_40

[9] NarasimhaCharyCh, CH. GVN Prasad. Human Deep Skin Surface Vibration Frequency Detection from CT and DT Signals Using Genetic Algorithms. International Journal of Algorithms Design and Analysis Review. 2024; 2(1): 1–8p.

[10] Huang, J., Zhou, W., & Li, H. (2015). Sign Language Recognition using 3Dconvolutional neural networks. IEEE International Conference on Multimedia and Expo(ICME) (pp. 1-6). Turin: IEEE.

[11] CHOLLETI, N., & HIRWARKAR, T. (2020). BIOMEDICAL DATA ANALYSIS IN PREDICTING AND IDENTIFICATION CANCER DISEASE USING DUO-MINING. *Advances in Mathematics: Scientific Journal*, *9*, 3487-3495.

[12] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., &Fei-Fei, L. (2009). ImageNet: ALargeScale Hierarchical Image Database. Computer Vision and Pattern Recognition, 2009.CVPR 2009. IEEE Conference on (pp. 248-255). IEEE. Miami, FL, USA .

[13] Soomro, K., Zamir , A. R., & Shah, M. (2012). UCF101: A Dataset of 101 HumanActions

Classes From Videos in The Wild. Computer Vision and Pattern Recognition,arXiv:1212.0402v1, 1-7.

[14] R. Shobarani, R. Sharmila, M. N. Kathiravan, A. A. Pandian, C. Narasimha Chary and K. Vigneshwaran, "Melanoma Malignancy Prognosis Using Deep Transfer Learning," *2023 International Conference on Artificial Intelligence and Applications (ICAIA) Alliance Technology Conference (ATCON-1)*, Bangalore, India, 2023, pp. 1-6, doi: 10.1109/ICAIA57370.2023.10169528

[15] Zhao, Ming & Bu, Jiajun& Chen, C..(2002). Robust background subtraction in HSVcolor space.Proceedings of SPIE MSAV, vol. 1.4861. 10.1117/12.456333.

[16] Chary, C. N., Krishna, A., Abhishek, N., & Singh, R. P. (2018). An Efficient Survey on various Data Mining Classification Algorithms in Bioinformatics. *International Journal of Engineering and Techniques*, *4*.