JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

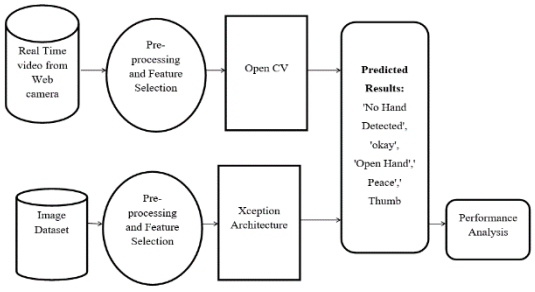
Sign Language Recognition: A Case Study

Harshitha, Ruchi, Sameer Kumar Singh, Saurav Kumar, Tejas V Kangod

ASSISTANT PROFESSOR, STUDENT
RNS INSTITUTE OF TECHNOLOGY

Abstract :

The "Real-Time Hand Gesture Detection for Sign Language Recognition using Python" project aims to develop a system for recognizing sign language gestures in real-time. It utilizes computer vision techniques and machine learning algorithms to detect and classify hand movements, achieving a training accuracy of 99.34% and validation accuracy of 99.00% with the Xception architecture model. The system, implemented in Python with OpenCV, detects gestures like "Ok," "Open Hand," "Peace," "Thumb," and "No Hand Detected." With a user-friendly interface, it facilitates non-sign language users, enhancing inclusivity in classrooms, workplaces, and public spaces.



Introduction:

Language is a language in which people express meaning through language patterns through visual communication. It also includes facial and body language, which are called non-book signs and play an important role in understanding the truth of the signs. People who speak badly have difficulty communicating with others. Because even if family members learn the language, most people in the world cannot learn it.

This is because they do not need the language in their daily work. People from many countries signed it. It includes algorithms inspired by the brain's patterns and functions, including decision-making, called neural networks. Language checkbooks help people with special abilities. Therefore, they are not alone or isolated. Science couldn't have been developed so quickly if the ideas of people like Stephen Hawking had not been heard. is about algorithms that respect decisions resulting from the structure and performance of the brain, called artificial neural networks. Sign language app will help special people communicate with all. Therefore, they will not be separated or excluded.

1.1 Objectives:

The main purpose of creating this application is to bring people closer to each other. Language detection will be performed in real time to ensure rapid response and provide instant feedback for communication and learning. There should be no commerce in the exchange of information. It is important to have a user-friendly graphical interface (GUI) that is intuitive and accessible to users with different levels of expertise as well as users without expertise. Applications should be developed assuming end users have no technical skills. With this in mind, an

application form should be created that allows users to practice and experience the application. The application should be able to convert sign language into different languages. Advanced systems that can accurately create 3D images and provide accurate outputs should be used.

2. Related Work:

There has been a lot of research and development to create the best word analysis application. Tanuj Bohrab Contact should be established immediately. The system has hand detection, skin color segmentation, average blur and visual detection features. These are used for images in the dataset to get better results. It was trained to use large dataset of 40 groups and can predict 17,600 image parameters with 99% accuracy in 14 seconds. Dr. Muthu Mariappan H. and Gomathi V. developed a fast annotation detection system using face, left hand and right hand detection and fuzzy C language algorithm to classify data books listed into categories. The system can reach a 75% accuracy rate. Suharjito implemented language perception with a 3D model for the first time using adaptive learning.
10-word LSA64 public dataset containing 500 videos. The data layer for training is split in the ratio 6:2:2.Public dataset LSA64 was used for 10 vocabularies with 500 videos. For training, the dataset is distributed in 6:2:2 ratio.

Figure 2 Real time detection

For checking the efficiency of the system 320+ videos is used for training, 110+ for

validation and 110+ for testing set. It has very low-validation accuracy. Aditya Das traineda CNN using Inception v3 model.

Figure 3 American Sign Lamguage

Before training, data augmentation is applied on the images to avoid overfitting. This model gives more than 90% accuracy on a dataset consisting of 24 class labels where each class has 100 images. The paper- Developed LSTM model for recognition of extended symbols using leap motion. A new framework for extended SLR using leap motion sensors is described. An improved LSTM architecture is also proposed the recognition of sign words and sentences. Average accuracy of 72.3% havebeen recorded on the signed sentences. For isolated sign-language words an accuracy of 89.5% was recorded. By increasing the training data, the performance increases.

3.1 Challenges in sign-language detection model

1. A single sign language is not following throughout the ASL, Indian Sign Language(ISL) are many other sign languages. It's difficult to guess the meaning.
2. Achieving real-time processing detection is a formidable challenge. The system will have to analyse and precisely tell the meaning ofsign and give output within a few seconds.
3. It is difficult to manually create the training dataset for signature. Recognition. There will be inconsistency in the existing training datasets.
4. The model will have to ignore the background objects which can include lighting, environment and people. The system will have to filter these out and give accurate output.

3.2 Dataset Generation

It's imperative for development of consistent Dataset:

Kaggle Dataset Link:

https://www.kaggle.com/datasets/jayaprakashpondy/hand-gesture-dataset

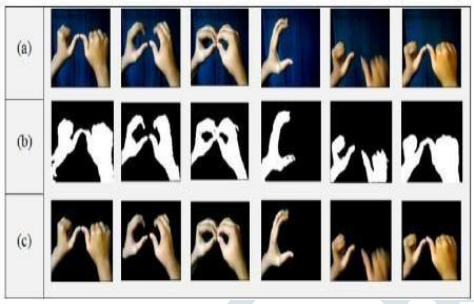


Figure 4 Sign Language Dataset

Finally, apply the gaussian blur filter to the dataset which helps to extract various features of the selected image.

3.3 Advantages

1. Basic Communication: The system allows for basic communication for people that don't know ASL to make iteasier for visitors. Personalization: Custom signs can be tailored to an individual's unique communication needs, preferences, and even regional variations, enhancing the user experience.
2. Real-Time Recognition: The real-time recognition of custom signs from moving frames input provides a natural and efficient means of communication, enabling faster and more fluid interactions.
3. Compact and Accessible: Accessible through mobile and web applications, the system offers a convenient and portable means of communication, allowing users to communicate fromvarious devices.
4. Simplified Model Training: Training models on a limited custom gesture dataset is simpler and requires less datacompared to large, generic sign language datasets, reducing training time and computing resources.
5. Focused Vocabulary: The system's focused vocabulary enhances recognition accuracy, as it is specifically optimized for practical and commonly used words and phrases.

3.4 Disadvantages

1) Limited to Simple Phrases: Thesystem's primary limitation is that it's designed for simple phrases and may not support the complexities of a full sign language, limiting its application scope.
2) Custom Gesture Collection: The process of collecting personalized custom gesture data can be time- consuming and may require user involvement, potentially posing achallenge.
3) Restricted Vocabulary: The system's vocabulary, while tailored to user needs, may not cover all communication requirements,potentially leading to gaps inexpression.
4) Confusion Risk: With a small custom gesture vocabulary, there's a higher riskof confusion, especially if gestures are similar for different words or phrases.
5) Not Accessible to Non-Signers: This system primarily benefits some knowledgeable sign language, making it less accessible to individuals who only know spoken languages.

4 Methodology

4.1 Importing the necessary libraries:

This module involves importing the necessary libraries such as NumPy, TensorFlow to implement the model. We will be using Python language for this. First, we will import the necessary libraries such as keras for building the main model, sklearn for splitting the training and test data, PIL for converting the images into array of numbers and other

4.2 Importing the necessary libraries:

This module involves importing the necessary libraries such as NumPy, TensorFlow to implement the model. We will be using Python language for this. First we will import the necessary libraries such as keras for building the main model, sklearn for splitting the training and test data, PIL for converting the images into array of numbers and other libraries such as pandas, NumPy, matplotlib and TensorFlow.

4.3 Retrieving the images:

We will retrieve the images and their labels. Then resize the images to (224,224) as all images should have same size for recognition. Then convert the images into numpy array.

4.4 Splitting the dataset:

This module involves splitting the preprocessed 80% train data and 20% test data.

4.5 Building the model:

Xception model:

•Inspired by Google's Inception model
•Xception is based on an 'extreme' interpretation of the Inception model
•Simple and modular architecture

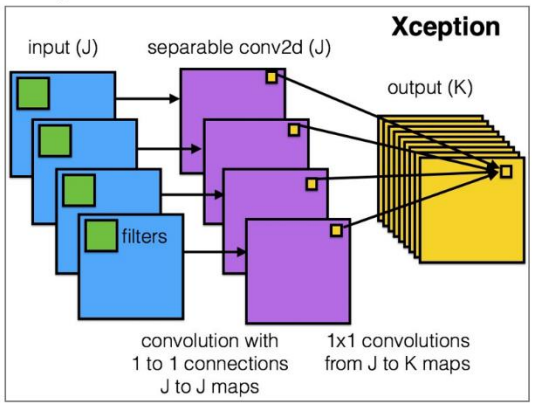
Depthwise Separable Convolution:

Regular Convolutions:

•Look at both channel & spatial correlations simultaneously

Depthwise separable convolution:

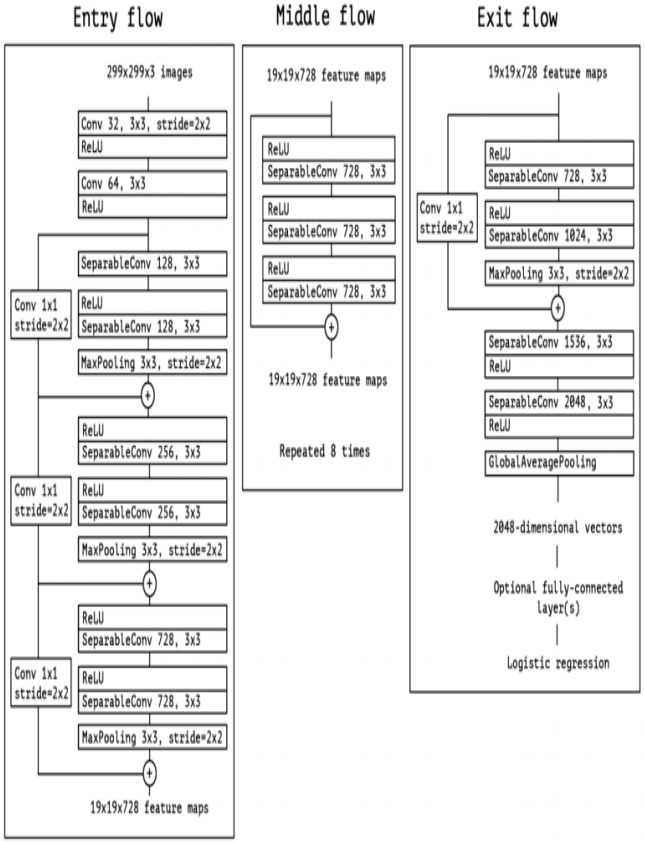
•Look at channel & spatial correlations independently in successive steps
•Spatial convolution: 3x3 convolutions for each channel
•Depthwise convolution: 1x1convolutions on concatenated channels



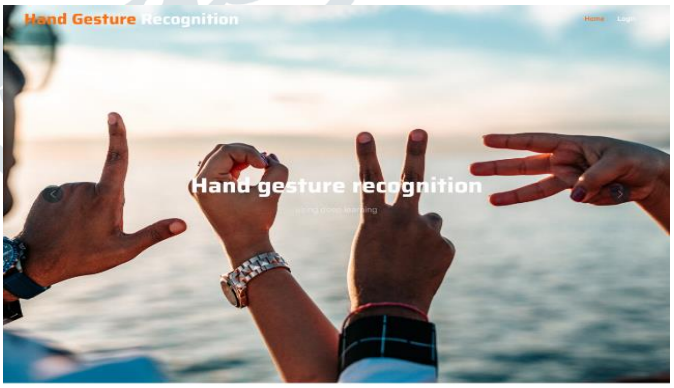
Example: take 3x3 convolutional layer on 16 input channels and 32 output channels.

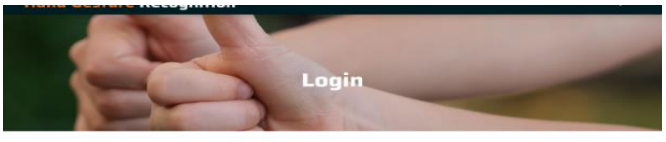
•regular convolution: 16x32x3x3 = 4608 parameters
•depthwise separable convolution: (spatial conv + depthwise conv) = (16x3x3 + 16x32x1x1) = 656 parameters
•greatly reduced parameter count
•more efficient complexity.

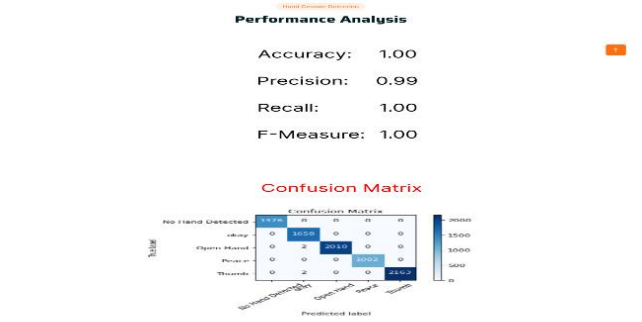
4.6 Xception architecture

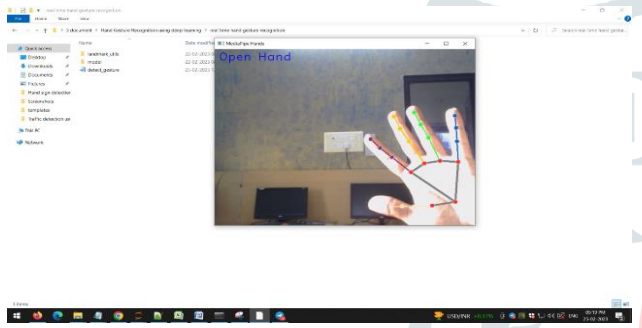


5. RESULTS









Conclusion:

This article describes the work done to date in research field of language analysis applications. Certain objectives and criteria have to be met for the application to be efficient. Explaining the functioning of existing systems and laying the foundation for the development of new systems. All the features of the sign-language detection system are described which would be implemented in the new system. The project's significance lies in its potential to foster inclusivity, understanding, and accessibility,creating a universe of effective communication.

References

1. Fig 3.1 by T. Bohra, S. Sompura, K. Parekh and P. Raut, "Real-Time Speech and Hearing Impaired Using Computer Vision and Deep Learning," 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2019, pp. 734-739,doi:10.1109/ICSSIT46314.201 9.8987908.

2. Fig 2 by H. Muthu Mariappan and V. Gomathi, "Real-Time Recognition of Indian Sign Language," 2019 in Data Science (ICCIDS), Chennai,India,2019,pp.1-6,doi: 10.1109/ICCIDS.2019.8862125.

3.Suharjito, Suharjito & Gunawan, Herman & Thiracitta, Narada & Nugroho, Ariadi. (2018). CNN Model. 1-5. 10.1109/INAPR.2018.8627014.

4.Fig 3 by Singha, Joyeeta & Das, Karen. (2013), "Recognition of Indian Sign Language in Live Video," International Journal of Computer Applications. 70. 10.5120/12174-7306.

5. Koller, Oscar & Zargaran, Sepehr & Ney, Hermann & Bowden,Richard. (2016). Deep Sign: Hybrid CNN- HMM for CSL Recognition. 10.5244/C.30.136

6. A. Mittal, P.Kumar, P.R Roy, R Balasubramanian, B.B. Chaudhuri, "A Modified-LSTM Model for Continuous Sign Language Recognition using Leap motion",2019. https://ieeexplore.ieee.org/abstract/

7. A. Mittal, P.Kumar, P.R Roy, R Balasubramanian, B.B. Chaudhuri, "A Modified-LSTM Model for Continuous Sign Language Recognition using Leapmotion",2019. https://ieeexplore.ieee.org/abstract/document/8684245

8. Fig 2 by Elizabth Cody, "Deaftravel", http://www.deaftravel.co.uk/signs.php?id=27