IDEATION PHASE

LITERATURE SURVEY

Date	13/10/2022
Team ID	PNT2022TMID22166
Project Name	Project - Real-Time Communication System Powered by AI for Specially Abled

Paper Title I : **Sign Language Recognition, Generation, and Modelling: A Research Effort with Applications in Deaf Communication**

Author Names : Eleni Efthimiou,(Institute for Language and Speech Processing) , Evita Fotinea,
Christian Vogler,
Thomas Hanke,(University of Hamburg)

Abstract : The avatar that signs Natural language knowledge, which is encoded in a lexicon of annotated signs, and a set of rules that permit the structuring of basic grammatical events, making significant use of feature attributes and structuring options, are both crucial to sign language synthesis. This is required to ensure that the signing is conducted in a linguistically appropriate manner. The Dicta-Sign project's annotated parallel corpora serve as the foundation for these rules, which cover both manual and non-manual aspects and the significance of where signs are placed in space. The conversion of a linguistic description of the signed utterance into a smooth animation using inverse kinematics, with precise positioning of the hands in contact with the body, is one of the trickiest difficulties in sign synthesisand creating adequate visual stress, among other realistic prosodic qualities.

To this goal, the sign language corpus is explained in the following section and includes prosodic information in addition to phonetic and grammatical information. This provides for a vastly higher level of

realism in the animations, especially when combined with the features gained from the visual tracking and recognition component..

Paper Title II : Sign Language Recognition using Deep Learning

Author Names : Dhruv Sood

Abstract : We'll employ a method called Transfer Learning along with Data Augmentation to create a deep learning model for the ASL dataset. A model created for one task is used as the basis for a model created for another task in a process known as transfer learning. Pre-trained models are a common approach in deep learning for computer vision and natural language processing applications due to the enormous computing and time resources required to construct neural network models for these problems as well as the significant skill gains they provide on related problems.

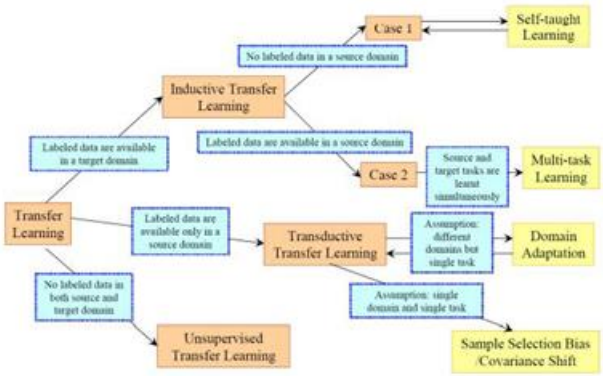

Fig 2. Brief on Transfer learning

After the model has been trained, it is integrated into the application. OpenCV is used to grab frames from a video feed. The application provides an area (within the green rectangle) where the indicators for detection or recognition can be shown. The signs are then photographed in frames, which are then processed for the model and supplied to the model. Based on the sign created, the model predicts the sign captured. If the model predicts a sign with a confidence of greater than 20%, the prediction is presented to the user.

Paper Title III : Towards Computer Assisted International Sign Language Recognition System

Author Names : Mikhaylyna Melnyk, Vira Shadrova,
Borys Karwatsky

Abstract : Sign languages pose the challenge that there is unfortunately no international sign language. All sign languages suffer from lack of rules and regulations.Each country needs to create its own method for recognising sign language. Despite the fact that facial expressions and gestures serve as the fundamental building blocks of all sign languages, designing a global recognition framework is challenging due to nation-specific needs. The development of computer-assisted sign language recognition systems is the main topic of this presentation. We specifically talk about whether the ongoing study might lead to the creation of an international sign language. We group the existing sign language recognition systems into categories and provide a summary of them. We also provide a list of openly accessible databases that can be used in the development of sign language recognition systems.

Paper Title IV : Sign Language Translator and Gesture Recognition

Author Names : Mohammed Elmahgiubi(University of Guelph)
Mohamed Ennajar,Nabil Drawil(University of Tripoli) ,
Mohamed Samir Elbun

Abstract : Individuals who are deaf or mute struggle to communicate with others. Such people struggle to communicate their ideas because not everyone can understand sign language. In this work, a Data Acquisition and Control (DAC) system that converts sign language into text that is readable by everyone is developed. Gesture Recognition and Sign Language Translator is the name of this system. We created a smart glove that records hand motions and translates them into words that can be read. This text

can be wirelessly transmitted to a smart phone or shown on a built-in LCD screen. The experimental findings demonstrate that a combination of low-cost sensors can record motions by measuring the positions and the orientation of the fingers. The current version of the system is able to interpret 20 out of 26 letters with a recognition accuracy of 96%.

Paper Title V : Sign Language Recognition Using Deep Learning on Custom Processed Static Gesture Images

Author Names : Aditya Dasl, Shantanu Gawde

Abstract : Despite a wide social group that could profit from it, the idea of sign language recognition by technology is underutilised. There aren't many technological advancements that assist this social group communicate with the rest of the world. One of the main tools for enabling sign language users to communicate with the rest of society is understanding sign language. Computers can recognise sign language with the use of image categorization and machine learning, which can then be translated by humans.

This study uses convolutional neural networks to identify sign language motions. The static sign language gestures were photographed using an RGB camera and comprise the image dataset used. The photos underwent preprocessing before being used as the input that had been cleaned. Inception v3 convolutional neural network model was used to retrain and test this dataset of sign language motions, and the findings are presented in this study.

Multiple convolution filter inputs are processed on a single input in the model. The resulting validation accuracy was greater than 90%. The numerous attempts at sign language detection using machine learning and image depth data are also reviewed in this work. It assesses the different difficulties involved in solving the issue at hand and also describes the problem's potential future.

Paper Title VI : Deepsign: Sign Language Detection and Recognition Using Deep Learning

Author Names : Deep Kothadiya, Chintan Bhatt, Krenil Sapariya, Kevin Patel ,Juan M. Corchado

Abstract : An RNN structure (containing forgetting units like long short-term memory (LSTM) and gated recurrent units (GRU)) is typically offered to enable the storing units in proposals to overcome the gradient vanishing problem. The best time delay and when to forget certain knowledge are determined using such a system. Given the known characteristics of GRU and LSTM, the methodology put forward in this study should be able to recognise and identify sign language motions from a video source and produce the English word that goes with them. This technique uses ISL sign language gesture sequences from a video source as an input. This system's main goal is to recognise the words from The first challenge is to break the video file containing the sequence of ISL motions for distinct words into separate sub-section films comprising different words. This is accomplished by determining the beginning and conclusion of each unique gesture.

The subsequent step is to split the resulting subsection videos into frames after the videos have been divided. By using sampling techniques on the video sequence, the frames are produced. The proposed architecture made use of InceptionResNetV2 to extract the features from the frames. InceptionResNetV2 outperforms its predecessor in terms of performance. The inceptionResNetV2 model receives the resized video frames and uses mobile net pre-trained weights to extract the features from the motions. The inceptionResNetV2 model generates an array that includes feature vector.