ABSTRACT

In last three decades, progress in Artificial intelligence and Machine learning has made tremendous development in the field of Computer Vision, Natural language processing (NLP), Automatic Speech Recognition (ASR), and many more. ASR, also known as speech to text or Automatic Lip Reading (ALR), is a visual speech recognition technology, which interprets the speech by observing the lip, tongue, and teeth movements of the speaker without speech signals. Lip movements are, especially, used by partially deaf people to perceive the meaning of speech. Generally, partially deaf or hearing of hard people communicate with each other using sign language. Sign language is manual communication in which gestures or symbols are organized in linguistic way. Just like other languages, sign language also changes according to different countries and regions. However, it becomes difficult for deaf people to communicate with people who do not know sign language. Through technological innovations and ALR techniques, this deficit can be filled. Using Lip reading, in their childhood only, deaf children can be taught their mother tongue easily. Also, the lip expressions can be converted into sentences by a Lip reading system which can be displayed on a deaf person's mobile screen which can be easily read. According to research, skills acquired in childhood, help a child to learn languages in a better way. Most of the Lip-reading work is carried out for the English language and other foreign languages. This thesis work is to create a dataset for Gujarati Alphabets, as no work has been carried out, from scratch, design and develop an algorithm for Lip detection and extraction and classification using CNN-LSTM model. In this work, a dataset *GJVarna* has been created with 3 shots of 25 speakers, uttering 34 consonants of Gujarati Language. A lip detection and extraction algorithm with extra frame removal, ViLiDEX, has also been developed. The CNN-LSTM based MobileNet model has been used for training and testing, recognition and classification of alphabets of Guttural, Palatal, Retroflex, Dental and Labial types.