Summaty Title

Techer name

Nilesh navalkar

60003200076

**Title**

The purpose of this study was to study the uses and ways of implementing lip reading using deep learning. It was a survey research study that looked at the features of different approaches for lip reading, and it comes with its own advantages and disadvantages. The end to end conversion of lip moments to the words is possible because of availability of huge data and development of different deep learning methods such as Convolution Neural Network and Recurrent Neural Networks. The use of Deep Learning in lip reading is a recent concept and solves upcoming challenges in real-world such as Virtual Reality system, assisted driving systems, sign language recognition, movement recognition, improving hearing aid via Google lens. Various other approaches along with different datasets are explained in the paper.

Lip perusing, otherwise called discourse perusing, visual discourse acknowledgment (VSR), is a strategy of knowing the discourse with outwardly translating developments of the facial features like lips, face, and tongue when ordinary sound isn't accessible. Researchers are starting to focus their attention-based mechanism with convolution neural networks (CNN) which is based on specific regions such as lip reading. Also, classification methods and target detection has drastically improved which supports the deep learning methods and making it more efficient.

**Lip reading based on Convolutional Neural Network:**

The proposed methods in the paper uses Convolutional Neural Network to extract visual features, the model has been trained by using different images along with labels. one of the most widely used neural networks from image classification and detection for extracting visual features from image.

The key advantages for the proposed method were easy implementation as separate model of lip shape or hand-labelled data are not required also, shift- and rotation- resistant image recognition is performed by CNN guarantees. The proposed mechanism is tested on 40 kinds of phoneme recognition evaluation experiments and attains a 58% recognition rate.

Data set used for the experiment was a Japanese audiovisual dataset [18], the dataset contains of 300 words with 6 different male speakers [19]

What the study found

My opinion on study

references