Deep Learning In Silent Sound Technology -An Efficient Application

Sweta Subhash Rajbhar[1], Sonali Vitthal More[2]
[1, 2] ASM's Institute of Management and Computer Studies, Thane, India

Abstract- *Artificial Intelligence, a broad field of Computer Science, aims to reproduce or simulate human Intelligence in machines. Machine Learning, a branch of Artificial Intelligence deals with machine intelligence i.e. enabling the machines to learn and perform the tasks from experience rather than programming them specifically about the task. Deep Learning is now considered as the most proficient subset of Machine Learning that learns by mimicking the inner workings of human brain to process data and make decisions based upon that data. Basically, it is a sort of Machine Learning that brings into play an Artificial Neural Networks. Rank Brain Algorithm is an example of Deep Neural Network, which is one amongst the factors within the Google Search Algorithm. Another efficient application of Deep Learning that has been emerging in recent years is the Silent Sound Technology. Silent Sound Technology (talking without talking), is a non-speech interaction embedded engine that captures a range of lip movement and convert it into computer generated sound that could be transmitted via various means. It will be useful for those who just want to ignore all other noise and turn it into a positive sound at the receiver and make a quite call. Within this article, I have also reviewed the benefits of having Silent Sound Technology, its methodologies and its future scope towards driving digitalization to a new path.*

Keywords- Silent sound Technology, electromyography, deep learning, muscular movement, ultrasound transducer

I. INTRODUCTION

The next computing revolution will be driven by artificial intelligence and machine learning. These technologies rely on the capacity to discern patterns and then forecast future events based on historical data. This explains why Amazon makes suggestions while you shop online or how Netflix knows you like awful 80s movies. Although AI-based robots are frequently referred to as "smart," the majority of these systems do not learn on their own; human programming is required. Data scientists pick the factors to be used in predictive analytics and prepare the inputs. Deep learning, on the other hand, can automate this process. Deep learning is still in its infancy, yet it has the potential to alter civilization in the next decades. Besides self-driving cars, which are tested worldwide, natural language processing in case of Siri, Alexa, Cortana, deep learning has now enriched its roots in the world of Speech. And one such application is Silent Sound Technology.

Nowadays, there is so much noise and disruption that when we converse on a cell phone in a crowd, we are essentially yelling rather than communicating. However, “Silent Sound Technology” has been introduced that would actually put an end to noise pollution. When this technology is employed, it monitors every movement of the lips and inwardly turns the electrical pulses into sound signals, which it then sends out, rejecting all other distractions. Silent sound technology is ideal for folks who have lost their voices but yet want to communicate on their phones. According to research, articulation muscles are always engaged whether or not air goes through them (the vocal cords vibrate if you tension these muscles slightly and release a little air). When you talk or hum, place your finger over your throat to demonstrate phonation, which is also known as fundamental frequency of voice). The electrical impulse is then converted into a sound signal, which is then delivered to the listener, who hears the speech.

This technological innovation will come in handy anytime a person loses his or her voice, when speaking to make silent calls without disturbing others, or even when we want to tell a trusted friend or relative our PIN information without having another person secretly listen to it. The listener can hear a distinct voice on the other end.

II. LITERATURE REVIEW

Detection every lip movement and transform them into sounds, which could be helpful for people who lose their voices to speak, and allow them to make silent calls without disturbing others, was illustrated in 2010 CeBIT's "future park", a concept (Schultz, 2010) "Silent Sound" Technology. Instead of creating sounds, your handset would decode your lip motions by sensing muscle activity, then turn this into speech that the person on the other end of the line could hear. In a nutshell, it reads your lips. As a result, people can converse on their phones in a crowded area without being disturbed.

"At the moment, we're employing skin-attached electrodes. Such electrodes may, for example, be included into mobile phones in the coming years," remarked (Schultz, 2010) from the KIT. The technology has a wide range of uses, including assisting persons who have lost their voices due to illness or injury. Anyone can also use innovation to become a fluent speaker in an instant. "Local speakers can discreetly express an explanation in their own tongue, and beneficiaries will hear the deciphered statement in their native tongue," Schultz explained, adding, "It appears as if the native speaker was speaking in a foreign language."

III. METHODS

The techniques which have been included for use in Silent Sound Technology are:

1. Electromyography (EMG)
2. Image processing.

1. ELECTROMAYOGRAPHY

Electromyography (EMG) is a technique for studying muscle contraction by monitoring electrical activity produced by the muscle. This electrical activity, which appears as a signal, is caused by neuromuscular activation, which is linked to muscle contraction. This mechanism is used in silent sound technology SST, which detects muscle activity when we talk quietly and converts it to electrical signals.

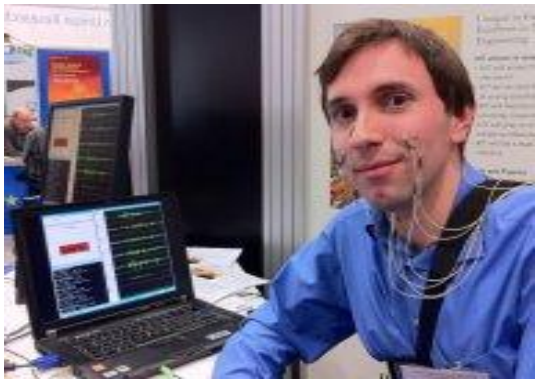

Fig 1. Electromyography

Wire electrodes are introduced into muscle tissue through the skin (fig 1). The electrical signal produced by the facial muscle is recorded and compared to the recorded signal of the spoken words, both of which are identical. It was proven that the electrical signal produced corresponded to the spoken words, so the electrical signal could now be transmitted and converted to the same sound signal at the listener's end.

This whole process is carried out with an instrument called electromyograph and the signals generated are called as electromyogram. Thus, in this manner speech can be communicated without the use of sound.

2. IMAGE PROCESSING

Image processing is manipulation of image in order to improve it or extract meaningful information from it. It is a type of signal processing in which an image is used as input and an image or imagecharacteristics/features are used as output.

The large proportion of image-processing methodologies involve treating the image as a two-dimensional signal and applying standard signal-processing techniques to it. The most basic type of digital image processing converts a digital data tape into a film image with the fewest corrections and computations. Large mainframe computers are then used for sophisticated interactive data manipulation.

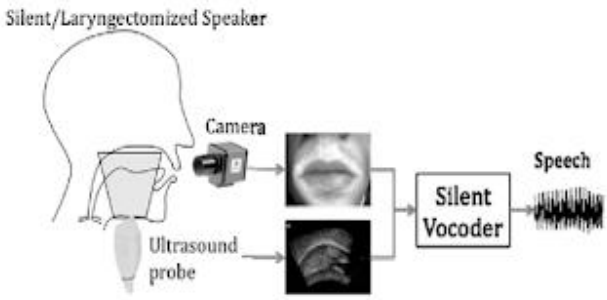

Fig 2. Image Processing

Image Processing techniques are employed in order to analyze remotely sensed data. Image Processing methods include:

1. Analog image processing
2. Digital image processing

1. ANALOG IMAGE PROCESSING

The analog image processing technique is used to process hard copies of data such as photographic images or printed notes.

Analog signal is a real-world processing technique however, it is not flexible for good quality images.

2. DIGITAL IMAGE PROCESSING

Digital Image Processing refers to a set of techniques used by computers to manipulate digital images. However, it

has some flaws. To overcome flaws and deficiencies and ensure the data's originality, it must go through several processing steps. Three general steps are involved in digital image processing:

1. Image Acquisition

Image Acquisition involves retrieving an image from a specific source. Image processing such as scaling is done on the image extracted.

2. Image Enhancement

Image enhancement involves sharpening of images, adjusting the brightness and contrast, removal of noise in order to obtain results that are quite suitable for display.

3. Image Restoration

Image Restoration involves improvement in the enhancement and appearance of image that might have been degraded due to additive noise.

4. Color Image Processing

This step involves processing of colored images.

5. Wavelets and multiresolution processing

Wavelets provide the foundation for representing images with varying degrees of resolution. For data compression, images are subdivided into smaller regions.

6. Compression

Compression techniques are used in order to reduce the bandwidth required to transmit the image.

7. Morphological Processing

Image components useful in representation and description of shape are done in this step.

8. Segmentation

Segmentation involves dividing the image into multiple segments in order to locate the objects.

9. Representation and Description

Computer processing requires image to be converted into suitable form. Representation deals with converting the data to suitable form.

Description involves obtaining the attributes of image to obtain qualitative information.

10. Recognition

The last step is recognition that assigns a label tothe object based upon the description obtained above.

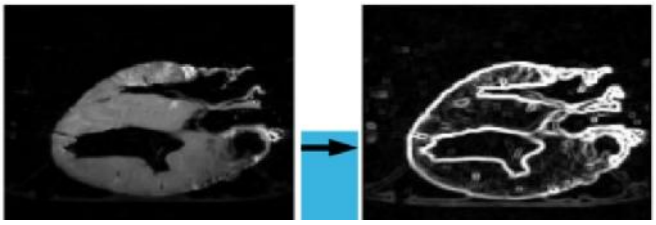

Fig 3. Digital Image Processing

IV. ALGORITHM USED

Convolutional Neural Network

CNN is a deep learning model used to process data with a grid pattern, such as images especially for identification of two-dimensional image information.

The CNN algorithm is comprised of two major processes: convolution and sampling. Convolution process: using a trainable filter Fx, deconvolution of the input image (the first stage is the input image, the input of the after convolution is the feature image of each layer, namely Feature Map), and then adding a bias bx, we can obtain convolution layer Cx.A sampling procedure: n pixels from each neighbourhood are pooled to form a pixel, which is then weighted by scalar weighting Wx + 1, bias bx + 1, and activated to produce a narrow n times feature map Sx + 1.

CNN's key technology is the local receptive field, weight sharing, and sub sampling by time or space to extract features and reduce the size of the training parameters. The advantage of the CNN algorithm is that it avoids explicit feature extraction and instead learns implicitly from training data.

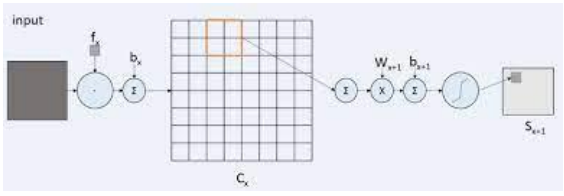

Fig 3. Main Process of CNN

Because the same neuron weights are used on the surface of the feature mapping, the network can learn in parallel, reducing the network's complexity. Adopting a time or space-based subsampling structure can achievesome degree of robustness, scale, and deformation displacement. It has distinct advantages in speech recognition and image processing when input information and networktopology are well matched.

V. ADVANTAGES OF USINGSILENT SOUND TECHNOLOFY

1. Share confidential Information

Confidentiality refers to an individual's commitment to refrain from disclosing confidential information to others unless the other party expressly consents. Loss of confidentiality affects the integrity of messages and could even lead to loss of billions of dollars. Thus, Silent Sound Technology could be used to share confidential information such as PIN Number on calls in public place.

2. Noise Cancellation

Talking on cell phone in a crowd makes one to yell rather than talk because of lots of disturbances and distractions around us. It becomes very difficult to convey the message thus resulting in wasting energy and increase in noise pollution as well. Silent Sound Technology will put an end to noise pollution to some extent and acts as perfect solution for those who wish to talk even in crowded place.

3. Translation to any Language

Due to the universal nature of electrical signals, they can be interpreted into any language. It can be interpreted by native speakers before being sent to the other side.Consequently, it could be transformed into any preferred language, presently German, English, and French.

And thereby, amongst the most recent trends in the field of informatics is Silent Sound Technology, which implements ‘Speaking without Actually speaking.’ It will become one of the most innovative and effective technologies, and will be used in everyday life in the near future.

VI. LIMITATIONS

1. Differentiation between emotions

It is impossible to distinguish between people and emotions. As a result, people will always have the impression that they could be conversing with a robot.

2. Difficulty in Recognition

Since, Silent Sound Technology translates the spoken words into electrical impulse signals which are transmitted to receiver at other end as a native voice thus, recognizing who is talking becomes tricky from a security standpoint.

3. Impractical usability

This device currently requires nine leads to be attached to the face before it can be used, making it difficult to use in practice.

4. Not Possible for Language that holds different meaning for different tones

Translation into the majority of languages is possible, but in languages such as Chinese, different tones have different meanings, despite facial movements being the same. As a result, this technology is difficult to apply in such circumstances.

VII. FUTURE SCOPE

Silent sound technology advancements pave the way for a bright future for speech recognition technology, from simple voice commands to transcripts dictated over the phone; all of this is reasonably possible in noisy public places. Rather than having electrodes all over the face, these electrodes will be assimilated into smart phones.Nanotechnology will be a significant step toward making the device useful.

With millions of phones in use, there is a great opportunity to increase earnings by saving 'lost calls' – phone calls that go unanswered or uninitiated because the user is unable to speak – not just in business meetings, but in everyday situations. According to research, these "lost calls" are worth $20 billion worldwide each year. So, this is the potential earning to the cellular operator underneath the table.

Silent Sound technology is a cutting-edge innovation. Scientists are working on it and attempting to overcome technical difficulties. It only supports a few languages, including English, German, and French. As a result, it is hoped that it will be improved and become the technology of the future.

As a result, we conclude that Silent Sound Technology, one of the most recent trends in information technology.

VIII. CONCLUSION

Silent Sound technology detects and translates every motion of the lip into sound, allowing the ones who have lost their voices to speak, including other folks to make silent calls without disturbing others. The device, according to the engineers, has a 99 percent efficiency rate. Instead of just making sounds, the phone would decipher mouth movements by evaluating muscle activity and then translate this to speech so that the fellow on the receiving end of the phone could understand. In other words, it reads lips. It would be a cutting-edge and practical new tech will be used in everyday life in the not-too-distant future.

IX. ACKNOWLEDGEMENT

This research was supported by Audyogik Shikshan Mandal's Institute of Management and Computer Studies.

REFERENCES

[1] https://www.digitalinformationworld.com/2020/12/this-new-ai-model-changes-silence-into.html
[2] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7922488/
[3] R. Sorensen, Hearing Silence: The Perception and Introspection of Absences. Website Available: https://pages.wustl.edu/files/pages/imce/sorense n/hearing_silence_chap_nudds.pdf
[4] https://www.geeksforgeeks.org/introduction-deep-learning/
[5] http://www.techpark.net/2010/03/04/silent-soundtechnology-an-end-to-noisy-communications/
[6] http://www.scribd.com/doc/114337130/silent -sound - technology-report/
[7] http://www.dellchallenge.org/projects/silent-soundtechnology
[8] http://en.wikipedia.orrg/wiki/Digital_signal_processing
[9] http://en.m.wikibooks.org/wiki/speechlanguage_patholog y/stuttering/print_version
[10] http://www.techpark.net/2010/03/04/silent-soundtechnology-an-end-to-noisy-communications/
[11] https://wiki.pathmind.com/neural-network