ISSN: 0975-766X
CODEN: IJPTFI

Available Online through ***Review Article***

www.ijptonline.com

HUMAN EMOTION DETECTION USING DEEP LEARNING NEURAL NETWORK-A REVIEW

[1]Swagata Sarkar*, [2]H Ranganathan

[1]Research Scholar, Department of ECE, Sathyabama University, Chennai, India.

[2]Professor, Department of ECE, Gojan School of Business and Technology, Chennai, India.

Email: swagata.b.sarkar@gmail.com

Received on 02-05-2017 *Accepted on: 22-06-2017*

Abstract:

Deep learning is a new era of machine learning which has been introduced with the objective of moving machine learning closure to artificial intelligence. It is a powerful set of techniques for learning in neural networks. In this paper a review has been done on deep learning neural network for human emotion analysis. Human emotions are majorly classified into six categories.

They are happy, fear, anger, sad, neutral and surprise. But it is very difficult to identify the emotions efficiently using conventional neural network. In conventional neural network only single layer is used as a hidden layer in between input and output layer.

But for accurate and precise result network must work synonymous to human brain, which can be possible by using deep learning neural network. Human emotion can be identified from different mode of signals. In this paper speech and facial expression are reviewed to identify emotions using deep learning neural network.

Key Words:

Deep Learning, Human Emotion, Neural Network

1. Introduction

Deep learning means using a neural network with several layers of nodes between input and output. Like our brain, multiple number of layers between input and output layer can do feature identification and processing in a series of stages. In the following figure we can get some idea about layered architecture of human brain similar to deep learning neural network.

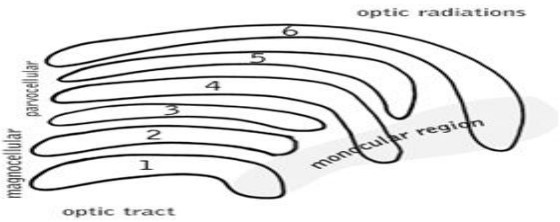

Fig.1. Layered architecture of human brain.

It is a concept of multilayer network. But main differences between conventional neural networks are the number of hidden layers and training algorithms. In conventional neural network single hidden layer is used along with single input and output layers. In figure 2 conventional layered architecture of neural network is depicted.

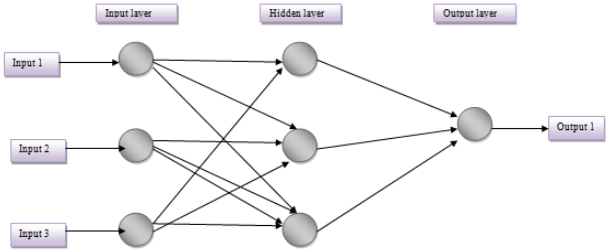

Fig.2. Conventional Neural Network.

In deep learning neural network we have several layers in between input layer and output layer. The number of hidden layer can be changed according to the complexity of the problem. It is observed that deep learning neural network can give better performance in many aspects such as emotional intelligence, emotion detection and many other abstract fields. In conventional network learning algorithms are good for adjusting weights for single hidden layer. But for deep learning neural network new algorithms are used for learning weights for networks with more hidden layers. In figure 3, architecture of deep learning neural network is shown.

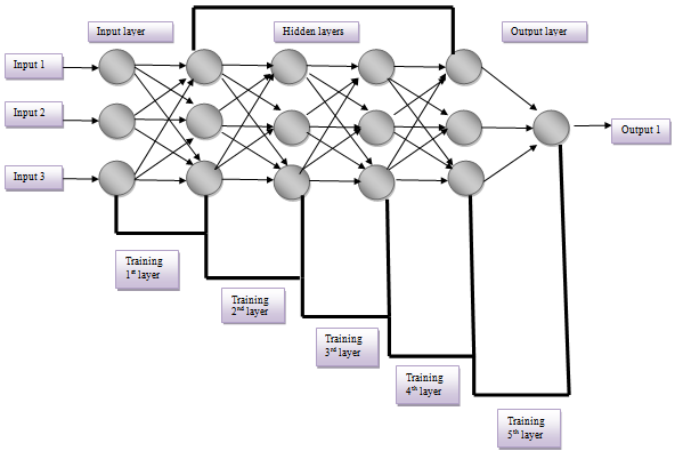

Fig.3. Deep Learning Neural Network.

To detect multiple emotions in an effective way multiple layers between input and output make sense. Our brain also works in same way with multiple numbers of layers shown above. So, deep learning neural network can best suited for complex problems. Much layer neural network architecture should be capable of learning the true underlying features and feature logic and therefore generalize very well.

The training is totally different for deep learning neural network. In the beginning input and first hidden layer will be trained. Next 1st and 2nd hidden layer will be trained. After that training will be done for 2nd and 3rd hidden layer. The last hidden layer and output layer will be trained finally. Therefore it is forced to learn good features that describe, what comes from the previous layer. Each intermediate layers are trained by auto encoders.

There are five major deep learning architectures are there. They are

1. Deep Neural Network
2. Convolution Deep Neural Network
3. Deep Belief Network
4. Convolution Deep Belief Network
5. Deep Q Network

In deep neural network, Feed forward and recurrent network are used. It can be trained by Back Propagation algorithm. Weight updating can be done by stochastic gradient descent. The mathematical expression is

$$Wij\ (t+1) = Wij\ (t) + n\ (\delta c/\delta Wij) + \xi\ (t) \qquad (1)$$

Where n= learning rate

c= cost function

ξ (t)= stochastic term

The main drawbacks for deep neural networks are over fitting and computation time.

The main advantage of convolution neural network is that it is very useful for 2D structure of input data. Compare to other conventional algorithm it is superior for speech and image processing. Back propagation algorithm with modifications can be used for convolution neural network.

Deep belief neural network is probabilistic and generative model made up of multiple layers of hidden units. A DBN can be used to generatively pre-train a DNN by using the learned DBN weights as the initial DNN weights. Then can be

trained by BPN. It can be efficiently trained in unsuppressed layer-by-layer manner where the layers are typically made of restricted Boltzmann machines. Convolution deep belief architecture is the combination of convolution deep learning and deep belief network. It can efficiently handle 2D images or dataset by pre-training process of deep belief network. It can be used for complicated data set.

Deep Q network is a combination of deep convolution neural network with Q learning, a form of reinforcement learning.

2. Review on Existing Network

In reference [1] speech processing is done by deep learning neural network. Feed forward and recurrent neural architectures are used for IEMOCP database. Reference [2] deals with Multiple Fusin Layer based Ensemble Classifier of Stacked Encoder for Stacked Encoder analysis. Three hidden layers are used. The results are compared with best classifier and got mean of classification and F-score improves by 5.26%. In reference [3] facial expression recognition is done by eight directional strengths for facial image. Modified local directional pattern with generalized discriminator analysis is used. Video data is used for analysis. It gives 96.25% efficiency compare to others as 91.67%. In reference [4] multimodal information convolution neural network is used to learn the location of emotional expression in a cluttered scene.

3. Discussions

It is seen from the above review is that the deep learning neural network can be use efficiently for human emotion analysis. Emotion is a very complicated human behavior. Emotional intelligence has a very big impact on human growth. From research, it is seen that high intelligent quotient people fail many cases due to lack of emotional intelligence. Emotions can be compound and complex rather than single and distinct. Deep neural network can be best fit for emotion analysis from various modes. In the above review deep learning network is used for speech and facial emotional expression recognition. They have used limited number of hidden layers. If the number of hidden layers increased, more accurate and precise results can be obtained.

4. Conclusions

In this paper it is clear that any work with speech and image domain, deep belief network is the best choice. Convolution deep belief network is best fit for 2D image analysis. Since numbers of hidden layers are higher and training method is more advanced, it can be successfully implemented in any complex problem**.**

5. Acknowledgement

The author likes to thank the university to provide excellent research facility.

6. References

1. Evaluating deep learning architectures for Speech Emotion Recognition Original Research Article, Haytham M. Fayek, Margaret Lech, Lawrence Cavedon *Neural Networks, March 2017.*
2. Recognition of emotions using multimodal physiological signals and an ensemble deep learning model, Zhong Yin, Mengyuan Zhao, Yongxiong Wang, Jingdong Yang, Jianhua Zhang, *Computer Methods and Programs in Biomedicine, Volume 140, March 2017, Pages 93-110.*
3. A facial expression recognition system using robust face features from depth videos and deep learning, Md. Zia Uddin, Mohammed Mehedi Hassan, Ahmad Almogren, Mansour Zuair, Giancarlo Fortino, Jim Torresen, *Computers & Electrical Engineering, April 2017*
4. Emotion-modulated attention improves expression recognition: A deep learning model Original Research Article, Pablo Barros, German I. Parisi, Cornelius Weber, Stefan Wermter *Neurocomputing*, March *2017.*