International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 9 Issue 7 ǁ 2021 ǁ PP. 06-11

A Survey of Deep learning Techniques on Medical Image Processing

Baah Barida[1], Wosu Ednah Nwereka[2]
[1]*Department of Computer Science, Ebonyi State University, Abakaliki- Nigeria*
[2]*Department of Computer Science, Captain Elechi Amadi Polytechnic,Rumuola,Port Harcourt-Nigeria*

ABSTRACTS
Deep learning is an improved method of artificial neural network, has been applied to various problem domain of application as many researcher has identified. However, in this research work we take a survey on four various types of deep learning techniques like convolution neural network, recurrent neural network, generative adversarial neural network and self organizing map which will be based on medical images with the view of looking at its various strength and weakness of these techniques and taking into cognizant the techniques that is more advantageous compare to its weakness level.
KEYWORD: *Deep Learning, CNN, RNN, GAN, and SOM.*

Date of Submission: 25-06-2021 Date of acceptance: 08-07-2021

I. INTRODUCTION

Deep learning is an aspect of machine learning that is born from artificial intelligence, deep learning is becoming a most used technique in most problem solving domain. It is widely used in the areas of medical imagery to serve as a decision support system for medical practitioner. It is also important to note that deep learning are utilized in the case of an unlabeled information during training; it is thus well-suited to addressing heterogeneous information and data, in order to learn and acquire knowledge [1]. In comparism with the traditional machine learning, deep learning is far superior as it can learn from raw data, and has multiple hidden layers which allow it to learn abstractions based on inputs [2].

The deep learning model works in three basic steps which involves: taking some data as input, training a model based on input data, and using trained model for making predictions concerned with the new data. The procedure of training the model can be stated as learning process in which the model is made exposed to unfamiliar new data at each step [4]. For every step, model predicts and receives feedback concerning the accuracy of predictions generated. The feedback is used for correcting errors made while prediction. If a parameter of model is tweaked, resulting in a correct prediction, model may end up portraying the previous prediction wrong, even if it was stated correct initially. Thus, representing learning process as a back-and-forth method in parameter space. It might take several iterations for training the model, with a blend of noble predictive performance. The process continues till these predictions from the model stops to improve[5]

Secret to deep learning skills lies in the capability of the neural networks to learn from data through general purpose learning procedure [2].The main reason for this research work is to review the applications of deep learning in medical diagnosis in a concise and simple manner and most importantly also take a look at different deep learning techniques there strength and weakness.

II. REVIEW OF RELATED LITERATURE

Many research works has been done in the aspect of deep learning utilization on medical images for images that are obtained either from the magnetic resonance image (MRI) or the computed tomography image scan. We review related works on deep learning techniques it relate to image processing:

Araújo and Costa [12] presented a new SOM with a variable topology for image segmentation. The proposed fast convergent network is capable of color segmenting images satisfactorily with a self-adaptive topology.

Sujaritha and Annadurai [14] proposed a fully automatic three-level clustering algorithm for color-texture segmentation which self organizing map is used to identify the number of components and initialize the Gaussian mixture model. Experimental results indicate that the proposed algorithm is efficient and competent with popular CTex and JSEG algorithms.

N. Srivastava, R. R. Salakhutdinov [16] suggested one of the earliest attempts to introduce category hierarchy in CNN, in which a discriminative transfer learning with tree-based priors is proposed. They use a hierarchy of classes for sharing information among related classes in order to improve performance for classes with very few training examples.

Xiao et al. [17] propose also a CNN training method that grows a network not only incrementally but also hierarchically. In their method, it utilizes classes that are grouped according to similarities and are self-organized into different levels

Wim De Mulder et.al [18] covers recurrent neural nets for language modeling. Other resources focus on a specific technical aspect such as Barak A Pearlmutter[19], which surveys gradient calculations in recurrent neural networks.

Andrej Karpathy et. al[20] proposed recently a recurrent neural networks which was used successfully for image captioning in his work he set up a training set which consists of input images x and target captions y. Given a large set of image-caption pairs, a model is trained to predict the appropriate caption for an image

2.2 Types of Deep learning Techniques

There are various types of deep learning techniques that are in used, some of this deep learning techniques are stated and discussed below:

2.2.1 Convolutional Neural Networks

The convolutional neural network is a branch of neural networks and consists of a stack of layers each performing a specific operation, e.g., convolution, pooling, loss calculation, etc. Each intermediate layer receives the output of the previous layer as its input as illustrated in Figure 1 below. The beginning layer is an input layer, which is directly connected to an input image with the number of neurons equal to the number of pixels in the input image. The next set of layers are convolutional layers that present the results of convolving a certain number of filters with the input data and perform as a feature extractor. The filters, commonly known as kernels, are of arbitrary sizes, defined by designers, and depending on the kernel size. Each neuron responds only to a specific area of the previous layer, called receptive field. The output of each convolution layer is considered as an activation map, which highlights the effect of applying a specific filter on the input. Convolutional layers are usually followed by activation

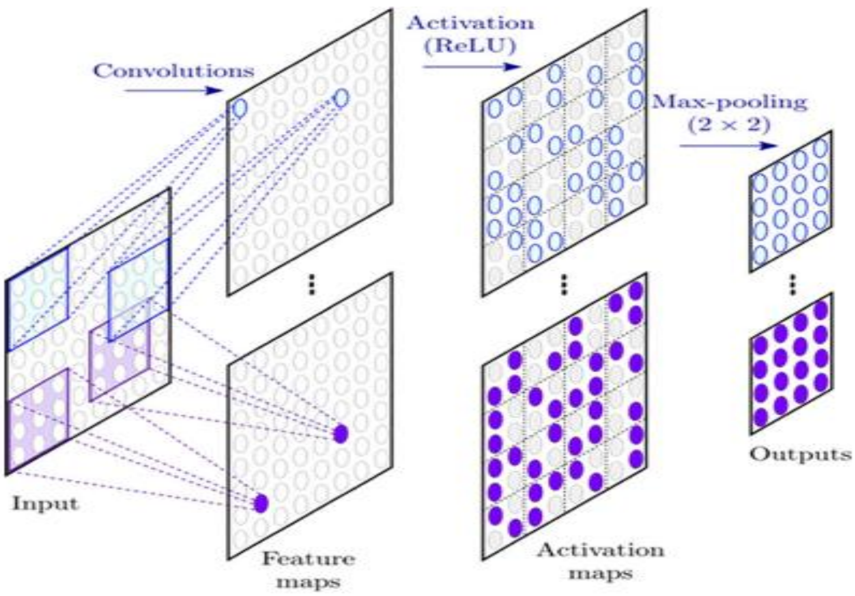

Figure 1: A typical structure of A CNN
Source: Commandeur F. et al.[6]

2.2.2 Recurrent Neural Networks(RNN)

The RNN is empowered with recurrent connections which enables the network to memorize the patterns form last inputs. The fact that the region of interest in medical images usually distributed over multiple adjacent slices (e.g., in CT or MRI), results in having correlations in successive slices. Accordingly, RNNs are able to extract inter-slice contexts from the input slices as a form of sequential data. The RNN structure consists of two major sections of intra-slice information extraction which can be done by any type CNN models, and the RNN, in charge of inter-slice information extraction. Figure 2 diagram shows a recurrent neural network.

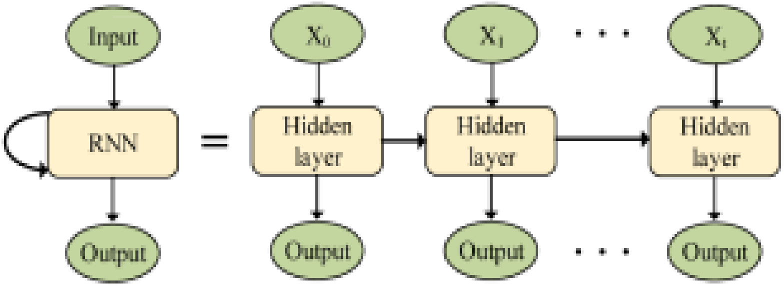

Figure 2: A schematic diagram of recurrent neural network
Source: Wei Chen et al. [7]

2.2.3 Generative Adversarial Network

A generative adversarial network (GAN) is a class of machine learning frameworks designed by Ian Goodfellow and his colleagues in 2014[8] Two neural networks contest with each other in a game (in the form of a zero-sum game, where one agent's gain is another agent's loss).

Given a training set, this technique learns to generate new data with the same statistics as the training set. For example, a GAN trained on photographs can generate new photographs that look at least superficially authentic to human observers, having many realistic characteristics. Though originally proposed as a form of generative model for unsupervised learning, GANs have also proven useful for semi-supervised learning,[9[fully supervised learning,[10] and reinforcement learning.[11]

The core idea of a GAN is based on the "indirect" training through the discriminator, which itself is also being updated dynamically. This basically means that the generator is not trained to minimize the distance to a specific image, but rather to fool the discriminator. This enables the model to learn in an unsupervised manner.

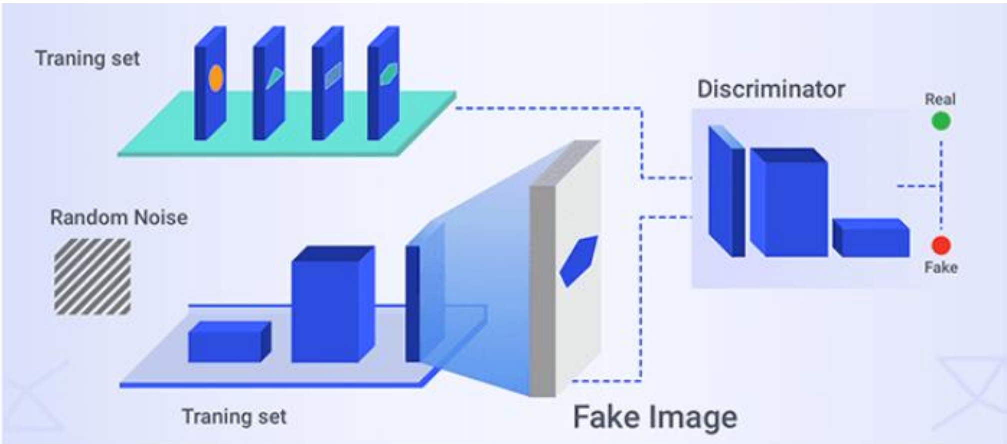

Figure 3: A diagram showing a generative adversarial network

2,2,3.1 Comparison of Generative Adversarial Networks (GAN) and CNN

If GAN is compared to CNN it is proven that GAN has generation capability and it can work with unlabeled data because it belongs to unsupervised learning techniques.

2.4 Self Organizing Maps (SOM)

Self Organizing Map (SOM) is method that utilizes unsupervised learning where the network does not utilize the class membership of sample training, but use the information in a group of neurons to modify the local parameter. The SOM system is adaptively classify samples (X image data) into classes determined by selecting the winning neurons are competitive and the weights are modified.

SOM, first put forward by Kohonen [14], is a kind of widely used unsupervised artificial neural network. The map is a group of node units represented by prototype vectors lying in a 2-dimension space usually though occasionally nodes are set in one or multi-dimensional space. These units are connected to adjacent units by a neighborhood function. Prototype vectors are initialized with random or linear methods and "folded" in the 2-dimension space. Then, they are trained iteratively by randomly selected input samples sequentially or in batches and updated according to the neighbor function. After the training, prototype vectors become stable and "unfolded" themselves in the 2-dimension-space map.

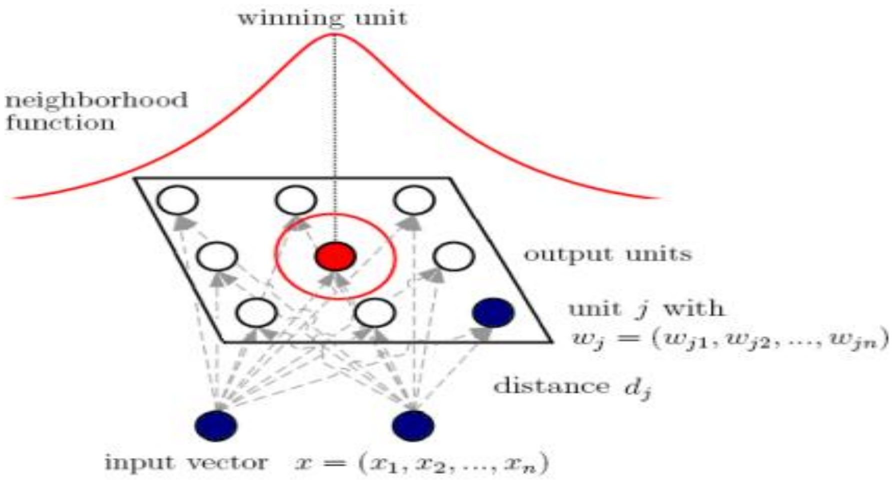

Figure 4: A Topology of Self Organizing Map
Source: Daniel T. Larose [15]

From the figure 4 above the implementation of a Kohonen Self-Organizing Map incorporates a m×m quadratic matrix where each input unit x = (x1, x2, ..., xn) is connected to all output units. Each output unit j is associated to the weight vector wj = (wj1, wj2, ..., wjn) whereas weights have the same dimensionality as input vectors. The distance of the output vector j to an input signal is referred to as dj

III. METHODOLOGY

3.1 Existing Stage of Processing Data

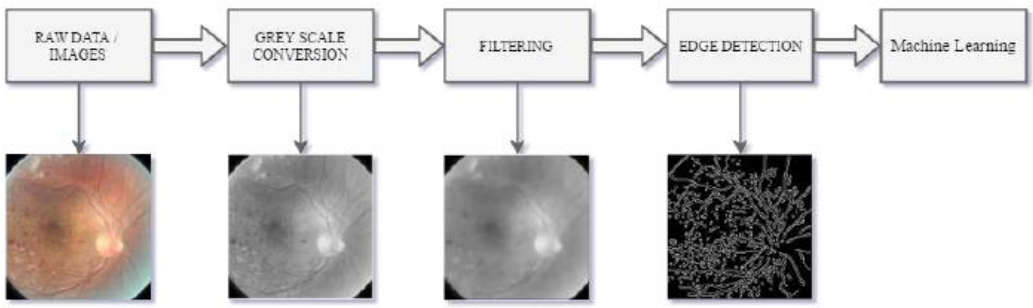

Figure 5: Stages of Processing of Medical Images
Sources: Suvajit Dutta et. al [3]

From figure 5 above the images or raw data has been extracted or collected contains noise like blur, high contrast. With this noise, it is difficult to extract desired features from the images, which may result in miss-classification of data during training, validation & testing. In order to extract proper features image processing is necessary for features detection or features extraction. The above-mentioned process is adapted for the image processing in this project. Where the raw data is first converted into grey scale i.e., converting RGB (3-bands) image into grey-scale (1-band) image. On this gravy-scaled image median filtered is applied to remove noise, so that all the pixels are normalized to the data around the pixel intensity values. From the median filtered image, we can extract features such as white lesion"s and few thin veins. When we subtract the median filtered image from the grey scaled image we can get the exudates. which plays the key role in the feature extraction. Edge detection is applied on this difference data to highlight the features. Then this edge detected image is taken as input for the machine learning algorithms as the input data to proper classification

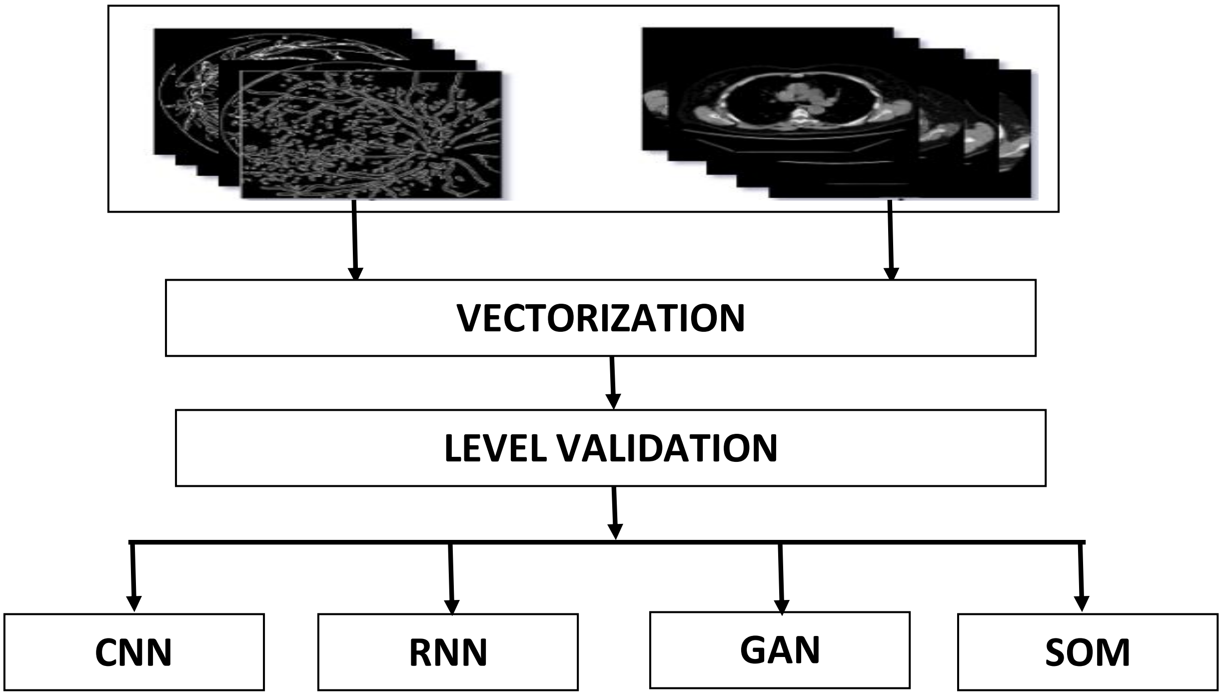

Figure 6: Proposed hierarchical diagram for Deep learning techniques

IV. SUMMARY OF DEEP LEARNING TECHNIQUES: STRENGTH AND WEAKNESS

In this section we present a summary table for the various deep learning techniques discussed based on their various strength and weakness as discussed in the table 1 below:

Table 1: Deep learning techniques strength and weakness

Deep Learning Techniques	**Strength**	**Weakness**
CNN	1. It provide a dense network which is use to perform prediction and identification 2. It produces a more accurate results when using larger data set.	1. It does not encode the position and orientation of the object 2. Lack the ability to be spatially invariant to the input data
RNN	1. It can model a collection of records (i.e. time collection) so that each pattern can be assumed to be dependent on previous ones. 2. Recurrent neural networks are even used with convolutional layers to extend the powerful pixel neighbourhood.	1. It has the vanishing or exploding gradient problem. 2. The RNNs cannot be stacked up. 3. It has slow and Complex training procedures. 4. It is difficult to process longer sequences
GAN	1. It has the capability of working with unlabeled data 2. It also reduces the load required for a deep neural network because the two systems share the burden.	1. Harder to train: You need to provide different types of data continuously to check if it works accurately or not. 2. Generating results from text or speech is very complex
SOM	1. It is easy to interpret and understand data. 2. It is easy to observe similarities in data due to reduction in dimensionality and grid clustering.	1. Lack of neuron weight sufficient to aid in the clustering of inputs 2. It computationally expensive compare to K-means or other algorithms 3. It may not be entirely accurate or informative due to grouping in the map.

V. CONCLUSION

Conclusively, there is absolute know preferable deep learning techniques that this the best, the choice of the techniques depend on what is best for you to achieve a greater result based on the kind of medical images, but however in the case of the recurrent neural network is very much effective in both cognitive and supervised learning and less effective in other areas of learning medical image, also for the others techniques discussed above all have their strength and weakness.

REFERENCES

[1]. Chen, X.-W.; Lin, X. Big data deep learning: Challenges and perspectives.IEEE Access 2014,2, 514–525.[CrossRef]

[2]. Miotto, R.; Wang, F.; Wang, S.; Jiang, X.; Dudley, J.T (2017), Deep learning for healthcare: Review, opportunities and challenges.Brief Bioinform.. [CrossRef] [PubMed

[3]. Suvajit Dutta, Bonthala CS Manideep, Shalva Rai and Vijayarajan V(2017), A Comparative Study of Deep Learning Models for Medical, IOP Conference Series: Materials Science and Engineering pp. 1-9

[4]. Glorot, Xavier, Antoine Bordes, and Yoshua Bengio. "Deep Sparse Rectifier Neural Networks."Aistats. Vol. 15. No. 106. 2011.

[5]. Bengio, Yoshua. "Learning deep architectures for AI."Foundations and trends® in Machine Learning2.1 (2009): 1-127.

[6]. Commandeur F, Goeller M, Betancur J, Cadet S, Doris M, Chen X, Berman DS, Slomka PJ, Tamarappoo BK, Dey D(20180 Deep learning for quantification of epicardial and thoracic adipose tissue from non-contrast CT. IEEE Trans Med Imaging 37(8):1835–1846, 2018

[7]. Wei Chen, Liang Zhang, Jia Song, Yanhua Wang, Yang L(2018), Polarimetric radar target recognition framework based on LSTM, IET International Radar Conference

[8]. Goodfellow, Ian; Pouget-Abadie, Jean; Mirza, Mehdi; Xu, Bing; Warde-Farley, David; Ozair, Sherjil; Courville, Aaron; Bengio, Yoshua (2014). *Generative Adversarial Networks(PDF)*. Proceedings of the International Conference on Neural Information Processing Systems (NIPS 2014). pp. 2672–2680.

[9]. Salimans, Tim; Goodfellow, Ian; Zaremba, Wojciech; Cheung, Vicki; Radford, Alec; Chen, Xi (2016). "Improved Techniques for Training GANs". *arXiv*:*1606.03498* [*cs.LG*].

[10]. Isola, Phillip; Zhu, Jun-Yan; Zhou, Tinghui; Efros, Alexei (2017). *"Image-to-Image Translation with Conditional Adversarial Nets"*. Computer Vision and Pattern Recognition.

[11]. Ho, Jonathon; Ermon, Stefano (2016). *"Generative Adversarial Imitation Learning"*. Advances in Neural Information Processing Systems. **29**: 4565–4573. *arXiv*:*1606.03476*. *Bibcode*:*2016arXiv160603476H*

[12]. A. R. F. Araújo and D. C. Costa, "Local adaptive receptive field self-organizing map for image color segmentation," *Image and Vision Computing*, vol. 27, no. 9, pp. 1229–1239, 2009.

[13]. M. Sujaritha and S. Annadurai (2010), "A three-level clustering algorithm for color texture segmentation," in *Proceedings of the International Conference and Workshop on Emerging Trends in Technology (ICWET '10)*, pp. 645–650, Mumbai, India.

[14]. T. Kohonen,(2001) *Self-Organizing Maps*, Springer, 3rd edition, 2001.

[15]. Daniel T. Larose (2004), Discovering Knowledge in Data: An Introduction to Data Mining. Wiley-IEEE Press, 2004.

[16]. N. Srivastava, R. R. Salakhutdinov (2013), Discriminative transfer learning with tree-based priors, in: Proceedings of the Advances in Neural Information Processing Systems (NIPS), 2013, pp. 2094–2102.

[17]. T. Xiao, J. Zhang, K. Yang, Y. Peng, Z. Zhang(2014), Error-driven incremental learning in deep convolutional neural network for large-scale image classification, in: Proceedings of the ACM Multimedia Conference, 2014, pp. 177–186.

[18]. Wim De Mulder, Steven Bethard, and Marie-Francine Moens(2015),. A survey on the application of recurrent neural networks to statistical language modeling. Computer Speech & Language, 30(1):61–98,.

[19]. Barak A Pearlmutter (1995), Gradient calculations for dynamic recurrent neural networks: A survey. Neural Networks, IEEE Transactions on, 6(5):1212– 1228, 1995

[20]. Andrej Karpathy and Li Fei-Fei.(2014), Deep visual-semantic alignments for generating image descriptions. arXiv preprint arXiv:1412.2306.