Deep learning applications for analysis in Medical Imaging and Computer Vision

Shital Patil
Research Scholar
EE,VJTI MUMBAI

Surendra Bhosale
EE,VJTI MUMBAI

Abstract - Machine learning has a vital role in Image Analysis and Computer Vision field. Problems ranging from image segmentation, image registration to structure-from-motion, object recognition and scene understanding use machine learning techniques to analyze information from visual data. The medical image analysis is growing field of deep learning. DL techniques and their applications to medical image analysis includes standard ML techniques in the computer vision field, ML models in deep learning and applications to medical image analysis. One of the most recently use of ML in computer aided diagnosis and medical image analysis is the classification of objects such as lesions into certain classes based on input features like contrast, area obtained from segmented objects. Artificial neural network conceptually inspired by neural systems.

The important deep-learning techniques including the Neocognitron, CNNs, neural filters. ML with image input including deep learning is a useful technology with higher performance. The deep learning will become the mainstream technology in medical image analysis in upcoming decades.

Keywords - Convolutional neural networks (CNNs), ANN, CBIR, Medical image analysis, Machine learning, Deep Learning.

I. INTRODUCTION

Machine learning techniques are widely used in medical imaging research field as successful classifier and clustering algorithms [1]. The Best classifiers used are support vector machine and clustering algorithms, such as k-nearest neighbor (k-NN) [3]. Now a days deep learning (DL) has come into the picture as the methodology to effectively improve the performance of existing machine learning techniques. Next, Deep learning is a generic methodology that has a disruptive impact in other scientific fields as well. Therefore, it has become imperative for medical imaging researchers to fully embrace Deep Learning technology.

Medical image processing refers to a set of procedures so as to get clinically meaningful information from various imaging modalities, commonly for diagnosis or prognosis. The modalities are typically in vivo types. The extracted information/data could be used further to enhance diagnosis and prognosis according to the need of patient's.

The main comparison between Machine learning with image input including "deep learning" and Machine learning with feature input is the direct use of pixel values with machine learning model. Machine Learning algorithms have the capability to be invested deeply in vast area of medicine right from discovery of drugs to clinical decision making, significantly altering the way medicine is practiced. The success of machine learning algorithms at computer vision tasks in recent years comes at time when medical records are increasingly digitalized. Therefore, it is ideal for medical image analysis to be carried out by an automated, accurate and efficient machine learning algorithm.

Deep learning has got great interest in almost each and every field and especially in medical image analysis and it is expected that it will hold $300 million medical imaging market by 2021. Therefore, by 2021, it will get more investment for medical imaging than the entire analysis industry spent in 2016. It is the most effective and supervised machine learning approach. This approach use models of deep neural network which is variation of Neural Network but with large approximation to human brain using advance mechanism as compare to simple neural network. The term deep learning implies the use of a deep neural network model. The basic computational unit in a neural network is the neuron, a concept inspired by the study of the human brain, which takes multiple signals as inputs, combines them linearly using weights, and then passes the combined signals through nonlinear operations to generate output signals.

Machine Learning (ML) and Artificial Intelligence (AI) have progressed rapidly in recent few years. ML and AI have played important role in medical field like medical image processing, computer-aided diagnosis, image interpretation, image fusion, image registration, image segmentation, image retrieval and analysis.ML extracts information from the images and represents information effectively and efficiently. The ML and AI together can diagnose and predict accurate and faster the risk of diseases and prevent them in time. These techniques enhance the abilities of doctors and researchers to understand that how to analyze the generic variations which will lead to disease. These techniques composed of conventional algorithms without learning like Support Vector Machine (SVM), Neural Network (NN), KNN etc. and deep learning algorithms such as Convolutional Neural Network (CNN), Recur-rent neural Network (RNN), Long Short term Memory, Extreme Learning Model etc.

A. *Types of Medical Imaging:*

Medical Imaging is the use of imaging modalities and processes to get pictures of the human body, which can assist diagnosis and treatment of patients. It can also be used to track any ongoing issues, and can therefore help with treatment plans. There are many different types of medical imaging techniques, which use different technologies to produce images for different purposes. Here the most common imaging techniques uses AI in radiology indicates how these techniques, mixed with AI, will direct the way for more accurate imaging.

Various medical imaging modalities and digital medical images incorporates as magnetic resonance imaging (MRI), computed tomography (CT), X-ray Computed Tomography and positron emission tomography (PET), Single Photon Emission Computed Tomography (SPECT) etc could provide specific information for the patient being imaged. Research in medical image processing mainly targets to extract important features that might be difficult to assess with the naked eye.

A histology slide is an image file of a few megabytes while a single MRI may be a few hundred megabytes. This has technical implications on the way the data is pre-processed, and on the design of an algorithm's architecture, in the context of processor and memory limitations.

B. *Background of medical image analysis:*

Initially, medical image analysis was done using sequential application of low level pixel processing and mathematical modeling to construct a rule-based system that could solve only specific task. Similarly there were some rules, likely in the area of Artificial Intelligence commonly known as *GOFAI* (Good Old Fashioned Artificial Intelligence) agent[3].

At the end of *1990s*, supervised techniques were becoming more popular where training data was used to train models and they were becoming increasingly popular in medical image analysis field. Examples may involves active shape model , atlas method. This pattern recognition and machine learning is more popular with the introduction of some innovative ideas. Therefore, Change in shift from systems that were designed by humans to systems that are trained by computers based on example data.

Machine learning includes constructing data-driven models to solve research problems [5]. There are two basic categories of ML as supervised learning and un-supervised learning. In supervised learning, The training data consist of a set of training examples and each example is a pair consisting of an input object and a desired output value. Supervised machine learning systems provide the learning algorithms with known quantities to support future judgments. we train the models using input data with matched labels [6]. The model is a mathematical model that can associate input data with the matched labels and a predictive model, which is validated using unseen test data. In Unsupervised Learning is a class of Machine Learning techniques to find the patterns in data. The data given to unsupervised algorithm are not labeled, which means only the input variables(X) are given with no corresponding output variables.

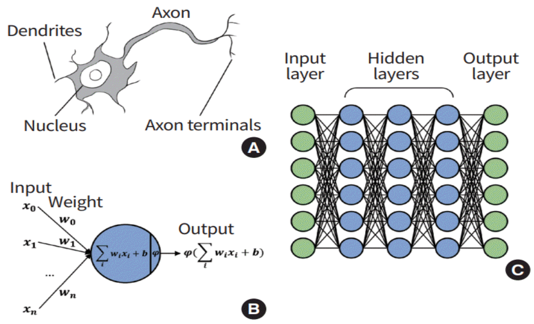

Figure 1: Artificial neural network. (A) A neuron, transforms the inputs from the dendrites into electric signals. (B) In a node , the input values are transformed by the weights, biases, and activation functions. The output values are transmitted to the next perceptron. (C) Multilayer perceptron consists of multiple perceptrons.

ANN is a statistical method inspired by brain mechanism from neuroscience as shown in figure 1 [7]. A typical neuron is the basic unit of the very important brain system. The neuron is an electrically excitable cell that receives signals from other neurons, it processes the received information, and transmits electrical signals to other neurons. The input signal to a given neuron needs to exceed a certain threshold for it be activated and further transmit a signal. The neurons are interconnected to each other and forms a network that collectively steers the brain mechanism. ANN is an abstraction of an interconnected network of neurons with layers of nodes, and it consists of an input layer aggregating the input signal from other connected neurons, a hidden layer responsible for training and an output layer [8]. Each node takes the input from nodes from the previous layer using various weights and computes the activation function, which is relayed onto the next layer of nodes. The activation function approximates the complex process of a physical neuron, which regulates the strength of the neuronal output in a non-linear manner. The mathematical processing in a node can be represented using the following equation:

$$Output = \varphi\left(W^T x + b\right).$$

A node takes an input value ‘x’ and multiplies it by weight ‘W,’ and then a bias of ‘b’ is added, which is fed to the activation function ‘φ.’

The losses are back-propagated through the network, and they are used to modify the weights and biases [6].

C. *Deep Learning in Image analysis :*

The accurate and most useful type of models for analysis of images till date are Convolutional Neural Networks as shown in figure 2. A single CNN model has many layers which work on identifying edges and normal features on shallower layers and more deep features in deeper layers. An image is convolved with filters and after that pooling is applied, this process may be applied for some layers and at last recognizable features are extracted.

Introduction of *GPUs* have favored the research in this field and since the introduction of challenge, a sudden rapid growth in development of such models may be seen.

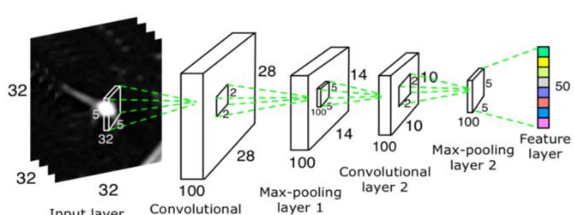

Figure 2 : Illustration of CNN

The medical image analysis community has taken notice of these pivotal developments. However, transition from systems that used handcrafted features to systems that learn features from data itself has been gradual.

D.Organization :

The aim of this paper is to provide an overview of machine learning algorithms and thrust areas of applications such as applied to medical imaging, with an prominence that will be most useful to the doctors, clinicians and provide more impact in medical science and patient healthcare system.
In **Section II**, It describes various machine learning architectures used in medical image analysis, with an prominence on CNNs. Machine learning is broadly classified into Supervised, Unsupervised, Semi-supervised and Reinforcement learning methods, it is the first two which are currently most applicable to image analysis. **Section III** divides into different application areas. **Section IV** gives conclusion with obstacles that the field of medical image analysis faces, and some of the future possible directions.

II. MACHINE LEARNING ARCHITECTURES

1) Convolutional neural networks:

a)Convolution Layer: A convolution is defined as an operation on two functions. In image analysis, one function consists of input values (e.g. pixel values) at a position in the image, and the second function is a filter (kernel) each can be represented as array of numbers.

Convolution shows ideas intrinsic to perform computationally efficient machine learning with sparse connections, parameter sharing [8]. Few neural networks where each input neuron is connected to every output neuron in the subsequent layer, In CNN neurons only few inputs are connected to the next level layer.

The convolution operation is defined by the * symbol. An output s(t) is defined below when input I(t) is convolved with a filter or kernel K(a) .

$$s(t)=(I*K)(t). \qquad (1)$$

Now if t can only take integer values, then discretized convolution can be expressed as,

$$s(t)=\sum I(a)\cdot K(t-a) \qquad (2)$$

The equation (2) assumes a one-dimensional convolutional operation.

Two dimensional convolution operation with input I (m, n) and a kernel K (a, b) is expressed below as:

$$s(t)=\sum_a\sum_b I(a,b)\cdot K(m-a,n-b).$$

(3)

Next the kernel is flipped and the above equation can be written as,

$$s(t)=\sum_a\sum_b I(m-a,n-b)\cdot K(a,b). \qquad (4)$$

Further Neural networks implement the cross-correlation function, and it is similar to as convolution but without flipping the kernel.

$$s(t)=\sum_a\sum_b I(m+a,n+b)\cdot K(a,b). \qquad (5)$$

b) Rectified Linear Unit (RELU) Layer: The Rectified Linear Unit is the most commonly used activation function in deep learning models. The function returns 0 if it receives any negative input, but for any positive value *x* it returns that value back.

So it can be written as

$$f(x)=max(0,x). \qquad (6)$$

where x is the input to the neuron. Other activation functions include the sigmoid, tanh, leaky RELUs.

It's wondering that such a simple function can allow your model to account for non-linearities and interactions so well. But the ReLU function works great in most applications, and it is very widely used as a result.

b) *Pooling Layer:*

Convolutional networks may include local or global pooling layers which combine the outputs of neuron clusters at one layer into a single neuron in the next layer [9]. The major concept of CNNs is pooling, which is a form of non-linear down-sampling. There are many non-linear functions to implement pooling out of which which max pooling is the common one. It partitions the input image into a set of non-overlapping rectangles and, for each such sub-region, outputs the maximum. The pooling layer operates independently on every depth slice of the input and resizes it spatially.

The pooling layer is there in between the Convolution and RELU layers to reduce the number of parameters to be calculated, as well as the size of the image.

c) *Fully Connected Layer:*

Fully connected layers connect every neuron in one layer to every neuron in another layer. It is in principle the same as the traditional multi-layer perceptron neural network . After several convolutional and max pooling layers, the high-level reasoning in the neural network is done via fully connected layers. Neurons in a fully connected layer have connections to all activations in the previous layer, as seen in regular (non-convolutional) artificial neural networks. Their activations can thus be computed as an affine transformation, with matrix multiplication followed by a bias offset.

2. *Recurrent neural networks (RNNs):*

A recurrent neural network (RNN) is an extension of a conventional feed forward neural network, which is able to handle a variable-length sequence input. It handles the variable-length sequence by having a recurrent hidden state whose activation at each time is dependent on that of the previous time.

Because of its ability to generate text [10], RNNs have been used in text analysis tasks as machine translation, speech recognition, language modelling, text prediction and image caption generation [11]. In general, the output of a layer is added to the next input, and this is fed back into the layer, resulting in a capacity for contextual 'memory'.

Therefore to avoid vanishing gradient problems with back propagation through time, plain RNNs have evolved into Long Short Term Memory (LSTM) networks and Gated Recurrent Units (GRUs). These are modifications of RNNs to hold long term dependencies and to discard or forget some of the accumulated information.

In the medical image analysis , RNNs have been used mainly in segmentation[12], together CNN as well RNN to segment neuronal and fungal structures from three-dimensional electron microscope images.

III. APPLICATIONS IN MEDICAL IMAGE ANALYSIS

- **Classification** : This is the first areas where in medical image analysis where deep learning was used. Diagnostic image classification includes classification of diagnosed images, in such setting every diagnosed exam is a sample and data size is less than that of a computer vision.
- **Detection** : Anatomical object localization such as organs is important pre-processing part of segmentation task. Localization of object in a image requires *3D* parsing of image, several algorithms have been proposed to convert *3D* space as composition of *2D* orthogonal planes.
- **Segmentation** : The segmentation of organs and other substructures in medical images allows quantitative analysis related to shape, size and volume. The task of segmentation is typically defined as identifying set of pixels that define contour or object of interest. Segmentation of lesions combines the challenge of object detection and organ and substructure segmentation in the application of deep learning algorithms.

- **Registration** : Referred as spatial alignment is common image analysis task in which coordinate transform is obtained from one image to another image. Though lesion detection and object segmentation are eyed as main use of deep learning algorithms but researchers have found that deep networks can be beneficial in getting best possible registration performance.

  - *Future Applications:*

There are many challenges in area of Deep Learning in medical image analysis, Unavailability of large dataset is often mentioned as one. Some of the innovative applications that span across traditional medical image analysis categories are described in medical imaging. Content based image retrieval **(***CBIR*) is a technique for knowledge discovery in large databases and offer similar data retrievals for case histories and understand rare disorders. Image generation and enhancement is another task that uses Deep Learning in improve image quality, normalizing images, data completion and pattern discovery. Combining Image data with reports is yet another task that seem to have a very large scale application in real world.

IV. CONCLUSION :

Deep learning is widely used in all areas and it will continue to grow in the nearer future in almost all fields of science. Scientists and expert peoples in this area are actively involved in research teams to solve critical severe medical problems. Medical image processing will benefit immensely from deep learning approaches as it has shown remarkable performance in non-medical regular imaging research compared to conventional machine learning approaches. Here we have explained a brief history from traditional machine learning to deep learning, highlighted various deep learning applications in medical imaging finally concluded with few drawbacks and future perceptions of deep learning in medical imaging. This will be definitely an essential tool for diagnosis and prognosis in the era of precision medicine.

REFERENCES

[1] Wang S, Summers RM. *Machine learning and radiology*. Med Image Anal 2012;16:933–51.
[2] J. Ker, L. Wang, J. Rao and T. Lim, "*Deep Learning Applications in Medical Image Analysis,*" in *IEEE Access*, vol. 6, pp. 9375-9389, 2018.
[3] Wernick MN, Yang Y, Brankov JG, Yourganov G, Strother SC. Machine learning in medical imaging. *IEEE Signal Process Mag 2010;27:25–38.*
[4] Nasrabadi NM. Pattern recognition and machine learning. J Electron Imaging 2007;16:49901.
[5] Hagan MT, Demuth HB, Beale MH, De Jesus O. Neural network design. 2nd ed. Boston (MA): PWS Publishing; 2014.
[6] Le QV, Ngiam J, Coates A, Lahiri A, Prochnow B, Ng AY. On optimization methods for deep learning. In: Getoor L, Scheffer T, editors. Proceedings of the 28th International Conference on Machine Learning; 2011 Jun 28-Jul 2; Bellevue, WA. Madison (WI): Omnipress; 2011. p. 265–72.
[7] Prospects of deep learning for medical imaging, Precis Future Med. 2018;2(2):37-52.
[8] Goodfellow, Y. Bengio, A. Courville, Deep Learning, Cambridge, MA, USA:MIT Press, 2016.
[9] Ciresan, Dan; Ueli Meier; Jonathan Masci; Luca M. Gambardella; Jurgen Schmidhuber, "Flexible, High Performance Convolutional Neural Networks for Image Classification". Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence-Volume Volume Two. 2: 1237–1242. Retrieved 17 November 2013.
[10] Sutskever, J. Martens, G. E. Hinton, "Generating text with recurrent neural networks", Proc. 28th Int. Conf. Mach. Learn. (ICML), pp. 1017-1024, 2011.
[11] A. Karpathy, L. Fei-Fei, "Deep visual-semantic alignments for generating image descriptions", *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 3128-3137, Jun. 2015.
[12] J. Chen, L. Yang, Y. Zhang, M. Alber, D. Z. Chen, "Combining fully convolutional and recurrent neural networks for 3D biomedical image segmentation", *Proc. Adv. Neural Inf. Process. Syst.*, pp. 3036-3044, 2016.

Author's Biography:

Ms. Shital Patil graduated from Pune University in 2006, Completed her Post graduation in Electronics Engineering from Mumbai University. Currently working on Machine learning algorithms for her Doctoral program. She has more than 10 y ing and research experience. Her research areas of interests are electronics for societal needs, instrumentation for agriculture, healthcare especially for children and women and medical applications of deep learning ML algorithms.

Surendra Bhosale is an Associate Professor in Electrical Engineering and is a Ph.D. guide. He had received B.E.Degree in Electrical Engg.from Shivaji University, Kolhapur, Maharashtra and M.E. Degree in Electrical from the University of Mumbai, Maharashtra. He has obtained his Ph.D. Degree in Electrical Engineering from the University of Mumbai, India. Currently, He is an Associate Professor of Electrical Engineering Department at VJTI, Mumbai. He has more than 31 years of teaching experience.
He was incharge of the state monitoring and evaluation cell for world bank project name TEQIP-2 (Technical Education Quality Improvement Programme) 2011-2016.
His teaching and research areas include Wireless Communications and Routing algorithm.

JETIR