DOI : https://doi.org/10.56452/7-3-109

ISSN: 0974-5823 Vol. 7 No. 3 March, 2022

International Journal of Mechanical Engineering

Review of application of Deep Learning based Image recognition in Disease diagnosis

Devesh Tiwari

Asst. Professor, Department of Comp. Sc. & Info. Tech., Graphic Era Hill University, Dehradun, Uttarakhand India 248002

Abstract:

Due to the environmental and life style changes, diseases are increased exponentially. For Radiologists and clinicians, to diagnose the condition or stages of disease by the conventional image analysis method becomes difficult as mutants changes and adapt the environment vary fast. Deep convolutional neural networks and deep leaning is a boon for medical imaging analysis in this era, that's why most deep learning image analysis approaches attracts the researcher and scientists. This paper reviews the applications in area of medical imaging analysis based on deep learning. Initially the summery of medical imaging analysis, its characteristics, principles, Deep learning algorithms are described. Furthermore, applications in area of medical imaging analysis are systematically reviewed. Additionally, hidden challenges and forthcoming exploration directions are concise.

Keywords*:* Deep learning algorithms , image analysis, medical imaging analysis, Deep convolutional neural networks.

I. INTRODUCTION

In area of image processing, medical imaging usually refers to visual representations of human body organs. Doctors can use medical imaging to diagnose diseases or treat diseases more efficiently [1]. In the past, medical imaging became an effective tool to diagnose diseases for patients. Modalities of medical imaging include Computed Tomography (CT), histology slides, mammography, Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), retinal photography, Ultra Sound (US) [2]. Clinically, medical imaging needs to be interpreted in detail by radiologists or clinicians. Therefore, clinicians or researchers improve their efficiency using computer-assisted diagnosis from the early 1980s. The computer-assisted diagnosis mainlyuses machine learning algorithms to extract features from medical imaging data and plays an important function in medical imaging analysis [3,4].

Machine learning technologies have great latent capacity in all fields of medicine, significantly changing the way medicine is practiced from drug discovery to clinical decision making. Machine learning technologies have enormous success in computer vision tasks, and provide an opportunity for medical imaging analysis when medical imaging is more and more digitalized. However, with the rapid accumulation of various medical imaging data, the interpretations of medical imaging are time-consuming and easily prone to error due to the difference of domain knowledge of experts. Every unreasonable diagnosis may cause harm to patients. So it has brought great challenges to domain experts who diagnose disease through medical image analysis. Moreover, the shallow structure of these machine learning model have limitation in their representational power, so we need new research directions to let learning algorithms learn automatically latent disease features from medical imaging.

Recently, with the deep application of artificial networks, deep learning (DL) [5] as a branch of machine learning was proposed. Be similar to our human brain, the artificial neural networks also comprises interconnections between neurons. All neurons are connected in the human brain,whereas there are discrete layers and connections in artificial neural networks. Deep learning architecture usually also includes many hidden layers. A representation of high-level feature is formed by combining low-level features in multiple hidden layers.

In past years, with the continuous improvement of computing technologies and machine learning algorithms, DL has been utilized in many fields including target detection, natural language processing, speech recognition, face recognition, and so on. Especially, the breaking success in fields of computer vision prompted the application of medical imaging analysis basedon deep learning [6,7].

Many experts have summed up, commented and discussed the research status and problems of DL based in medical imaging analysis. Lately, the review comprehensively summarized the research of deep learning in medical imaging classification, detection, registration.

Section 2 briefly introduces the basic principles of deep learning model, especially various popular neural networks in deep learning. In Section 3, the many applications in deep learning based medical imaging analysis are systematically reviewed. More importantly, we discuss research challenges and present possible future research interests in Section 4.

II. DEEP LEARNING ARCHITECTURE

Basic principles of DL model are introduced, which have been utilized far and wide in medical imaging analysis. Various popular neural networks in learning models are briefly introduced. These models include convolutional neural networks, stacked auto-encoders, recurrent neural networks and graph convolutional networks and so on.

2.1 Stacked Auto-Encoders

Firstly, an auto-encoder (AE) is a basic deep feed-forward network including an input layer, a hidden layer and an output layer. AE can learn latent patterns of input data through an unsupervised learning, and then reconstruct output data using the latent patterns. According to different functions, an AE includes encoder and decoder. The former ($f(x)$) generates a reducedfeature representation from an initial input x through a hidden layer h, and the latter ($g(f(x))$) reconstructs the input from the output of the former by minimizing the loss function:

$$L(x, g(f(x))) \quad (1)$$

Converting high-dimensional data to low-dimensional data through encoder and decoder. The representation power is very limited because of the shallow architecture. However, the representation power of a Stacked Auto-Encoders (SAE) stacked by multiple AE, can be very improved. Moreover, architecture of SAE is shown below in Fig 1. Due to the structural characteristics of SAE, it can discover and learn more complex potential patterns in the input data. Simple latent patterns can be extracted from the lower layer, and more complex latent patterns can be extracted from the higher layer. In short, different levels of data information can be represented in different layers of SAE.

DOI : https://doi.org/10.56452/7-3-109

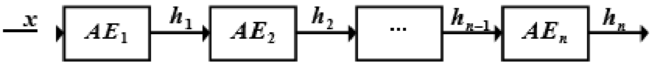

Fig 1: a stacked auto-encoder with *n* auto-encoders

2.2 Convolutional Neural Networks (CNNs)

CNNs were proposed to learn latent features from input data. CNNs have two most characteristics, one is translation, and the other is rotation. Translation and rotation in image are very important in radiology. Therefore, these characteristics make CNNs very popular in fields where an object shape is an important feature. LeNet, the most typical CNNs, consist of convolutional layers, pooling layers (also known as subsampling) and full connection layers, andshown in Fig 2.

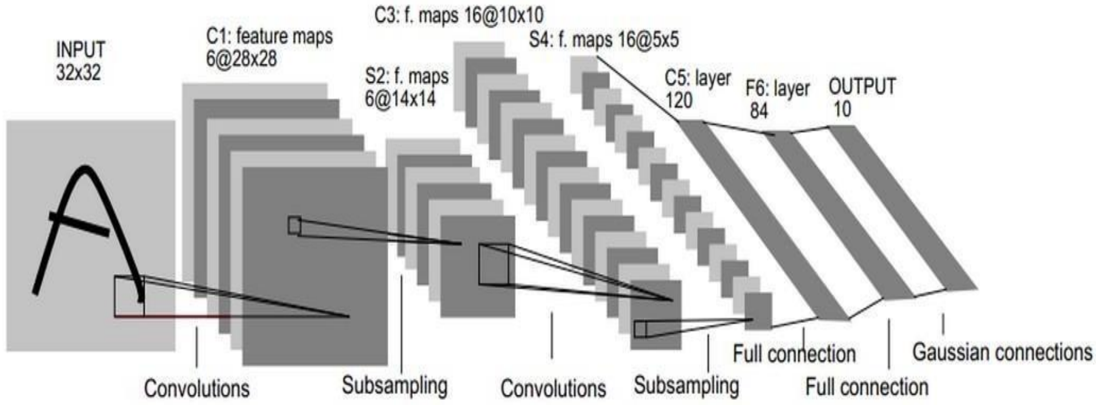

Fig 2: convolutional neural network

Convolution layer is utilized to extract latent features from input data, and the output is feature maps. There are multiple convolution kernels in each convolution layer. Each convolution kernel corresponds to a shared weight coefficient. In formal terms, let $I(\cdot,\cdot):\mathbb{R}^2 \to \mathbb{R}$ be the input value, depending on its position in the space, $W_k(\cdot,\cdot):\mathbb{R}^2 \to \mathbb{R}, k=1,\cdots,K$ a set of $K \in \mathbb{N}$ shared weight coefficient, named *kernels*; for every position $(x,y) \in \mathbb{R}^2$, the feature map $s_k(x,y)$ is given by

$$s_k(x,y) := \sum_m \sum_n I(x-m, y-n) W_k(m,n). \tag{2}$$

The feature maps are computed repeat until the entire image is covered, then becomes inputs of the next layer in CNNs architecture.

A **pooling layer** usually receives feature maps above as input and produces a popular statistical index as output. The main characteristics of pooling layer include translation and rotation. There are fewer parameters to be estimated in CNNs models through pooling operation, simultaneously the dimensions of the input is reduced. Therefore, pooling layers not only avoid over-fitting model training but also accelerate training. Most used in pooling operation includes max-pooling, average-pooling and Lp normalization pooling.

Full connection layer is the final layer in convolutional neural networks. For the sake of improving performance of CNNs network, *ReLU* function is usually utilized as activation function of each neuron in the full connection layer, which not only accelerates calculations, butalso avoids vanishing gradient issue. Formally the *ReLU* function is given below:

$$f(x) = max(0, x) \tag{3}$$

where x is the input. Additionally, other activation functions including *tanh, sigmoid* can be used. The output of last full connection layer is usually used as input of a classifier for classification.

There are lots of popular architectures of CNNs, including LeNet, AlexNet, DPN, GoogleNet, Inception ResNet, ResNet, SENet, VGGNet, and so on. Due to the space, we will not introduce them one by one.

2.3 Recurrent Neural Networks (RNNs)

RNNs were proposed, which have usually been utilized in analyzing sequential and time-series data, past data is implicitly stored in 'memory units'. In plain RNNs, the output of the currentsequential input is computed by all previous input data.

In medical imaging analysis field, RNNs are usually used in image segmentation. Image segmentation based on full convolutional network and RNNs was proposed, which is the first framework for 3D image segmentation. Stollenga et al. rearranged cuboid order of computationsin multi-dimensional long short term memory networks in pyramidal fashion, and segmented both 3D images such as electron microscope images and MRI brain images. A deep learning model was proposed, which was used to annotate disease name mined from an X-ray image using this model.

2.4 Graph Convolutional Networks (GCNs)

Graph convolutional networks were developed to analysis graph data based on CNNs, which is able to effectively handle complexity graph structure data by modeling relations between samples. Therefore, GCNs can be naturally used to model spatial relations for graph classificationtask. Mathematically, a graph defined as $G=(V, A)$ is a connected, undirected graph with real-

value weights on the edges, where V represents the vertex set including nodes $\{v_1, \dots v_n\}$, and

$A \in \mathbb{R}^{n \times n}$ is a symmetric adjacency matrix where a_{ij} denotes the weight coefficient between nodes v_i and v_j. A missing edge is represented through $a_{ij}=0$. In GCNs, the hidden states depend on the input and the connections among the nodes represented by the adjacency matrix. So for medicalimages, it needs to transform them to graph data in pre-processing, then the nodes and edges information of graph is regarded as input of GCNs. GCNs have been utilized in many applicationfields, including medical data analysis [8], applied chemistry, computer vision, citation networks,social networks, and natural language processing.

2.5 Open Source Toolkits in Deep Learning

DL is a complicated technology. Researchers and technician have to spend much energy and time to implement the abovementioned DL. Due to great applications based on DL technology, many famous companies and research groups have developed their open source toolkits in deep learning. Therefore, we are able to easily build deep learning models based on these toolkits, evenif we are not familiar with deep learning technology. In **TABLE I**, some commonly used and widely used open source toolkits for DL are shown; we should select the most suitable toolkit forour applications.

TABLE I. Some toolkits for DL

TOOLKITS	CREATOR	REFERENCE
Caffe	Berkeley Center	https://github.com/BVLC/caffe/
Chainer	Preferred Networks	https://chainer.org/
CNTK	Microsoft	https://github.com/Microsoft/CNTK
DL4J	Skymind	https://github.com/deeplearning4j/
Keras	Franois Systems	https://github.com/keras-team/keras
MXNet	Apache Software Foundation	https://github.com/apache/incubator-mxnet
PyTorch	Adam Paszke et al.	https://github.com/pytorch
SINGA	Apache Software Foundation	http://singa.incubator.apache.org/
TensorFlow	Google	https://github.com/tensorflow

III. APPLICATIONS OF DEEP LEARNING BASED MEDICAL IMAGING ANALYSIS

3.1 Deep Learning for Alzheimer's Disease (AD) Analysis

AD is the most common brain disorder of neurological, irreversible disease among older people, seriously affects daily life; even the life quality of patients. So far, scientists do not fully understand what causes AD. According to the estimation, the number of patients will double in the next 20 years, and by 2050 years one out of 85 peoples will have the disease [9]. With the increasing of patients with AD, the cost of nursing care for patients is rising rapidly. Therefore accurate diagnosis is very critical, particularly in early stage of diagnosis and treatment. For disease research, the ADNI is the best known open neuroimaging datasets. ANDI researchers collect, utilize and validate data, including blood biomarkers, cognitive tests, cerebrospinal fluid, genetics, MRI images and PET images as predictors of diseases for study. The ANDI dataset contains about more than one thousand patients. Those with mild cognitive impairment (MCI) easily cause more memory problems than normal for people with same age, and will develop to

AD.

Therefore, there are many papers about AD diagnosis at the early stage have been published using deep learning technologies. A feature representation based on deep learning with a SAE has been proposed [10]. With this feature representation, AD/MCI classification model has been proposed combining latent complicated low-level patterns, and gets higher accuracy. A novel deep learning based diagnostic framework has been developed utilizing a zero-masking method for diagnosis of AD, which can combine complementary information in multiple modals for data fusion in one setting. A robust deep learning framework has been presented, which prevented over-fitting and incorporated adaptive learning factor, multi-task learning strategies utilizing the dropout technique in deep learning training. A nonlinear learning model was proposed by Shi et al. [11], which fuse cross-sectional and longitudinal features learned from MR brain images utilizing stacked de-noising sparse AE. A novel framework was proposed by Lu et al. to identify and discriminate subjects at the MCI stage with presymptomatic AD from other subjects with MCI using FDG-PET metabolism imaging.

A new classification model was proposed to extract intra-slice features and inter-slice features utilizing combination of RNNs and 2D CNNs decomposed from the 3D PET images into 2D slices. In addition, the intra-slice features is learned from the hierarchical 2D CNNs, and the inter-slice features is learned from the gated recurrent unit (GRU) of RNNs cascaded, then these features are integrated and used in the final classification task. Because the spatial information in 3D brain images are discarded in the preprocessing step. There several 3D CNNs models directly using 3D brain images were proposed. Karasawa et al. proposed 3D CNNs model with about forty layers to improve performance, and the features are automatically learned from images without great effect of image pre-processing [12].

Recently, models based on GCN have been proposed, which combine medical images and demographic relationship for MCI classification tasks. A multi-class GCN classifier was implemented by Song et al. for classification of subjects including four categories: cognitively normal, early MCI, late MCI, and AD. GCN based PETNET model was proposed to analyze PETsignals by Guo et al.

3.2 Deep Learning for Pulmonary Nodule Disease Analysis

Pulmonary nodule is usually a round or oval spot on your lung. In the chest CT scans, 25% will appear nodules; More than 90% nodules are benign, not cancerous. But nodules need to be examined carefully because they may be a small cancer. The goal of the screening program is to detect small, curable cancers as early as possible. Almost 80 percent of small lung cancer patients are considered cured, and survive at least five years after diagnosis. Unfortunately, for patients with large lung cancer, the survival rate is low. Accurate and early detection is the key to cure patients.

Currently, Kaggle provides a large open data set, which attracts extensive interest from academia and industry. Therefore, a large number of DL based lesion detection algorithms have been proposed. Ding et al. proposed a novel CNNs based pulmonary nodule detection approach. They introduced a de-convolutional architecture to Faster Region-based CNNs for candidate

detection on axial slices. DeepLung was proposed in literature [13], which was automated system for lung CT cancer diagnosis. This system contains two 3D networks, one 3D Faster R-CNNs is implemented to effectively learn nodule features for nodule detection, and the other is designed for classification respectively. S4ND, a new lung nodule detection based on DL, is designed for detection using a single feed-forward without any further user guidance. The detection pipeline is implemented as a single 3D CNNs with dense connections. In literature [14], a novel hierarchical semantic convolutional neural network (HSCNN) is presented to predict whether a given pulmonary nodule is malignant or not. This network consists of low-level semantic features and a high-level prediction of nodule malignancy. Its advantage is that HSCNN can not only produce interpretable lung cancer prediction, but also obtain better results than 3D CNN alone.

3.3 Deep Learning for Breast Cancer Disease Analysis

Breast cancer is a malignant tumor occurring in the epithelial tissue of the breast. The incidence rate of breast cancer has been increasing since the end of 1970s, and that in China is increasing rapidly, which is 3% to 4% per year and 5 years survival rate is 73%. In the United States, breast cancer is diagnosed more than 80 percent, while that in China is less than 20%. Millions of people die of breast cancer every year. Therefore, for the treatment of breast cancer, early diagnosis is particularly important. An automatic disease detection system for malignant breast cancer in patients' images will help doctors or experts more effectively, and make the diagnosis more accurate, more scalable and less error prone. In area of breast cancer diagnosis, many scholars use deep learning method to do related research and have achieved good results. Spanhol et al. proposed an image patches based extraction method, which uses the combination of pre-training CNN and image patches for final classification. These breast cancer histopathological images are taken from Breast Cancer Histopathological Database (a publicly dataset available at http://web.inf.ufpr.br/vri/breast-cancer-database). DeCAF features learned using previous trained CNNs was used as input of a classifier; and these features also can be usedto develop high accuracy breast cancer diagnosis system. A model based on CNNs was proposedin literature for classification of breast biopsy images. The architecture is implemented to learn features from both nuclei and overall tissue organization, and the features extracted by the CNNsare also used in Support Vector Machine classifier. In paper [15], a novel CNNs including a convolution layer, full connection layer, and small SE-ResNet module is designed. Among them, the SE-ResNet module achieves similar performance with less parameter using combination of residual module and Squeeze-and-Excitation block.

IV. CHALLENGES AND FUTURE RESEARCH INTERESTS

DL is to learn all levels of abstract features in a data-driven way, which shows strong feature representation ability and robustness in many application fields. Although DL has shown excellent performance in computer vision tasks with natural images as analysis and processing objects, it is still a great challenge to successfully apply DL methods in field of medical imaginganalysis.

Firstly, compared with most medical images, natural images have higher spatial resolution,

contrast and visual features such as brightness, color, texture. However, there are only intensity values of some special signal in most medical images, and the ratio of signal-to-noise is usually very small. For most medical images, the anatomical structure of tissues/organs and the boundaries between lesion areas are not clear, and the texture difference is not significant. At the same time, it is significantly different due to individual differences and imaging principles. Therefore, it is more difficult to analyze medical image than to natural image.

Secondly, medical imaging in different modality can only reflect the specific anatomical and functional information of human body, so there are some limitations in various medical imaginganalysis methods, and each has its advantages and disadvantages. Moreover, different imaging equipment and image reconstruction methods have great differences. In clinical practice, it is usually necessary to use multi-modality medical imaging for auxiliary analysis and diagnosis. Asa result, medical image analysis is more complex and difficult than natural image.

Thirdly, for task of natural image classification, there are many large-scale sample datasets with manual labels for learning and training, such as CIFAR, ImageNet, MNIST. However, it isdifficult to get large scale sample datasets when DL is applied to medical imaging analysis. Especially, there is a large difference between lesion sample datasets need labeled by clinical experts; therefore the available datasets is relatively small.

Finally, the commercial application of clinical medical image based on deep learning system may have legal and ethical problems, because the performance and accuracy of diagnosis systemis very dependent on the quantity and quality of training data. At same time if the deep learning system is used in the radiology filed, which is independent of the supervision of radiologists, therewill be legal liability problems.

At present, in order to meet the challenges above, the following strategies can be adopted in the future deep learning applications.

First of all, to solve insufficient training samples in medical imaging analysis, data enhancement strategy should be adopted in the future. At the same time, transfer learning shouldbe adopted as another solution, which transfers the model trained on large scale sample datasets to model on small sample datasets for further training.

Secondly, the emergence of challenges and large public databases in the field of medical imaging has provided great help for the pre training. Then, network parameters are fine-tuned through the existing limited labeled data sets to improve the repeatability of the experiment and the generalization ability of the model.

Thirdly, an effective multimodality data fusion also has been a challenge in medical imaging analysis. Multimodality data reflect the morphology and structure of normal organs and tissues from different angles; therefore there is a strong complementarity between them. At the same time this complementarity will be benefit to the feature extraction for DL model. Previous studiesfor multimodality data fusion have two categories, including data level fusion and decision levelfusion. The former focuses on fusing data from different modalities and the latter focuses on assembling classifiers respectively. However, the deep neural network architectures allows the intermediate multimodal data fusion of learning representation, which provides a truly flexible

method for multimodal data fusion. Further research is that which level of fusion is optimal for the current problem.

Finally, with the construction of digital hospital and the establishment of medical big data center, clinical medical data such as electronic medical record and text report can also be used asthe supplement of medical imaging. Combined with the analysis of medical text data and electronic medical record by RNN, LSTM and CNN, it has technical support for medical imaginganalysis and medical information.

V. CONCLUSION

In conclusion, this review provides researchers with valuable insights for the applications of DL model in medical imaging analysis. We review the most recent studies on medical imaging analysis using DL and focus on three typical diseases analysis. More importantly, introduce potential challenges and future research interests.

REFERENCES

[1] Brody H. (2013) Medical imaging. Nature 502(7473): S81-S81.

[2] Heidenreich A, Desgrandschamps F, and Terrier F. (2002) Modern approach of diagnosis and management of acute flank pain: Review of all imaging modalities. European Urology 41(4): 351-362.

[3] Erickson BJ, Korfiatis P, Akkus Z, and Kline TL. (2017) Machine learning for medical imaging. RadioGraphics 37(2): 160-130.

[4] Wernick M, Yang Y, Brankov J, Yourganov G, and Strother S. (2010) Machine learning in medical imaging. IEEE Signal Processing Magazine 27(4): 25-38.

[5] Bengio Y, Lamblin P, Popovici D, Larochelle H, editors (2006) Greedy layer-wise training of deep networks. Proceedings of the Twentieth Annual Conference on Neural Information Processing Systems; Vancouver, British Columbia, Canada.

[6] Epistemological View: Data Ethics, Privacy Trust on Digital Platform, Harsh, R., Acharya, G., Chaudhary, S., 2018 IEEE International Conference on System, Computation, Automation and Networking, ICSCA 2018, 2018, 8541166

[7] Enhance the Data Security in Cloud Computing by Text Steganography, Sanghi, A., Chaudhary, S., Dave, M., Lecture Notes in Networks and Systemsthis link is disabled, 2018, 18, pp. 241–248

[8] Marinka Z, Monica A, Jure L. (2018) Modeling polypharmacy side effects with graph convolutionalnetworks. Bioinformatics 34(13): i457-i466.

[9] Barnes D, Yaffe K. (2011) The projected impact of risk factor reduction on alzheimer's disease prevalence. Alzheimer's and Dementia 7(4): S511.

[10] Suk HI, Shen D, editors. (2013) Deep learning-based feature representation for ad/mci classification. International Conference on Medical Image Computing and Computer-Assisted Intervention.