# Scope of Deep learning in medical image analysis: A survey

**Asim Rafiq Bhat Teaching Faculty NIT Srinagar**

***Abstract*— Deep learning algorithms, in particular convolutional neural networks, have rapidly become a methodology of choice for analyzing medical science images. This paper reviews the major deep learning concepts pertinent to medical image analysis and summarizes various contributions to the field, most of which appeared in the last year. Survey includes the use of deep learning for object detection, image classification, segmentation, registration, and other tasks. Brief overviews are provided of studies per application area: neuro, retinal, digital pathology, cardiac, and breast. We end with a summary of the present state-of-the-art, an acute discussion of open challenges and directions for future research.**

***Index Terms*— deep learning, convolutional neural networks, medical imaging, survey.**

## I. Introduction

With the advent of computers in early 1970's the analysis of images became popular. The growth followed from low level processing and mathematical models to rule based expert systems to machine learning or pattern recognition. The pattern recognition is still very powerful. The systems evolved from being completely developed by humans to being trained by powerful machines. Convolutional networks are the most successful models used by human researchers for designing such systems. The first successful application of LeNet was reported in 1998 for hand written digit recognition [1]. In computer vision the Deep CNN has become first choice of human researchers.

## II. Deep learning models

There are various models of machine learning that have been proposed over time. Neural networks form the crux of the most deep learning models. Networks have a neuron with respective activation functions and set of parameters called as weights and bias. The neural networks are single layered and multi-layered as well. The multi-layered neurons have been extensively used to learn complex and hidden features from raw hypermedia. The three models that have found extensive use in medical image analysis are explained below.

### *A. Convolutional neural networks*

This architecture varies from multi layered perceptron in that the weights are shared in such a way that convolution operation is performed on images. This way the model not needs to sense and learn indivual detectors for the same image at different locations of the images thus making it equivalent to the translation of the inputs. This way number of parameters is also drastically reduced.

### B. Auto Encoders and stacked AEs

This model reconstructs the input **x** on the output layer **x'**. The network has a hidden layer **h** whose dimensions are taken less than| **x**|. This way the data is projected onto lower dimensional space to learn hidden features latent in the input. Weight matrix is given by $\mathbf{W}_{x,h}$ and bias $\mathbf{b}_{x,h}$ from input layer to the hidden layer and $\mathbf{w}_{h.x'}$ with corresponding bias $\mathbf{b}_{h,x..}$ and $\boldsymbol{\alpha}(.)$ is a linear transfrorm function.

$$\mathbf{h} = \alpha\,(\mathbf{W}_{x,h}\mathbf{x} + \mathbf{b}_{x,h})$$

Sparsity constraints or regularization can also be employed to enhance feature discovery. If the hidden layer is same as input then the network simply learns the identity functions. Deep Auto encoders or Stacked AEs networks are formed by placing auto-encoder layers in stacked fashion.

### *C. Deep Belief Networks*

These are essentially SAEs where AE layers are replaced by Restricted Boltzmann machines. Training is again in supervised fashion.

RBMs are markov random field having a visible layer and hidden layer that carries the hidden features the connections between the nodes is bidirectional in that the hidden features can be obtained from input data and vice versa. The training proceeds in this fashion.

## III. Deep learning uses in medical science field

Application of deep learning can be in:

### *A. Classification*

This field of medical science involves detecting the presence of any disease from a given set x ray diagnosis images. An image is given with a label associated with it of the presence of disease or not. The training dataset is either employed for scratch learning or transfer learning is used with pre tra4ned models. The field evolved from being initially focussed on unsupervised learning. [2][3][4][5] Applied deep belief networks and stacked auto encoders to classify reported patients of Alzheimer's disease based on brain MRI (magnetic resonance imaging). Authors in [6] used CNN with pre trained models. Authors in [7][8] used 3d CNN rather 2d CNN to detect Alzheimer's disease. Authors in [9] used CNN architecture to brain identified graph detected from MRI DTI (diffusion tensor imaging). CNNs have been shown to adapt to implicit medical image features.

The other field of classification is lesion or object classification. In this area of medical science a part of object is classified rather than the whole object. Authors in [10][11] used multi stream CNN architectures to classify nodule and skin patches, [12] proposed a combination of RNN and CNN for classifying and ranking nuclear cataracts in slit-lamp images.

*B. Detection*

This field involves organ/ landmark localization. It involves parsing of 3d volumes. The deep learning solves the 3d parsing by assuming them as orthogonal 2d planes. Authors in [13] solved the parsing issue by processing three 2d planes and taking intersection with highest classification rate as the output. Authors in [14] used area of interest as the landmark around anatomical regions by defining a bounded 3d box after 2d parsing of 3d volume data.

Thus localization from 2d models based on deep CNNs has been identified as the most popular strategy to identify organs, landmarks and regions. The field is now moving for accurate localisation with healthy results. Recurrent neural networks are showing great promise of exploration in localisation filed and multi-dimensional RNNs are playing an important part in spatial field as well.

Object or area of interest detection has been one of predominant time consuming affairs in manual clinics. A lot of research has been done to aid such activities with the power of computing and processing to improve the detection accuracy. Early object detection model was based on four layered CNN to detect and identify nodules in x-ray images [14]. CNN is applied to perform voxel classification, and then some form of after-processing is performed for accurate object detection. Authors in [15] used 3d convolutional neural networks to search for micro-bleeds in brain MRI. Authors in [16] used supervised deep learning for detection of lesions in mammography and nodules in chest radiographs.

*C. Segmentation*

Segmentation allows the analysis of medical parameters like shape and volume in brain and or cardiac analysis. The objective of segmentation is to identify the set of pixels/voxels that make up the boundary or the interior of the area of interest. Segmentation has been the most common field that has been invariant deep learning methodologies being applied , including the design of Unique CNNs known as U-NET by [17].U-net architecture proposes equal number of up sampling and down sampling layers. In practical training scenario it means that an image can be parsed in a single pass to give segmentation map. Thus whole context of image can be taken into account rather than a single patch, in patch based CNNs. Although U-net works good for orthogonal 2d slices, a 3d model has also been proposed called as V-net with objective function based on dice coefficient. Some papers have been also published based on patch trained CNNs. Authors in [18] applied sliding window scan over electron microscopy imagery with patch based training rather single pass image parsing.

U-net also finds application in lesion based segmentation to take advantage of both local and global context. Methodology in [18] is also based on U-net with up and down sampling paths with no skip connections. Researchers in [19] also used U-net to segment white matter cuts in brain MRI with 3d CNN and one skip connection between first and layer CNN layer. Class imbalance is the major challenge in segmentation as most voxels are from regions of not interest. This problem is solved from incorporation of loss functions. Segmentation being the prominent in medical science analysis has seen custom architectures coming to forth tailored for this purpose only.

D. Other fields

These include *registration, content based image retrieval, Image generation and Enhancement, Combining image data with reports*.

Registration involves spatial alignment of medical science images in which coordinate translation is performed from one image to another. It is often achieved from recursive framework where a specific non-parametric translation is assumed and to be optimized is a pre-determined metric (e.g., L2-norm). Authors in [20] [21][22] used deep learning networks to measure the similarity between the two images to run an recursive optimization strategy. Authors in [23][24] used learning algorithms to directly regress the registration translation parameters given input images.

CBIR is a knowledge discovery technique and offers the possibility to understand rare disorders, same case histories to improve patient care. The major challenge is to extract workable features from pixel based representations. Authors in [25][26] used already trained CNNs to extract invariant features from medical imagery. The x-ray images were taken as the test set by the authors. The network was 5 layered fully connected CNN. This field has not seen as much success as of others as yet.

Image generation and enhancement involves removing noise, improving image quality, and pattern discovery and image normalization. Image generation usually uses 2d and 3d CNN models from image transformation. In [27] multi-stream CNN architecture reconstructed high quality cardiac MRI from low resolution multiple MRI images.

Combining image data with reports has seen the rise from recent caption generation work for natural images. The first work in field of medical image analysis was reported from [28]. They used supervised learning with caption or label for each image in the dataset to regress correct labels at test time. They also showed that classification accuracy is improved from the semantic information of labels.

## IV. Functional application areas

Deep learning finds its application in various medical imaging fields. These include brain, cardiology and eye. The contributions in these fields are briefly discussed.

### *A. Brain*

Belief networks have been widely employed for brain image analytics. The focus of the research has been on the Alzheimer's disease classification, brain tissue segmentation, segmentation and detection of lesions including tumors and micro-bleeds. Most methodologies have focussed on learning mappings from local patches to representation levels and then to captions. Although local patches have implicit problem of lack of contextual semantics. The brain images are 3d models and analysed orthogonally in 2d slices as discussed earlier. In [29] the authors have focussed on brain network estimation, Huntington's disease classification and Schizophrenia based on deep belief networks. In [30], the work based on CNN, tumor grading or 2d tumor patch classification was analysed with good results.

### *B. Eye*

The focus of work in this field has been to analyse fundus color images. The applications range from eye disease diagnosis, detection and segmentation of retinal abnormalities to image quality assessment. Authors in [31] performed the analysis of the inception V3 network from Google for diabetic retinopathy detection and performance was better than panel of seven certified eye specialists. Authors in [32] performed blood vessel segmentation. CNN along with CRF was used as an underlying model to model long range pixel interaction. The work was reformed in [33] with CRF being replaced by recurrent neural networks.

### *C. Digital pathology and microscopy*

This field is a popular application area for machine learning techniques due to the growing availability of large scale giga-pixel whole-slide images (WSI). Deep learning has been employed in normalisation of histopathology images which includes color normalization. In [34] deep sparse auto-encoders have been used for color stain normalization of haematoxylin and eosin (H & E). In [35] the importance of color normalization was demonstrated for CNN based tissue classification in H&E stain images. A challenge was recently held called TUPAC to address detection of mitosis in breast cancer tissue, and prediction of tumor grading at WSI level. Authors of [36] were best performers. There methodology had three main steps i) finding high density cell regions, ii) using CNN to detect mitosis in areas of interest and iii) conversion of results to vector for each WSI and using SVM classifier to compute molecular data scores and tumor proliferation.

### *D. Breast*

Since most breast imaging procedures have been 2 dimensional, all the natural image techniques can easily be employed with the exclusion of breast cancer detection. The problem with the field is that available datasets are small, therefore performance metrics vary. Authors in [37][38][39] applied semi-supervised learning, weakly supervised learning and transfer learning respectively to address the issue with little success. Authors in [40] employed CNN for direct classification of future risk of developing breast cancer based on negative mammograms.

### *E. Cardiac*

Deep learning has been employed for many dimensions of cardiac imagery analysis. MRI is the most analysed modality and segmentation of left ventricle is the most common task. Other tasks involve automated calcium scoring,, super-resolution, 2d slice classification and coronary centreline tracking. Work in [41] focussed on left ventricle segmentation. Deep belief network has been used to initialize a level set framework based on MRI. Authors in [42] used CNN for automatic generation of text descriptors for Doppler ultrasound images of heart vessels using doc2Vec.

Thus from the above summary it is obvious that deep learning is slowly prevailing each and every aspect of medical science imagery analysis and take over is happening extremely quickly. The current standard practice seems to be the use of end-to-end trained CNNs. In this approach a pre trained CNN is downloaded and handcrafted features could be easily extended over the network downloaded. The key aspects for success of deep learning in the field has been shown by data augmentation, a pre-processing step, that make networks robust and improved performance without the need to change network topology. Another aspect to be kept in consideration is input size to prevent context loss and last and obvious aspect is the hyper parameters (learning rate, dropout rate).

## V. Challenges and Future Work

There are obvious challenges for deep learning in medical imagery field. Lack of large dataset is the major obstacle. This involves data acquisition and relevant annotations for the said images. This requires help from domain experts which is cumbersome.

Secondly learning from large amount is time consuming because the images are available in 3d from radiologists and have to be sliced and slice wise annotations fixed. Thus learning from limited and small dataset is important area of research in medical science image analysis field.

Also, even if large raw dataset is available proper annotations or labelling requires expert opinion which is again cumbersome. The radio diagnosis persons analyse the images in the form of text reports. Converting them into appropriate annotations requires text-mining techniques.

In medical sciences the classification and segmentation is always seen as binary problems ie. Ill or not, normal or not, background vs object etc. This is unsophisticated simplification as the parameters may be heterogeneous. This approach leads to excluding sub classes but fails miserably to capture rare cases. This requires turning automation and deep learning systems into multi class models which again require detailed subclass labelling.

Another problem with data is data imbalance i.e., unavailability of test cases for a particular subclass. Some researchers have found data augmentation as a good solution to the problem but still a lot can be done and is an open field for research.

All of the challenges discussed above have not been fully addressed yet and open research areas but high success of deep learning has and is being reported. Going by the current trends unsupervised learning seems to be the field that receiving high amount of interest from researchers in medical science field, since it allows training to be from wealth of data without the requirement of annotations. Also unsupervised learning is pretty much analogous to humans and seems to be data efficient where humans learn the object labels without any knowledge of specific labels. Variational auto encoders and generative adversarial networks are the two strategies that are highly rated to be future unsupervised learning strategies of future.

## References

[1]. LeCun, Y., Bottou, L., Bengio, Y., Ha_ner, P., 1998. Gradient-based learning applied to document recognition. Proceedings of the IEEE86, 2278–2324.

[2]. Brosch, T., Tam, R., 2013. Manifold learning of brain MRIs by deep learning. In: Med Image Comput Comput Assist Interv. Vol. 8150

[3]. Plis, S. M., Hjelm, D. R., Salakhutdinov, R., Allen, E. A., Bockholt,H. J., Long, J. D., Johnson, H. J., Paulsen, J. S., Turner, J. A., Calhoun,V. D., 2014. Deep learning for neuroimaging: a validation study. Frontiers in Neuroscience.

[4]. Suk, H.-I., Shen, D., 2013. Deep learning-based feature representationfor AD/MCI classification. In: Med Image Comput Comput AssistInterv. Vol. 8150 of Lect Notes Comput Sci. pp. 583–590.

[5]. Suk, H.-I., Lee, S.-W., Shen, D., 2014. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. NeuroImage 101, 569–582.

[6]. Menegola, A., Fornaciali, M., Pires, R., Avila, S., Valle, E., 2016. Towardsautomated melanoma screening: Exploring transfer learning schemes. arXiv:1609.01228.

[7]. Hosseini-Asl, E., Gimel'farb, G., El-Baz, A., 2016. Alzheimer's disease diagnostics by a deeply supervised adaptable 3D convolutional network. arXiv:1607.00556.

[8]. Payan, A., Montana, G., 2015. Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks arXiv:1502.02506.

[9]. Kawahara, J., Brown, C. J., Miller, S. P., Booth, B. G., Chau, V.,Grunau, R. E., Zwicker, J. G., Hamarneh, G., 2016b. Brain-NetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment. NeuroImage.

[10]. Shen, W., Zhou, M., Yang, F., Yang, C., Tian, J., 2015b. Multi-scale convolutional neural networks for lung nodule classification. In: Inf Process Med Imaging. Vol. 9123 of Lect Notes Comput Sci. pp. 588–599.

[11]. Kawahara, J., Hamarneh, G., 2016. Multi-resolution-tract CNN with hybrid pretrained and skin-lesion trained layers. In: Machine Learning in Medical Imaging. Vol. 10019 of Lect Notes Comput Sci. pp. 164–171.

[12]. Gao, X., Lin, S., Wong, T. Y., 2015. Automatic feature learning to grade nuclear cataracts based on deep learning. IEEE TransBiomed Eng 62 (11), 2693–2701.

[13]. Yang, D., Zhang, S., Yan, Z., Tan, C., Li, K., Metaxas, D., 2015. Automated anatomical landmark detection on distal femur surface using convolutional neural network. In: IEEE Int Symp Biomedical Imaging. pp. 17–21.

[14]. Lo, S.-C., Lou, S.-L., Lin, J.-S., Freedman, M. T., Chien, M. V., Mun, S. K., 1995. Artificial convolution neural network techniques and applications for lung nodule detection. IEEE Trans Med Imaging 14, 711–718.

[15]. Dou, Q., Chen, H., Yu, L., Zhao, L., Qin, J., Wang, D., Mok, V. C., Shi, L., Heng, P.-A., 2016c. Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks. IEEE Trans Med Imaging 35, 1182–1195.

[16]. Hwang, S., Kim, H., 2016. Self-transfer learning for fully weakly supervised object localization. arXiv:1602.01625.

[17]. Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: Med Image Comput Comput Assist Interv. Vol. 9351 of Lect Notes Comput Sci. pp. 234–241.

[18]. Wang, C., Yan, X., Smith, M., Kochhar, K., Rubin, M.,Warren, S. M., Wrobel, J., Lee, H., 2015. A unified framework for automatic wound segmentation and analysis with deep convolutional neural networks. In: Conf Proc IEEE Eng Med Biol Soc. pp. 2415–2418.

[19]. Brosch, T., Tang, L. Y., Yoo, Y., Li, D. K., Traboulsee, A., Tam, R.,2016. Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to Multiple Sclerosis lesion segmentation. IEEE Trans Med Imaging 35 (5), 1229–1239.

[20]. Wu, G., Kim, M., Wang, Q., Gao, Y., Liao, S., Shen, D., 2013. Unsupervised deep feature learning for deformable registration of MR brain images. In: Med Image Comput Comput Assist Interv. Vol. 8150 of Lect Notes Comput Sci. pp. 649–656.

[21]. Simonovsky, M., Guti´errez-Becker, B., Mateus, D., Navab, N., Komodakis,N., 2016. A deep metric for multimodal registration. In:Med Image Comput Comput Assist Interv. Vol. 9902 of Lect Notes Comput Sci. pp. 10–18.

[22]. Cheng, X., Zhang, L., Zheng, Y., 2015. Deep similarity learning for multimodal medical images. Computer Methods in Biomechanics and Biomedical Engineering, 1–5.

[23]. Miao, S., Wang, Z. J., Liao, R., 2016. A CNN regression approach for real-time 2D/3D registration. IEEE Trans Med Imaging 35 (5), 1352–1363.

[24]. Yang, X., Kwitt, R., Niethammer, M., 2016d. Fast predictive image registration. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci.pp. 48–57.

[25]. Anavi, Y., Kogan, I., Gelbart, E., Geva, O., Greenspan, H., 2016. Visualizing and enhancing a deep learning framework using patients age and gender for chest X-ray image retrieval. In: Medical Imaging. Vol. 9785 of Proceedings of the SPIE. p. 978510.

[26]. Liu, X., Tizhoosh, H. R., Kofman, J., 2016b. Generating binary tags for fast medical image retrieval based on convolutional nets and Radon transform. In: International Joint Conference on Neural Networks. ArXiv:1604.04676.

[27]. Oktay, O., Bai, W., Lee, M., Guerrero, R., Kamnitsas, K., Caballero, J., Marvao, A., Cook, S., O'Regan, D., Rueckert, D., 2016. Multiinput cardiac image super-resolution using convolutional neural networks. In: Med Image Comput Comput Assist Interv. Vol. 9902 of Lect Notes Comput Sci. pp. 246–254.

[28]. Schlegl, T., Waldstein, S. M., Vogl, W.-D., Schmidt-Erfurth, U., Langs, G., 2015. Predicting semantic descriptions from medical images with convolutional neural networks. In: Inf Process Med Imaging. Vol. 9123 of Lect Notes Comput Sci. pp. 437–448.

[29]. Plis, S. M., Hjelm, D. R., Salakhutdinov, R., Allen, E. A., Bockholt, H. J., Long, J. D., Johnson, H. J., Paulsen, J. S., Turner, J. A., Calhoun, V. D., 2014. Deep learning for neuroimaging: a validation study. Frontiers in Neuroscience.

[30]. Pan, Y., Huang, W., Lin, Z., Zhu, W., Zhou, J., Wong, J., Ding, Z., 2015. Brain tumor grading based on neural networks and convolutional neural networks. In: Conf Proc IEEE Eng Med Biol Soc. pp.699-702

[31]. Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, Cuadros, J., Kim, R., Raman, R., Nelson, P. C., Mega, J. L., Webster, D. R., Dec. 2016. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316, 2402–2410.

[32]. Fu, H., Xu, Y., Wong, D. W. K., Liu, J., 2016b. Retinal vessel segmentation via deep learning network and fully-connected conditional random fields. In: IEEE Int Symp Biomedical Imaging. pp. 698–701.

[33]. Fu, H., Xu, Y., Lin, S., Kee Wong, D. W., Liu, J., 2016a. Deepvessel: Retinal vessel segmentation via?deep learning and conditional random?field. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 132–139.

[34]. Janowczyk, A., Basavanhally, A., Madabhushi, A., 2016a. Stain normalization using sparse autoencoders (StaNoSA): Application to digital pathology. Comput Med Imaging Graph, in press.

[35]. Sethi, A., Sha, L., Vahadane, A. R., Deaton, R. J., Kumar, N., Macias, V., Gann, P. H., 2016. Empirical comparison of color normalization methods for epithelial-stromal classification in H and E images. J Pathol Inform 7, 17.

[36]. Paeng, K., Hwang, S., Park, S., Kim, M., Kim, S., 2016. A unified framework for tumor proliferation score prediction in breast histopathology. arXiv:1612.07180.

[37]. Sun, W., Tseng, T.-L. B., Zhang, J., Qian, W., 2016a. Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data. Comput Med Imaging Graph.

[38]. Hwang, S., Kim, H., 2016. Self-transfer learning for fully weakly supervised object localization. arXiv:1602.01625.

[39]. Kooi, T., van Ginneken, B., Karssemeijer, N., den Heeten, A., 2017. Discriminating solitary cysts from soft tissue lesions in mammography using a pretrained deep convolutional neural network. Medical Physics.

[40]. Qiu, Y., Wang, Y., Yan, S., Tan, M., Cheng, S., Liu, H., Zheng, B., 2016. An initial investigation on developing a new method to predict short-term breast cancer risk based on deep learning technology. In: Medical Imaging. Vol. 9785 of Proceedings of the SPIE. p. 978521.

[41]. Ngo, T. A., Lu, Z., Carneiro, G., 2017. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. Med Image Anal 35, 159–171.

[42]. Moradi, M., Guo, Y., Gur, Y., Negahdar, M., Syeda-Mahmood, T., 2016a. A cross-modality neural network transform for semiautomatic medical image annotation. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp.300–307.

**Asim Rafiq Bhat**: The author has completed M.Tech in CSE from SMVDU katra India. The author is currently working as AP in NIT Srinagar India. The areas of research are Deep Learning and IOT.