International Journal of Scientific Research and Engineering Development-– Volume 4 Issue 1, Jan-Feb 2021
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Image Segmentation and Recognition Based on Deep Learning

David Taylor[1],Emily Smith[1]

1.LimeStoneUniversity, South Carolina,USA

Abstract: Image segmentation these days have gained lot of interestfor the researchers of computer vision and machine learning. Deep Learning techniques have achieved accuracy in many fields like medicine, automobiles- self driving, indoor navigation etc., We discuss first about what is deep learning and how it has evolved, and how it is helpful in different fields and convolutional neural networks, architectures, and their related calculations of accuracy. We also discuss about some research paper and their achieved accuracy in different fields of medicine diagnosis using Deep learning techniques.

Introduction:

Deep learning is subset of machine learning, and machine learning is a subset of artificial intelligence. Artificial intelligence is a technique which will let computers behave like humans. Machine learning is a set of algorithms that is trained on data from different sources and will enable computers behave like this. Deep learning is also a form of machine learning which is taken as inspiration from human brain. Deep learning continuously analyzes data with the given logics and inorder to complete this process deep learning used a set of algorithms called neural networks.Neural network's design is based on human brain structure. Our brain will identify different objects and different information same way neural networks

Convolutional Neural networks:

Visual systems structure is taken into inspiration for convolutional Neural Networks (CNNs). CNN is a kind of artificial neural network and are proved immensely powerful [20]. Every ANN will have three layers input layer, hidden layer and output layer. To the input layer image is fed as input and the hidden layer is used for training the architecture based on input images that are fed, and the output layer gives the output. CNNs will have two components, out of which one component will extract the features and other component will do the classification [21]. ANN learn by using weights, by adjusting the weights ANN will decide what to pass to the next signal. While training the ANN we can control the weights. Activation function is also applied to the neural network and it will decide based on the conditions whether to pass the signal along or not to the next layer. There are different models of CNNs. Based on local connectivity which is between neurons and on image transformations in a computational model are found in Neural networks. This specifies that Neurons which has same parameters are applied on previous layer patches at multiple locations. CNN contains three neural layers.

- Convolutional Layers
- Pooling Layers
- Fully connected layers

Each Layer is used for different purposes. Each Layer of CNN takes the input and gives output which will be input for next layer, which will have fully connected layers. CNNs are successfully used these days successfully in vision applications using computer vision applications, like image recognition, and detecting objects, robotics, and in self driving cars. CNNs have been extensively used in medicine for skin cancer classification [9].

Convolutional Layers: CNN uses various kernels to convert the whole image and also feature maps that are intermediate, which generates several feature maps in a CNN.As this can be used as advantage of convolutional layer it is replaced for fully connected layer

Pooling Layers: Spatial volume of input is reduced by pooling layers for the next convolutional layer.The depth is not affected by pooling. This process is called down sampling as the reduction in size leadsto loss of information. But this is not considered as disadvantage for this layer, but it is considered as advantage as decrease in size is considered beneficial as it is less computational overhead for the upcoming layers of network.Max pooling and Average pooling are commonly used.[26] Max pooling could lead to faster convergence and there is an improvement will be seen in generalization.

Fully Connected Layers: Fully connected layers perform high level reasoning, following convolutional and pooling layers. Fully connected layers neurons have connections fully to all the activation in the previous layers.

There are different layers in fully connected layers.

- **Fully Connected Input Layer:** This layer will take the output of the layer before it and fallen in and turn them into single vector and this will be input for next layer.
- **The first fully connected layer:**This layer will take input from the analysis of features and applies weights to identify the correct label.
- **Fully Connected Output Layer:**This Layer will give the final probabilities for each layer.

Artificial Neural Networks work based on activation functions: Activation function is nothing but mathematical equation that will determine the output of a neural network. This function determines if it should be activated or not and is attached to each neuron and works based on input of each neuron. Activation will help normalize each neuron output between range 1 an 0 or -1 to 1. Another important factor about activation function is that they need to perform computationally efficient because they are calculated across thousands or more neurons of each sample. A technique called backpropagation is used by neural network to train the model. which increases computational strain on the activation function. The necessity of speed led to development of new functions like ReLu and Swish.

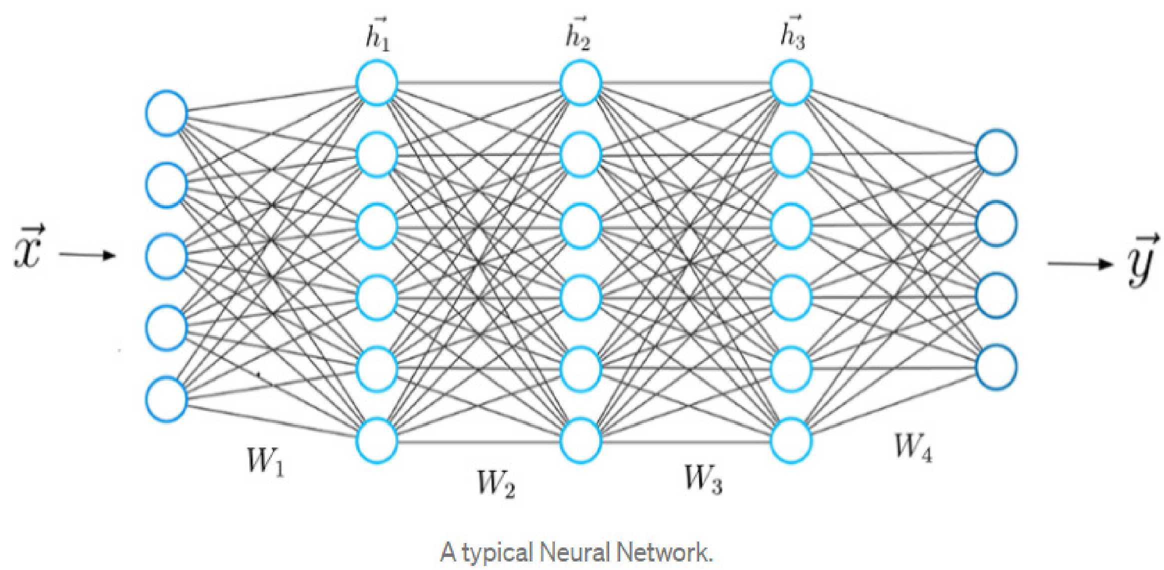


Fig.1: A typical Neural Network [1]

Activation Functions and its role in Neural Networks:

To the input layer in neural network inputs are fed which are numeric data points. Multiplying input with the weight of the neuron gives output of the neuron, and this output will be fed as input to next layer.so activation function acts like a mathematical gate between the input fed to the neuron and the output of the neuron. It can turn the neuron output off or on when its a step function depending on a rule or threshold. It can also map the input to the output and acts as a transformation. The representation of activation function is done by using g. Some activation functions are:

Softplus: Softplus represents smooth approximation of ReLU, which is one of variants of ReLU.

$$g(a) = \log(1+e^{a})$$

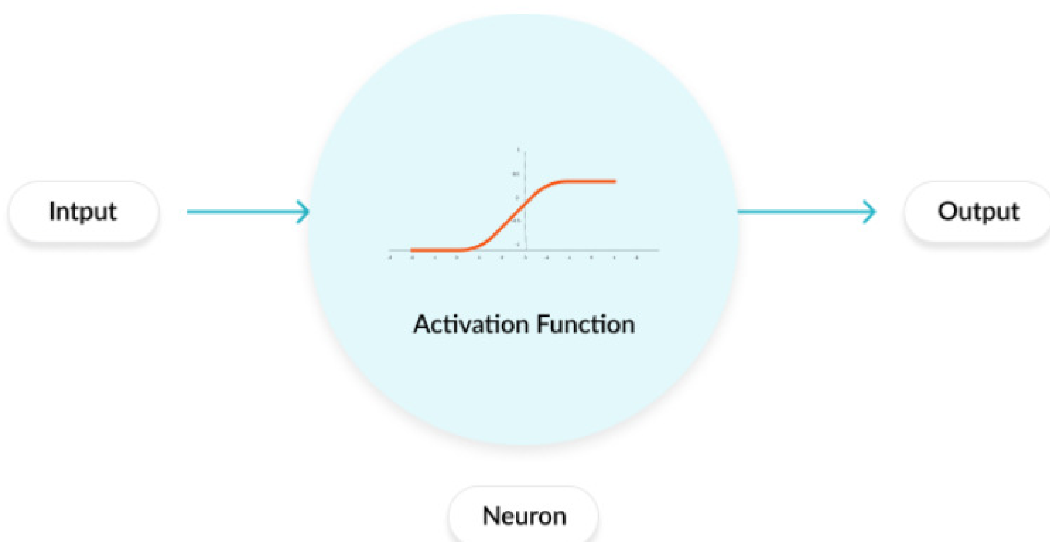
Rectified linear unit (ReLU): Superior Performance is shown by this activation function in many cases,and gradient diffusion problem can be solved by ReLU [7,8]. Some of the most popular deep learning activation functions are [5,9,10,11].

$$g(a) = \max(0, a)$$

Maxout: In this function the weight matrix is a three-dimensional array, and the third array corresponds to the connection of neighboring layers [12].

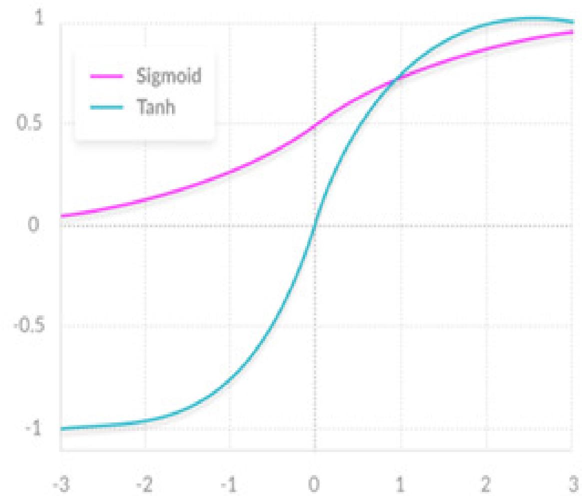
Threshold Function: This function passes 0, if the sum of the values of the inputs reach a certain threshold. It will pass 1 it the sum value is more than 0. This is like a step function as shown in the below image. It is a yes or no function and is straight forward.

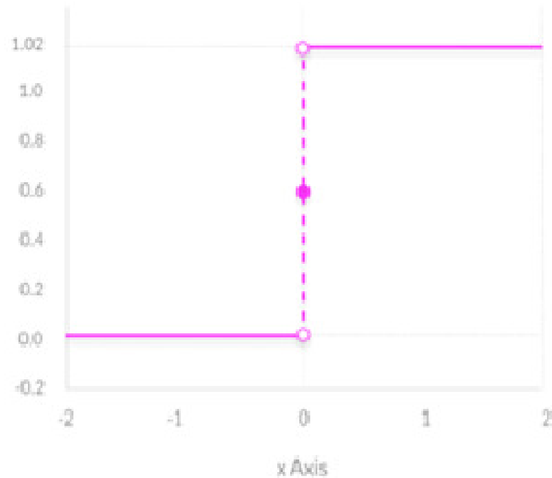
Softmax: The output of Softmax, is usually used in the last layer which is considered as a probability distribution over the categories.



There are different types of activation functions.

- Binary Step Function: Binary step function is an activation function based on threshold. Based on threshold value if the input value is above that threshold, the neuron will get activated and sends the value signal to next layer. The problem with that is it does not allow multi value outputs, which means it cannot classify the inputs based on several categories or requirements.





Frequently used Deep Neural Network Architectures

Deep neural networks has made significant contributions to several fields. AlexNet, VGG-16, GoogleNet, and ResNet are some of the commonly used deep network architectures.

AlexNet:AlexNet was one of the top deep CNN architectures that had accuracy rates of 84.6% compared to the one which uses traditionaltechniques instead of deep architectures and this one has accuracy rate of 73.8% when both are performed on same challenge.[2] represented an architecture.

VGG:Visual Geometry group fromOxford University has introduced VGG CNN. These authors of [15] have proposed various models and also configurations of CNNs.One of these modelsis used for a challenge on visual recognition in 2013 [3]. That model is VGG-16 as it has 16 weight layers and achieved an accuracy of 92.7%.

GoogleNet: Authors [4] has introduced GoogleNetit has an accuracy of 93.3% in an ILSVRC -2014 challenge. This CNN architecture consists of 22 layersand a newly introduced building block called inception module. Rather than a typical sequential manner CNN layers can be stacked in more ways. These modules contain Network in Network layer, pooling operation,convolution layer, and a convolution layer of small size.

ResNet:ResNet was introduced by Microsoft and it has an accuracy rate of 96.4% in ILSVRC-2016.This CNN architecture is well known due to its depthwhich has 152 layers and design of residual blocks was introduced with this. These residual blocks address the issue of training a deep architecture by using identity skip connections,in order for layers to copy inputs to next layer.The idea behind this model is, the next layer learns something new from the input, which is already encoded, and this also helps overcome the vanishing gradient problems.

ReNet:Authors [6] proposed Multi-dimensional recurrent Neural Network with an idea to extend Recurrent Neural Networks (RNNs) to multi-dimensional tasks which will replace each single recurrent connection from standard RNN with connections d, where the number of spatial dimensions is "d". Author [7] using this approachproposed ReNet, which instead of using multi-dimensional RNNs they use sequence RNNs.

Performance metrics of these neural network architectures is calculated by following metrics.

3.2 Performance Metrics:

Model Evaluation is performed by below following metrics [8,9].

Specificity: It gives what fraction of all negative classes are correctly identified as negative my model classifier. Also called True Negative Rate.

SP = TN/ (TN + FP)

Sensitivity: It is the ability of the process to correctly identify the disease condition or situation.

SE = TP/ (TP + FN)

ROC AUC: Area under receiver operating characteristic. It is the probability that the classifier will identify TPR against FPR. It is the graph between true positive rate vs false positive rate.

Precision: It gives what fraction of all classes that are correctly predicted positive, are actually positive.

PREC = TP / (TP + FP)

Negative Predictive Value (NPV): It is the probability of the inputs that tested negative, truly does not have the disease.

NPV = TN / (TN + FN)

Positive Predictive Value (PPV): It is the probability of the classes that are tested positive, truly have the disease.

PPV = TP/ (TP + FP)

Dice coefficient: It gives the overlap measure between the automatic and the ground truth segmentation. It is also called as overlap index.

DC = 2TP/ (2.TP + FN + FP)

Where

- True negative (TN): negative class is correctly predicted by the classifier model
- True positive (TP): positive class is correctly predicted by the classifier model
- False negative (FN): negative class is incorrectly predicted by the classifier model
- False positive (FP): positive class is incorrectly predicted by the classifier model

Use of Deep learning in image segmentation and Recognition:

Deep learning has been used in various fields to prevent human intervention and for better accuracy in image segmentation and classification. For example, in medicinal field it has been very helpful in different departments, that can easily avoid human mistakes in manually determining underlying issue vs with deep learning. In medical field it is used for Brain MR segmentation [10], segmentation of knee cartilage [11], segmentation of brain tumor[12], classification of skin cancer[9], detection of breast cancer[13], detection of skin lesions [14].

Author [15] collected the dataset from T1w in order to create algorithm to segment brain and there is four sections in the architecture fully connected section, section for input, section for output, and the last one is classification section. Algorithm is validated using 240 images from a hospital in Texas, USA. And they calculated DSI values and the accuracy was predicted as 0.75. Segmentation in the head and neck images is quiet difficult processing and also takes time. If the node lesion is large the anatomy of the site will be significantly affected. In head and neck cancer patients ,use of CNN speeds up the process and improves the accuracy. Author [16] has used images of 50 patients for neck and head radiotherapy and used a CNN with has multiple layers of drop out layer, normalization layer, and pooling unit. Another strategy similar to above has been proposed by [17] on the patients who are undergoing stereotactic treatments and used a 3D CNN to segment and train brains for patients.

Another one is pelvis segmentation, and for the segmentation several strategies has been implemented and an efficient technique using deep learning has been implemented by [18]. Abdomen segmentation is effectively done by deep learning techniques by author [19] , who is first to attempt this segmentation and 30 patients are used who are affected with lung cancer. And a DSI of 0.69 to 0.9 is produced as an accuracy.

Conclusion: Deep learning has a huge impact on many fields. The accuracy of these techniques have been improving continuously with the development of new architectures. We focused on what deep leaning is along with convolutional neural networks and artificial neural networks. Different formulas to calculate the accuracy and precision of a neural network is provided in this paper. There are few challenges associated with deep learning techniques, researchers are continuously trying to overcome these, and a successful event will help us better use of Deep learning techniques.

References:

1. **https://towardsdatascience.com/what-is-deep-learning-and-how-does-it-work-2ce44bb692ac**
2. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097–1105
3. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.
4. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9
5. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9
6. A. Graves, S. Fernandez, and J. Schmidhuber, "Multi- ´ dimensional recurrent neural networks," CoRR, vol. abs/0705.2011, 2007. [Online]. Available: http://arxiv.org/ abs/0705.2011
7. F. Visin, K. Kastner, K. Cho, M. Matteucci, A. C. Courville, and Y. Bengio, "Renet: A recurrent neural network based alternative to convolutional networks," CoRR, vol. abs/1505.00393, 2015. [Online]. Available: http://arxiv.org/abs/1505.00393
8. https://towardsdatascience.com/confusion-matrix-for-your-multi-class-machine-learning-model-ff9aa3bf7826
9. Ravi Manne,SnigdhaKantheti,Sneha Kantheti,(2020),"Classification of Skin cancer using deep learning, ConvolutionalNeural Networks - Opportunities and vulnerabilities- A systematic Review", International Journal for Modern Trends in Science and Technology,ISSN: 2455-3778,Vol. 06,Issue 11,pp. 101-108.https://doi.org/10.46501/IJMTST061118
10. 14,15,13,16,19,18 P. Moeskops, M.A. Viergever, A.M. Mendrik, L.S. de Vries, M.J.N.L. Benders, I. Isgum Automatic segmentation of MR brain images with a convolutional neural network IEEE Trans Med Imaging, 35 (2016), pp. 1252-1261
11. A. Prasoon, K. Petersen, C. Igel, F. Lauze, E. Dam, M. NielsenDeep feature learning for knee cartilage segmentation using a triplanar convolutional neural network Med Image ComputComput Assist Interv, 8150 (2013), pp. 246-253
12. S. Pereira, A. Pinto, V. Alves, C.A. SilvaBrain Tumor segmentation using convolutional neural networks in MRI images IEEE Trans Med Imaging, 35 (2016), pp. 1240-1251
13. M. Kallenberg, K. Petersen, M. Nielsen, A.Y. Ng, P.F. Diao, C. Igel, et al.Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring IEEE Trans Med Imaging, 35 (2016), pp. 1322-1331

14. T. Kooi, G. Litjens, B. van Ginneken, A. Gubern-Mérida, C.I. Sánchez, R. Mann, et al.Large scale deep learning for computer aided detection of mammographic lesions Med Image Anal, 35 (2017), pp. 303-312
15. Liu Y, Stojadinovic S, Hrycushko B, Wardak Z, Lau S, Lu W, et al. A deep convolutional neural network-based automatic delineation strategy for multiple brain metastases stereotactic radiosurgery. PLoS ONE. (2017) 12:e0185844. doi: 10.1371/journal.pone.0185844
16. Ibragimov B, Xing L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. Med Phys. (2017) 44:547–57. doi: 10.1002/mp.12045
17. Charron O, Lallement A, Jarnet D, Noblet V, Clavier JB, Meyer P. Automatic detection and segmentation of brain metastases on multimodal MR images with a deep convolutional neural network. Comput Biol Med. (2018) 95:43–54. doi: 10.1016/j.compbiomed.2018.02.004
18. Gambacorta MA, Boldrini L, Valentini C, Dinapoli N, Mattiucci GC, Chiloiro G, et al. Automatic segmentation software in locally advanced rectal cancer: READY (REsearch program in Auto Delineation sYstem)-RECTAL 02: prospective study. Oncotarget. (2016) 7:42579–84. doi: 10.18632/oncotarget.9938
19. Ibragimov B, Toesca D, Chang D, Koong A, Xing L. Combining deep learning with anatomical analysis for segmentation of the portal vein for liver SBRT planning. Phys Med Biol. (2017) 62:8943–58. doi: 10.1088/1361-6560/aa9262
20. Boldrini L, Bibault J-E, Masciocchi C, Shen Y and Bittner M-I (2019) Deep Learning: A Review for the Radiation Oncologist. Front. Oncol. 9:977. doi: 10.3389/fonc.2019.00977.
21. Chensi Cao, Feng Liu, Hai Tan, Deshou Song, Wenjie Shu, Weizhong Li, Yiming Zhou, Xiaochen Bo, ZhiXie, Deep Learning and Its Applications in Biomedicine, Genomics, Proteomics &Bioinformatics,Volume 16, Issue 1,2018,Pages 17-32,ISSN 1672-0229,https://doi.org/10.1016/j.gpb.2017.07.003.

