AUTOMATIC OPTIMIZED CNN BASED COVID-19 LUNG INFECTION SEGMENTATION FROM CT IMAGES

S. Rajaseela
PG Scholar, Department of Electronics and Communications Engineering
Pandian Saraswathi Yadav Engineering College, Sivagangai, Tamil Nadu

Abstract

The Corona Virus disease 2019 (COVID-19) has become a global pandemic since the beginning of 2020. The disease has been regarded as a Public Health Emergency of International Concern (PHEIC) by the World Health Organization (HIO) and the end of January 2020. Automated detection of lung infections from computed tomography (CT) images offers a great potential to augment the traditional healthcare strategy for tackling COVID19. However, segmenting infected regions from CT slices faces several challenges, including high variation in infection characteristics, and low intensity contrast between infections and normal tissues. Further, collecting a large amount of data is impractical within a short time period, inhibiting the training of a deep model. To address these challenges, a novel Seg Net based Convolutional Neural Network is proposed to perform COVID-19 Lung Infection Segmentation that automatically identifies infected regions from chest CT slices. In CNN, a parallel partial decoder is used to aggregate the high-level features and generate a global map.

Keywords: *Deep Learning, Convolutional Neural Network (CNN), COVID-19 Segmentation, SegNet*

Introduction

Computer Vision is a field of science that visualizes and understands the visual world using the concepts of image processing and machine learning. A computer vision technique tries to match the level of human perception in visual object recognition. However the system should be robust against noise, unwanted backgrounds, occlusion, lighting conditions, etc. Numerous machine learning techniques and algorithms are proposed for the task of image recognition. Since low-cost imaging devices are wide popular in today's camera market and due to the availability of free datasets in the internet, computer vision applications are being developed in great numbers among which food image recognition has recently gained wide attention.

Medical imaging is the process of producing visible images of inner structures of the body for scientific and medicinal study and treatment as well as a visible view of the function of interior tissues. This process pursues the disorder identification and management. This process creates data bank of regular structure and function of the organs to make it easy to recognize the anomalies. This process includes both organic and radiological imaging which used electromagnetic energies (X-rays and gamma), sonography, magnetic, scopes, and thermal and isotope imaging. Medical imaging produces the images of the internal structures of the body without invasive procedures. Those images were produced using fast processors and due to conversion of the energies arithmetically and logically to signals [2]. Those signals later are converted to digital images. Those signals represent the different types of tissues inside the body.

Deep Learning is a subdivision of Machine Learning which is rising quickly over the recent times. Deep learning techniques facilitates automatic feature extraction (which is a statistical representation of images) from the images where the user is unworried about manual feature extraction and the deep learning architectures like CNN extracts the best representative features from the images on its own as shown in Fig 1. Convolutional Neural Network (CNN) is best known for their ability to recognize patterns present in images. Machine learning approaches needs to be fed with hand crafted features where different applications works well on different features and fmding the appropriate feature for a particular application is a time consuming process.

The World Health Organisation (WHO) has declared the corona virus disease 2019 (COVID-19) a pandemic. A global coordinated effort is needed to stop the further spread of the virus. Corona viruses are a family of viruses that cause illness such as respiratory diseases or gastrointestinal diseases. Lung CT scan is one way to diagnose the COVID disease.

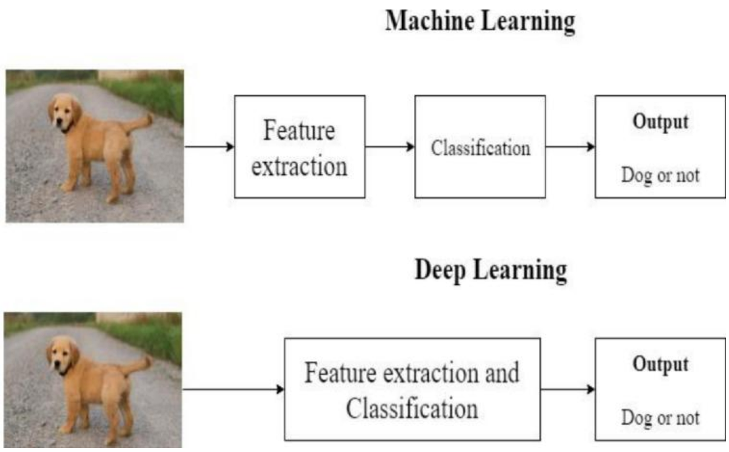

Figure 1 Machine learning Vs Deep learning

In this work, a novel and practical deep fully convolutional neural network architecture for semantic pixel-wise segmentation termed SegNet is proposed for the segmentation of COVID-19 infection in CT images. The paper is divided into the following sections: Section II explains the existing works related to this application. Section III briefs the working of CNN architecture for the image segmentation task. In section IV, the experimentation is explained. Section V gives the results and discussion and in Section VI, conclusion and the direction for the future work are given

Related Works

Semantic medical image segmentation are recently more popular using deep learning techniques [2] [3]. Florin C. et al proposed a pipeline for object detection and segmentation in the context of volumetric image parsing, solving a two-step learning problem: anatomical pose estimation and boundary delineation. Ilkay Oksuz et al discuss the implications of image motion artefacts on cardiac MR segmentation and compare a variety of approaches for jointly correcting for artefacts and segmenting the cardiac cavity. In this paper, they propose a segmentation network coupled with this in an end to-end framework. Subhankar Roy et al studies the application of DL techniques for the analysis of lung ultrasonography (LUS) images. A novel deep network, derived from Spatial

Transformer Networks, which simultaneously predicts the disease severity score associated to a input frame and provides localization of pathological artefacts in a weakly-supervised way is proposed. Ling Mang et al proposed a deep stacked transformation approach for domain generalization. Liang Sun et al proposed an anatomical attention guided deep learning framework for brain ROI segmentation of structural MR images, containing two sub networks. The first one is a segmentation sub network, used to simultaneously extract discriminative image representation and segment ROIs for each input MR image. The second one is an anatomical attention sub network, designed to capture the anatomical structure information of the brain from a set of labeled atlases. Zhe Guo et al propose an algorithmic architecture for supervised multimodal image analysis with cross-modality fusion at the feature learning level, classifier level, and decision-making level. Lua Ngo et al propose an automated segmentation method for OCT images based on a feature-learning regression network without human bias.

Convolutional Neural Networks

CNN is a multilayer neural network in which input to each layer is fed with results of the previous layer. CNN architecture consists of convolutional layer next to the input layer and followed by various other layers such as pooling layer, Relu layer, normalization layer, loss layer (drop out), fully connected layer etc. Convolutional layer is the first layer in CNN that performs automatic feature extraction using filters or kernels. Each filter performs convolution operation with the input image where each filter may learn some feature and the matrix obtained after convolution operation all through the image is called the feature map'. Each convolution layer is followed by a RELU (Rectified Linear Unit) layer which adds non-linearity to the model and this function **will** output the same input if it is positive, otherwise, it will **output** zero. To increase the performance and stability of a neural network further, normalization techniques are used that normalizes the output values of a previous feature map. Pooling layer generally reduces the spatial size of the feature map which reduces the number of parameters and the complexity of computation in the network. Dropout layer prevents the drawback called overfitting while training deep neural networks. The fully connected layer is supplied with values from the final pooling or fmal convolutional layer, where the values are vectorized and then fed as a vector into the fully connected layer. The fmal layer uses the softmax classifier which is used to get probabilities of each input belonging to a particular class.

SegNet is a convolutional neural network for semantic image segmentation. The network uses a pixel Classification Layer to predict the categorical label for every pixel in an input image. This core trainable segmentation engine consists of an encoder network, a corresponding decoder network followed by a pixel-wise classification layer. The architecture of the encoder network is topologically identical to the 13 convolutional layers in the VGG16 network. The role of the decoder network is to map the low resolution encoder feature maps to full input resolution feature maps for pixel-wise classification. The novelty of SegNet lies is in the manner in which the decoder upsamples its lower resolution input feature map(s). Specifically, the decoder uses pooling indices computed in the max-pooling step of the corresponding encoder to perform non-linear upsampling.

This eliminates the need for learning to upsample. The upsampled maps are sparse and are then convolved with trainable filters to produce dense feature maps. We compare our proposed architecture with the widely adopted FCN and also with the well known DeepLab-LargeFOV, DeconvNet architectures. This comparison reveals the memory versus accuracy trade-off involved in achieving good segmentation performance. SegNet was primarily motivated by scene understanding applications. Hence, it is designed to be efficient both in terms of memory and computational time during inference. It is also significantly smaller in the number of trainable parameters than other competing architectures. We show that SegNet provides good performance with competitive inference time and more efficient inference memory-wise as compared to other architectures. Implementing a CNN model requires significantly huge data to train the parameters of the network.

Experimentation

Data Set

The publically available "Covid-19" dataset from Kaggle is used for experimentation that consists of 21 food categories that Indian people loves [11]. This dataset has 20 CT scans with the expert segmentations of lung and Covid — 19 infection regions. The 20 CT scans have a total of 3250 2D slices in it. Among the 3250 slices, 2314 slices visualize the lung region and only 1630 slices has the COVID-19 infection region.

Proposed Methodology

Seg Net has an **encoder** network and a corresponding **decoder** network, followed by a final pixel wise classification layer. There are 13 convolutional layers from VGG-16. (The original fully connected layers are discarded.)At the encoder, convolutions and max pooling are performed. While doing 2x2 max pooling, the corresponding max pooling indices (locations) are stored. At the decoder, up sampling and convolutions are performed. At the end, there is soft max classifier for each pixel. During up sampling, the max pooling indices at the corresponding encoder layer are recalled to up sample as shown above. Finally, a K-class soft max classifier is used to predict the class for each pixel. One key ingredient of the SegNet is the use of max-pooling indices in the decoders to perform up sampling of low resolution feature maps. This has the important advantages of retaining high frequency details in the segmented images and also reducing the total number of trainable parameters in the decoders. The entire architecture can be trained end-to-end using stochastic gradient descent.

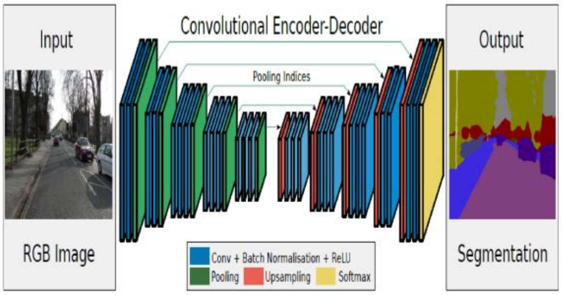

Fig 2: Proposed SegNet Architecture

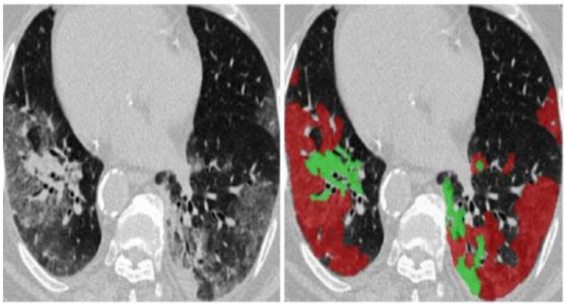

Figure 3 Example of COVID-19 infected regions (B) in CT axial slice (A), where the red and green masks denote the GGO and consolidation, respectively

Results and Discussion

The SegNet is trained on CPU, Intel 15 processor and on 8 GB RAM and the working platform is Matlab. 80% of the total images were used for training, 10% for validation and 10% for testing. The optimal network parameters are: number of epochs — 20, learning rate — 0.0001, mini batch size — 10, optimizer — SGDM (Stochastic Gradient Descent). Using the above parameters and the Covid-19 dataset, the network is trained using SegNet's architecture and has achieved a validation accuracy of 85%. In-order to improve the performance of the system, data augmentation is proposed

Data Augmentation

The performance of the CNN can be enhanced using data augmentation [12] technique. This is done because the deep neural network model has many parameters and the model should be trained with a huge amount of data so that the model is trained with the optimal parameters. For each original image, the images are horizontally flipped, vertically flipped and horizontally vertically flipped. The same 80%, 10%, 10% convention is used for training validation and testing the proposed model. Data augmentation improved the accuracy from 75% to 96.6% using the same SegNet architecture.

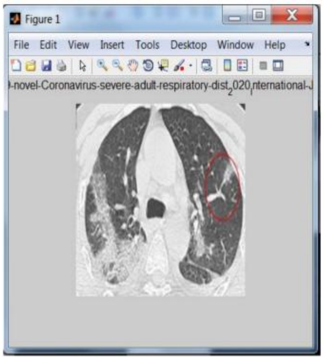

Fig 4: Input image

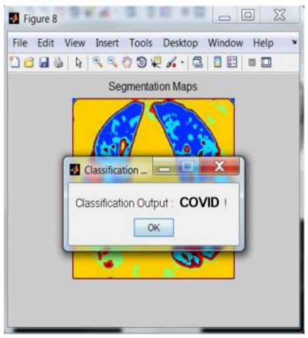

Fig 5: Layer processing and result

Table 1 Comparison of our results with existing works

	No of images	Epoch	Learning Rate	Accuracy
[4]	**60,000**	**10**	0.001	96.9%
[5]	5822	100	0.001	94%
Proposed work without Data Augmentation	1630	5	0.0001	85%
Proposed work with Data Augmentation	**6520**	**20**	**0.0001**	**96.6%**

While comparing the proposed work to the other existing works in the literature related to COVID-19 segmentation, our work has achieved a far good result. [2] has used around 60,000 images to train for 10 epochs and [3] has used 5822 images to train for 100 epochs which had been trained for a long time using CPUs and the proposed model has achieved a best accuracy of 96.6% with just 6520 images trained for 20 epochs where the performance of our system has been improved using data augmentation. The training time is also very less even while using a CPU. The testing results from the dataset are shown below in fig 4 and in fig 5.

Conclusion

Deep learning practices are an area where high scientific achievements are obtained in different scientific fields day by day. One of these fields is medical practices and studies such as disease detection, disease classification, and location of the disease are carried out. Dataset were performed as input data to the SegNet network using image processing techniques. The main motivation behind SegNet was the need to design an efficient architecture for road and indoor scene understanding which is efficient both in terms of memory and computational time. We analyzed SegNet and compared it with other important variants to reveal the practical trade-offs involved in designing architectures for segmentation, particularly training time, memory versus accuracy. The network, achieved higher accuracy. SegNet structure, which has been used less than other popular deep learning methods in previous studies, combined with image processing methods, has shown a successful result.

References

1. World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. Available from: https://www. who. int/dg/speeches/detail/who-director-general-s- opening-remarks-at-the-media-briefing-on-covid-19-11-march-2020 (Accessed 14 Marc 2020).
2. Marriam Webster Dictionary. Pandemic. Available from: https://www.merriam-webster.com/dictionary/pandemic (Accessed 14 March 2020).
3. World Health Organization. Novel Corona virus - China. Disease outbreak news : Update 12 January 2020.
4. Wikipedia. Timeline of the 2019-20 corona virus pandemic in November 2019 - January 2020.
5. World Health Organization. Director-General's remarks at the media briefing on 2019-nCoV on 11 February 2020.2020/2/18)[2020-02-21]. https://www. who. Int/dg/speeches/detail/who-director-generals-remarks-atthe-media -briefing -on-2019-ncov -on-11 -february-2020.2020.
6. Public Health England. COVID-19: epidemiology, virology and clinical features. Available from:https://www.gov.uk/government/publications/wuhan-novel-coronavirus-background-information/wuhan-novel-coronavirus-epidemiology-virology-and-clinical-features (Accessed 14 March 2020)
7. World Health Organisation. Novel coronavirus (2019-nCoV). Published on 31 January 2020. Available from: https://www.youtube.com/watch?v=mOV1aBVYKGA&t=88s [last accessed 16 March 2020]
8. World Health Organisation. WHO: Coronavirus - questions and answers (Q&A). Published on 16 January 2020. Available from:https://www.youtube.com/watch?v=0aRD9fV7jo&t=8s [last accessed 16 March 2020]
9. Osmosis. COVID-19 (Coronavirus disease 2019) March update- causes, symptoms, diagnosis, treatment, pathology. Published on 15 March 2020. Available from:https://www.youtube.com/watch?v=JKpVMivbTfg [last accessed 16 March 2020]
10. World Health Organization. Coronavirus. Available from:https://www.who.int/health-topics/coronavirus (Accessed 14 March 2020)
11. Oksuz, I., Clough, J. R., Ruijsink, B., Puyol-Anton, E., Bustin, A.,Cruz, G., ... Schnabel, J. A. (2020). Deep Learning Based Detection and Correction of Cardiac MR Motion Artefacts During Reconstruction for High-Quality Segmentation. IEEE Transactions on Medical Imaging, 1-1.
12. Ghesu, F. C., Krubasik, E., Georgescu, B., Singh, V., Zheng, Y., Hornegger, J., & Comaniciu, D. (2016). Marginal Space Deep Learning: Efficient Architecture for Volumetric Image Parsing. IEEE Transactions on Medical Imaging, 35(5), 1217-1228.

13. Roy, S., Menapace, W., Oei, S., Luijten, B., Fini, E., Saltori, C., ... Demi, L. (2020). Deep learning for classification and localization of COVID-19 markers in point-of-care lung ultrasound. IEEE Transactions on Medical Imaging, 1-1.
14. Duan, J., Bello, G., Schlemper, J., Bai, W., Dawes, T. J. W., Biffi, C., ... Rueckert, D. (2019). Automatic 3D bi-ventricular segmentation of cardiac images by a shape-refined multi-task deep learning approach. IEEE Transactions on Medical Imaging, 1-1.