

# Medical Image Segmentation using Deep Learning using Segnet

**SURESH SAPPA**

Assistant Professor, Department of Computer Science and Engineering, BVC Institute of Technology and Science, Amalapuram

**GARAGA SUBBARAO**

Assistant Professor, Department of Computer Science and Engineering, VSM College of Engineering, Ramachandrapuram

**Abstract:** Convolutional neural networks (CNNs) have achieved state-of-the-art performance for automatic medical image segmentation. However, they have not demonstrated sufficiently accurate and robust results for clinical use. In addition, they are limited by the lack of image-specific adaptation and the lack of generalizability to previously unseen object classes (a.k.a. zero-shot learning). To address these problems, we propose a novel deep learning-based interactive segmentation framework by incorporating CNNs into a bounding box and scribble-based segmentation pipeline. We propose image-specific fine-tuning to make a CNN model adaptive to a specific test image, which can be either unsupervised (without additional user interactions) or supervised (with additional scribbles). We also propose a weighted loss function considering network and interaction-based uncertainty for the fine-tuning. We applied this framework to two applications: 2D segmentation of multiple organs from fetal Magnetic Resonance (MR) slices, where only two types of these organs were annotated for training; and 3D segmentation of brain tumor core (excluding edema) and whole brain tumor (including edema) from different MR sequences, where only the tumor core in one MR sequence was annotated for training. Experimental results show that 1) our model is more robust to segment previously unseen objects than state-of-the-art CNNs; 2) image-specific fine-tuning with the proposed weighted loss function significantly improves segmentation accuracy; and 3) our method leads to accurate results with fewer user interactions and less user time than traditional interactive segmentation methods.

**Keywords:** Interactive Image Segmentation; Convolutional Neural Network; fine-Tuning; Fetal MRI; Brain Tumor;

## I. INTRODUCTION

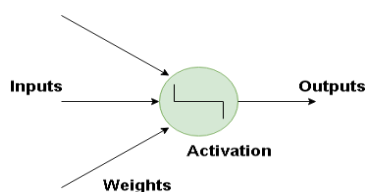
### Deep Learning

This perspective gave rise to the “Neural Network” terminology. The brain contains billions of neurons with tens of thousands of connections between them.

Deep learning algorithms resemble the brain in many conditions, as both the brain and deep learning models involve a vast number of computation units (Neurons) that are not extraordinarily intelligent in isolation but become intelligent when they interact with each other.

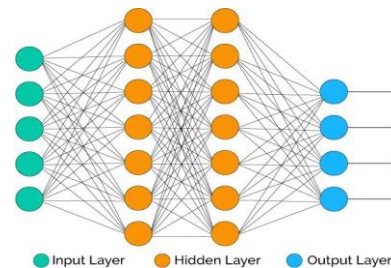
### Neurons

The basic building block for neural networks are artificial neurons, which imitate human brain neurons. These are simple, powerful computational units that have weighted input signals and produce an output signal using an activation function. These neurons are spread across the several layers in the neural network.



### How Does Artificial Neural Network Works?

Deep learning consists of artificial neural networks that are modelled on similar networks present in the human brain. As data travels through this artificial mesh, each layer processes an aspect of the data, filters outliers, spots familiar entities, and produces the final output.



**Input layer :** This layer consists of the neurons that do nothing than receiving the inputs and pass it on to the other layers. The number of layers in the input layer should be equal to the attributes or features in the dataset.

**Output Layer:**The output layer is the predicted feature, it basically depends in the type of model you’re building.

**Hidden Layer:** In between input and output layer there will be hidden layers based on the type of model. Hidden layers contain vast number of neurons. The neurons in the hidden layer apply

transformations to the inputs and before passing them. As the network is trained the weights get updated, to be more predictive.

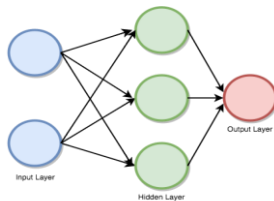
### Neuron Weights

Weights refer to the strength or amplitude of a connection between two neurons, if you are familiar with linear regression you can compare weights on inputs like coefficients we use in a regression equation. Weights are often initialized to small random values, such as values in the range 0 to 1.

### Feedforward Deep Networks

Feedforward supervised neural networks were among the first and most successful learning algorithms. They are also called deep networks, multi-layer Perceptron (MLP), or simply neural networks and the vanilla architecture with a single hidden layer is illustrated. Each Neuron is associated with other neuron with some weight,

The network processes the input upward **activating** neurons as it goes to finally produce an output value. This is called a forward pass on the network.



### Activation Function

An activation function is a mapping of summed weighted input to the output of the neuron. It is called an activation/ transfer function because it governs the inception at which the neuron is activated and the strength of the output signal.

## II. RELATED WORK

Computer vision and image analysis is the most important task of image segmentation. Several proposals to divide object feature extractions have been put forward. However, research challenges in the design of efficient and robust segmentation algorithms, owing to the complexity and variety of the images, remain (Yang and Tianzi, 2009). The aim of image segmentation is the division of the image into sectors which overlap with each other and are inconsistent in relation to definite properties like density, tone, color, and defined texture homogeneity. Four categories of image segmentation can be identified: groups, expos, threshold levels and extractions on the edge of the area. Each has their own strengths and weaknesses.

Benign diseases and malignant tumors cannot always be distinguished by Magnetic resonance imaging (MRI). This will prevent the observation

of defects in the brain due to the brains complex nature in terms of the shape size, location, tissues and the fact that it also contains 100 billion nerves. The problem is that brain tumors exist within a very complex human brain system (Balafar et al., 2008).

The level set method was first proposed by Osher and Sethian( 1988) as a way of using numerical methods to track contour evolution. A fresh volume of data input is generated by this technique in solving partial differential equations (PDE) with a function term extraction. Segmentation accuracy has been shown to be improved by the application of this technique. (Tsai and Oshe, 2011) developed numerical approximations for the level set method, regularizing solutions, representing object boundaries with curvature based velocities, regarding an image as a set of continuous functions etc. Considerable advancement has been made owing to the methods specified level concept in improving the evolution and implementation of the algorithms.

*CNNs for Image Segmentation:* For natural image segmentation, FCN [9] and DeepLab [10] are among the state- of-the-art performing methods. For 2D biomedical image seg- mentation, efficient networks such as U-Net [11], DCAN and Nabla-net [3] have been proposed. For 3D volumes, patch-based CNNs have been proposed for segmentation of the brain tumor [7] and pancreas, and more powerful end-to-end 3D CNNs include V-Net, HighRes3DNet [8], and 3D deeply supervised network.

*Interactive Segmentation Methods:* A wide range of interactive segmentation methods have been proposed [2]. Representative methods include Graph Cuts [7], Random Walks [8] and GeoS [9]. Machine learning has been popularly used to achieve high accuracy and interaction efficiency. For example, GMMs are used by GrabCut [10] to segment color images. Online Random Forests (ORFs) are employed by Slic-Seg for placenta segmentation from fetal Magnetic Resonance images (MRI). In [2], active learning is used to segment 3D Computed Tomography (CT) images. They have achieved more accurate segmentations with fewer user interactions than traditional interactive segmentation methods

To combine user interactions with CNNs, DeepCut [3] and ScribbleSup [3] propose to leverage user-provided bounding boxes or scribbles, but they employ user interactions as sparse annotations for the training set rather than as guidance for dealing with test images. 3D U-Net learns from annotations of some slices in a volume and produces a dense 3D segmentation, but is not responsive to user interactions. In [4], an FCN is combined with user interactions for 2D RGB image segmentation,

without adaptation for medical images. DeepGeoS [5] uses geodesic distance transforms of scribbles as additional channels of CNNs for interactive segmentation, but cannot deal with previously unseen object classes.

*Model Adaptation:* Previous learning-based interactive segmentation methods often employ image-specific models. For example, GrabCut and Slic-Seg [1] learn from the target image with GMMs and ORFs, respectively, so that they can be well adapted to the specific target image. Learning a model from a training set with image-specific adaptation in the testing stage has also been used to improve the segmentation performance. For example, an adaptive GMM has been used to address the distribution mismatch between the training and test images [6]. For CNNs, fine-tuning [5] is used for domain-wise model adaptation to address the distribution mismatch between different training sets. However, to the best of our knowledge, this paper is the first work to propose image-specific model adaptation for CNNs.

### III. PROPOSED METHOD

In order to do segmentation, image blocks are taken (with an odd number of pixels – the central pixel plus neighborhood) to determine the class (vessel or non-vessel) of the central pixel. Network training is performed on patches extracted from a set of images for which a manual segmentation exists. After such training, the network can be used to classify each pixel in the new examples of images after alternating four steps of C and MP layers two FC layers further combine the outputs into a 1D feature vector. The last layer is always a FC layer with one neuron per class (two in our case due to binary classification). In the output layer by using Softmax activation function each neuron's output activation can be taken as the probability of a particular pixel.

### IV. IMPLEMENTATION

#### Image sharpening and restoration

Image sharpening and restoration refers here to process images that have been captured from the modern camera to make them a better image or to manipulate those images in way to achieve desired result.

It refers to do what Photoshop usually does. This includes Zooming, blurring, sharpening, gray scale to colour conversion, detecting edges and vice versa, Image retrieval and Image recognition. Our human eye can only see the visible portion, in which we saw all the objects. But a camera can see the other things that awaked eye is unable to see.

The common examples are:

#### The original image:

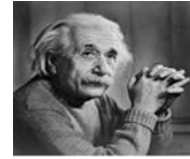


Figure 1 a Original Image

#### The zoomed image:

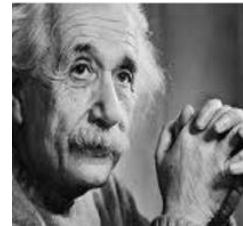


Figure 2 b Zoomed Image

These two are the representation of

1. Original Image
2. Zoomed image

#### Sharp image:

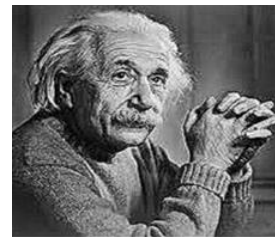


Figure 3 c Sharp Image

#### Blur image:

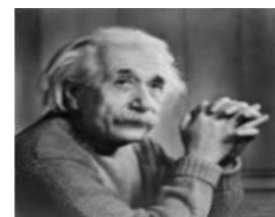


Figure 4 d Blur Image

The above Images are

- (3) Sharp Image
- (4) Blur Image

#### Edges:



Figure 5 e Edges

**Medical field:**

The common applications of DIP in the field of medical is

1. Gamma ray imaging
2. PET scan
3. X Ray Imaging
4. Medical CT
5. UV imaging

**UV imaging:**

In the field of remote sensing, the area of the earth is scanned by a satellite or from a very high ground and then it is analyzed to obtain information about it. One particular application of digital image processing in the field of remote sensing is to detect infrastructure damages caused by an earth quake. As it takes longer time to grasp damage, even if serious damages are focused on.

Since the area effected by the earthquake is sometimes so wide, that it not possible to examine it with human eye in order to estimate damages. Even if it is, then it is very hectic and time consuming procedure. So a solution to this is found in digital image processing. An image of the affected area is captured from the above ground and then it is analyzed to detect the various types of damage done by the earthquake.

- a. The key steps include in the analysis are
- b. The extraction of edges
- c. Analysis and enhancement of various types of edges



Figure 6 a UV Imaging

**Transmission and encoding:**

The very first image that has been transmitted over the wire was from London to New York via a submarine cable. The picture that was sent is shown below. The picture that was sent took three hours to reach from one place to another.

Now just imagine, that today we are able to see live video feed, or live CCTV footage from one continent to another with just a delay of seconds. It means that a lot of work has been done in this field too. This field doesn't only focus on transmission, but also on encoding. Many different formats have been developed for high or low bandwidth to

encode photos and then stream it over the internet or e.t.c.



Figure 7 Transmission & Encoding

**Machine or Robot vision**

Apart from the many challenges that a robot face today, one of the biggest challenge still is to increase the vision of the robot. Make robot able to see things, identify them, and identify the hurdles e.t.c. Much work has been contributed by this field and a complete other field of computer vision has been introduced to work on it.

**Hurdle detection**

Hurdle detection is one of the common task that has been done through image processing, by identifying different type of objects in the image and then calculating the distance between robot and hurdles.

**Line follower robot**

Most of the robots today work by following the line and thus are called line follower robots. This helps a robot to move on its path and perform some tasks. This has also been achieved through image processing.



Figure 8 Line Follower Robot

**Colour processing**

Colour processing includes processing of colored images and different colour spaces that are used. For example RGB colour model, YCbCr, HSV. It also involves studying transmission, storage, and encoding of these colour images.

**Pattern recognition**

Pattern recognition involves study from image processing and from various other fields that includes machine learning abranchofartificial intelligence.

In pattern recognition, image processing is used for identifying the objects in an image and then machine learning is used to train the system for the change in pattern. Pattern recognition is used in computer aided diagnosis, recognition of handwriting, recognition of images e.t.c



### Video processing

A video is nothing but just the very fast movement of pictures. The quality of the video depends on the number of frames/pictures per minute and the quality of each frame being used. Video processing involves noise reduction, detail enhancement, motion detection, frame rate conversion, aspect ratio conversion, colour space conversion e.t.c. The quality of the video depends on the number of frames/pictures per minute and the quality of each frame being used.

Most of the robots today work by following the line and thus are called line follower robots. This helps a robot to move on its path and perform some tasks. This has also been achieved through image processing.

Pattern recognition involves study from image processing and from various other fields that includes machine learning artificial intelligence. In pattern recognition, image processing is used for identifying the objects in an image and then machine learning is used to train the system for the change in pattern. Pattern recognition is used in computer aided diagnosis, recognition of handwriting, recognition of images e.t.c

Apart from the many challenges that a robot face today, one of the biggest challenge still is to increase the vision of the robot. Make robot able to see things, identify them, and identify the hurdles e.t.c. Much work has been contributed by this field and a complete other field of computer vision has been introduced to work on it. Colour processing includes processing of colored images and different colour spaces that are used.

For example RGB colour model, YCbCr, HSV. It also involves studying transmission, storage, and encoding of these colour images. In pattern recognition, image processing is used for identifying the objects in an image and then machine learning is used to train the system for the change in pattern. Pattern recognition is used in computer aided diagnosis, recognition of handwriting.

### V. SIMULATION RESULTS

TABLE

	ACCURACY	LOSS
TRAINING	0.8170	0.5219
TESTING	0.7663	0.6853
VALIDATION	0.7838	0.6389

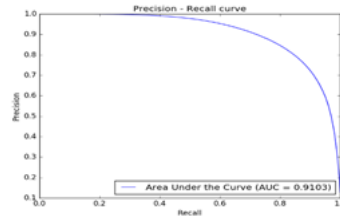


Fig 9 precision recall curve

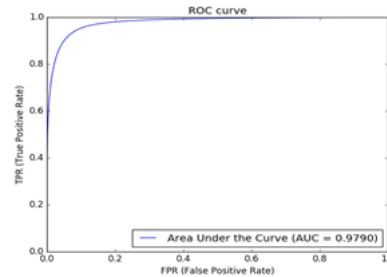


Fig. 10 ROC curve for AUC

### VI. CONCLUSION

In conclusion we give an overview of plans for future work which would lead to enhancements. Semantic segmentation is an important step towards understanding and inferring different objects and their arrangements observe dina scene. This has wide array of applications ranging from estimating scene geometry, inferring support-relationships among objects to autonomous vehicle driving. Early methods that relied on low-level vision cues have fast been superseded by popular machine learning algorithms. In particular, deep learning has seen huge success lately in hand written digit recognition, speech, categorizing whole images and detecting objects in im- ages also seen growing interest in semantic pixel- wise labelling problems. However, these re- cent approaches have tried to directly adopt deep architectures designed for category prediction to pixel-wise la- belling.

### REFERENCES

- [1] Bühler, K., Felkel, P., La Cruz, A., 2004. Geometric methods for vessel visualization and quantification—a survey. Springer.
- [2] Caffe | Deep Learning Framework [WWW Document],
- [3] n.d. URL <http://caffe.berkeleyvision.org/> (accessed 10.2.14).
- [4] Cireşan, D.C., Giusti, A., Gambardella, L.M., Schmidhu- ber, J., 2013. Mitosis detection in breast cancer histol- ogy images with deep neural networks, in: Medical Image Computing and Computer-Assisted Interven- tion–MICCAI 2013. Springer, pp. 411–418.

- [5] Cirezan, D.C., Meier, U., Masci, J., Maria Gambardella, L., Schmidhuber, J., 2011a. Flexible, high performance convolutional neural networks for image classification, in: IJCAI Proceedings-International Joint Conference on Artificial Intelligence. p.1237.
- [6] Cirezan, D.C., Meier, U., Masci, J., Maria Gambardella, L., Schmidhuber, J., 2011b. Flexible, high performance convolutional neural networks for image classification, in: IJCAI Proceedings-International Joint Conference on Artificial Intelligence. p.1237.
- [7] Cirezan, D., Meier, U., Schmidhuber, J., 2012. Multi-column deep neural networks for image classification, in: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, pp. 3642–3649.
- [8] Faust, O., Acharya, R., Ng, E.Y.-K., Ng, K.-H., Suri, J.S., 2012. Algorithms for the automated detection of diabetic retinopathy using digital fundus images: a review. *J. Med. Syst.* 36, 145–157.
- [9] Felkel, P., Wegenkittl, R., Kanitsar, A., 2001. Vessel tracking in peripheral CTA datasets-an overview, in: Computer Graphics, Spring Conference On, 2001. IEEE, pp. 232–239.
- [10] Fraz, M.M., Remagnino, P., Hoppe, A., Uyyanonvara, B., Rudnicka, A.R., Owen, C.G., Barman, S.A., 2012. Blood vessel segmentation methodologies in retinal images – A survey. *Comput. Methods Programs Bio-med.* 108, 407–433. doi:10.1016/j.cmpb.2012.03.009