



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirce.com

Vol. 5, Issue 3, March 2017

Survey on Semantic Image Segmentation by Convolution Nets and Transformation

Sahana Patil, Sandeep Kumar

M.Tech Student, Dept.of Computer Science and Engg, Cambridge institute of Technology, Bangalore, India

Associate Professor, Dept.of Computer Science and Engg, Cambridge institute of Technology, Bangalore, India

ABSTRACT: In recent years, substantial research has been carried out in the field of image processing to evaluate different structures and information from images. Image processing techniques have played a pivotal role in a wide range of applications. Convolution Neural Networks (CNNs) are an alternative type of neural network that can be used to reduce spectral variations and model spectral correlations which exist in image. Image segmentation is the process of partitioning an image into multiple segments, so as to change the representation of an image into something that is more meaningful and easier to analyze. Segmentation technique, basically convert the complex image into the simple image. For the image segmentation process the Watershed technique is used the segmented image is further post-processed using Morphological operation form the post-processing the fully segmented image can be obtained with high accuracy.

KEYWORDS: DCNN, LANEZOSE Interpolation, Watershed Transformation, Morphological operation.

I. INTRODUCTION

The image classification tasks are quite artificial. Typically, it is assumed that the object of interest is centered and at a fixed scale, i.e. that the localization problem has been solved using the neural network. Natural scenes rarely contain a single object or object class. Object detection and object-class segmentation are thus the logical step towards general images. In the modern applications of the neural network which involves more than one hidden layers- layers that sit between the input and the output layers, to allow the network to model non-linearity in the feature space. The number of hidden layers and the number of hidden nodes in each layer are hyper-parameters that are not always easy to determine. Much number of strategies has been used to include spatial and topological properties in the image segmentation stage. It is generally accepted that grouping of nearby pixels by modeling neighborhood relationships as (a, b) connected graphs may lead to meaningful image objects. The topological concepts suffer from ambiguity due to image pixels are two dimensional entities. CNNs are a type of deep models in which trainable filters and local neighborhood pooling operations are applied alternating on the raw input images, resulting in a hierarchy of increasingly complex features. The information in the image is need for the training part of the DCNN and taken as inputs and then generated the segmentation maps as outputs. The multiple intermediate layers applied convolution, pooling, normalization, and other operations to capture the highly nonlinear mappings between inputs and outputs. The output of the DCNN is given to the LANEZOSE interpolation in which filtering of the image using the low pass filter by that the smoothing of the image is done. The smoothed image further transformed using the watershed transformation in which the image is segmented properly by this the final segmented image is obtained.

II. LITERATURE SURVEY

The author Matthew Lai et.al [01] have presented overview of the current state of the art deep learning architectures and optimization techniques, and uses the ADNI hippocampus MRI dataset as an example to compare the effectiveness and efficiency of different Convolution architectures on the task of patch-based 3-dimensional hippocampus segmentation, which is important in the diagnosis of Alzheimer's Disease. They have found that a slightly unconventional "stacked 2D" approach provides much better classification performance than simple 2D patches



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirce.com

Vol. 5, Issue 3, March 2017

without requiring significantly more computational power. They have also examined the popular "tri-planar" approach used in some recently published studies, and found that it provides much better results than the 2D approaches, but also with a moderate increase in computational power requirement. Finally, evaluated a full 3D Convolution architecture, and found that it provides marginally better results than the tri-planar approach, but at the cost of a very significant increase in computational power requirement.

In this paper the author D. Marmanis et.al [02] have discussed on an end-to-end trainable deep Convolution neural network (DCNN) for semantic segmentation with built-in awareness of semantically meaningful boundaries. Semantic segmentation is a fundamental remote sensing task, and most state-of-the-art methods rely on DCNNs as their workhorse. A major reason for their success is that deep networks learn to accumulate contextual information over very large windows (receptive Fields). However, this success comes at a cost, since the associated loss of effective spatial resolution washes out high-frequency details and leads to blurry object boundaries. Here, they have proposed a counter this effect by combining semantic segmentation with semantically informed edge detection, thus making class boundaries explicit in the model. Firstly, construct a comparatively simple, memory-efficient model by adding boundary detection to the SEGNET encoder-decoder architecture. Secondly, include boundary detection in FCN-type models and set up a high-end classifier ensemble. Also showed boundary detection significantly improves semantic segmentation with CNNs. Our high-end ensemble achieves > 90% overall accuracy on the ISPRS Vaihingen benchmark.

The author George Papandreou et.al [03] has explored the use of weak or partial annotation in training a state of art semantic image segmentation model. Extensive experiments on the challenging PASCAL VOC 2012 dataset have shown that: (1) Using weak annotation solely at the image-level seems insufficient to train a high-quality segmentation model. (2) Using weak bounding-box annotation in conjunction with careful segmentation inference for images in the training set suffices to train a competitive model. (3) Excellent performance is obtained when combining a small number of pixel-level annotated images with a large number of weakly annotated images in a semi-supervised setting, nearly matching the results achieved when all training images have pixel-level annotations. (4) Exploiting extra weak or strong annotations from other datasets can lead to large improvements.

The author Evan Shelhamer et.al [04] have discussed on Convolutional networks are powerful visual models that yield hierarchies of features. They have discussed on Convolutional networks by themselves, trained end-to-end, pixels-to-pixels, improve on the previous best result in semantic segmentation. Our key insight is to build "fully Convolutional" networks that take input of arbitrary size and produce correspondingly-sized output with efficient inference and learning. They have defined and detail the space of fully convolutional networks, explain their application to spatially dense prediction tasks, and draw connections to prior models. They have adapted contemporary classification networks (Alex Net, the VGG net, and Google Net) into fully Convolutional networks and transfer their learned representations by fine-tuning to the segmentation task. Then define a skip architecture that combines semantic information from a deep, coarse layer with appearance information from a shallow, fine layer to produce accurate and detailed segmentations. Our fully Convolutional network achieves improved segmentation of PASCAL VOC (30% relative improvement to 67.2% mean IU on 2012), NYUDv2, SIFT Flow, and PASCAL-Context, while inference takes one tenth of a second for a typical image.

The author Tadej Vodopivec et.al [05] has proposed the Convolutional neural networks. Deep Learning has already been applied to segmentation, and recent architectures tend to be made of two parts: The first part applies convolution and pooling layers to the input image to produce a compact, low-resolution representation; the second part applies deconvolutional layers to this representation to produce the final segmentation, at the same resolution as the input image. This typically results in over smoothed segments.

The author Peng Gu et.al [06] have discussed on the automated algorithm to segment 3D ultrasound volumes into three major tissue types: cyst/mass, fatty tissue, and fibro-glandular tissue. To test its efficacy and consistency, the proposed automated method was employed on a database of 21 cases of whole breast ultrasound. Experimental results show that proposed method not only distinguishes fat and non-fat tissues correctly, but performs well in classifying cyst/mass. Comparison of density assessment between the automated method and manual segmentation demonstrates good consistency with an accuracy of 85.7%. Quantitative comparison of corresponding tissue volumes, which uses overlap ratio, gives an average similarity of 74.54%, consistent with values seen in MRI brain segmentations. Thus proposed method exhibits great potential as an automated approach to segment 3D whole breast ultrasound volumes



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirce.com

Vol. 5, Issue 3, March 2017

into functionally distinct tissues that may help to correct ultrasound speed of sound aberrations and assist in density based prognosis of breast cancer.

The author Jean-Charles Bricola et.al [07] have discussed on the problem of computing a depth map from a pair of rectified stereo images is undoubtedly a classic one in computer vision. When a point of the scene projects onto the two image planes, it does so with the same ordinates but with different abscissa. The difference of abscissa corresponds to what is commonly referred to as the disparity and is inversely proportional to the point's depth being sought for. Finding point correspondences between the left and right views of the stereo pair is relatively easy across non-uniformly textured areas. However, homogeneous regions are the source of matching ambiguities whilst the occlusion phenomenon makes it impossible for some pixels to have a correspondence and thus require their disparity to be estimated according to a suitable model. In order to overcome these two difficulties, it is usual to devise algorithms which ensure, on the one hand, that disparities evolve smoothly across the low- textured areas of the image for which the depth map is estimated and, on the other hand, that disparities remain consistent with the resulting warping of stereo images.

In this paper, authors Tara N. Sainath et.al [08] have explored applying CNNs to large vocabulary speech tasks. First, determine the appropriate architecture to make CNNs effective compared to DNNs for LVCSR tasks. Specifically, they focus on how many convolution layers are needed, what is the optimal number of hidden units, what is the best pooling strategy and the best input feature type for CNNs. Then they explore the behavior of neural network features extracted from CNNs on a variety of LVCSR tasks, comparing CNNs to DNNs and GMMs. They find that CNNs offer between a 13-30% relative improvement over GMMs, and a 4-12% relative improvement over DNNs, on a 400-hr Broadcast News and 300-hr Switchboard task.

In this paper, the author Isabella Nogues et.al [09] has discussed on deep convolution neural networks to computer-aided detection problems. They first explore and evaluate different CNN architectures. The studied models contain 5 thousand to 160 million parameters, and vary in numbers of layers. Then they have evaluated the influence of dataset scale and spatial image context on performance. Finally, examine when and why transfer learning from pre-trained Image Net (via fine-tuning) can be useful. They achieve the state-of-the-art performance on the meditational LN detection, with 85% sensitivity at 3 false positive per patient, and report the first five-fold cross-validation classification results on predicting axial CT slices with ILD categories. Our extensive empirical evaluation, CNN model analysis and valuable insights can be extended to the design of high performance CAD systems for other medical imaging tasks.

In this paper, author Emmanuel Maggiori et.al [10] have propose an end-to-end framework for the dense, pixel wise classification of satellite imagery with convolutional neural networks (CNNs). In our framework, CNNs are directly trained to produce classification maps out of the input images. They first devise a fully convolutional architecture and demonstrate its relevance to the dense classification problem. Then they address the issue of imperfect training data through a two-step training approach: CNNs are first initialized by using a large amount of possibly inaccurate reference data and then refined on a small amount of accurately labeled data. To complete our framework multi-scale neuron modules have designed and that alleviates the common trade-off between recognition and precise localization. A series of experiments show that our networks take into account a large amount of context to provide fine-grained classification maps.

The author Danyal Maheshwari et.al [11] proposes the use of image processing techniques in the medical field. The objective of this paper is to develop a MATLAB based algorithm that can be used to extract a brain tumor from a MRI Image. In this research, they have performed some noise removal functions, segmentation techniques and morphological operations for detection and extraction which are the basic concepts of image processing. They have developed a Watershed Transform technique based on internal and external markers. The detection and extraction of tumor from MRI image of the brain is done by using MATLAB software.

In this paper author Soumen Biswas et.al [12] discusses about the processing and analysis of microscopic image is a broad field in research area. White Blood Cells (WBC), Red Blood Cells (RBC), Platelets and Plasma are four important constituents. Identifying the different blood cells helps to find out disease detected cells. Approaching a new automated blood cell detection algorithm reduce the detecting time and produce accurate results. Segmentation over an image is performed to detect the required information in image analysis. This image processing technique of blood cell detection can also be used for different disease detection by analyzing the each cell accurately.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirce.com

Vol. 5, Issue 3, March 2017

Table 1: Survey Papers on the Various Algorithms

Title	Year	Algorithm	Advantage	Performance
Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning	2016	Deep Convolution Neural Network	Deep learning Algorithm	80% of the improvement is seen
Adaptive Rational Interpolation: Arnoldi and Lanczos-like Equations	2014	Arnoldi and Lanczos interpolation	Better filtering Algorithm	Better Performance can be seen
Blood Cell Detection using Thresholding Estimation Based Watershed Transformation with Sobel Filter in Frequency Domain	2016	Sobel Filter and Watershed Algorithm	Proper segmentation of image	90% of the improvement is seen
Individual Tree Crown (ITC) Delineation Using Watershed Transformation Algorithm For Tropical Lowland Dipterocarp	2015	Watershed Algorithm	Improved segmentation Algorithm	90% of the improvement is seen

III. PROPOSED SYSTEM

The architecture of proposed system is shown in figure 1. In which the segmentation and classification of the image is done using the convolution nets and by this the image is classified, firstly the image is preprocessed using the preprocessing algorithm in which the smoothing and removal of the noise is done and then the image is passed to the neural network in which the training of the image is done and then the object in the image is segmented and classified. In the next step the classified output from the neural network is taken and interpolation is applied for the smoothing of the image is done. For the proper segmentation of the image the transformation algorithm is applied on the image, after transformation the post-processing is carried out for proper segmentation of the image. From all these process the proper segmented image can be obtained. Figure 2 show the input image and output image for the segmentation process.

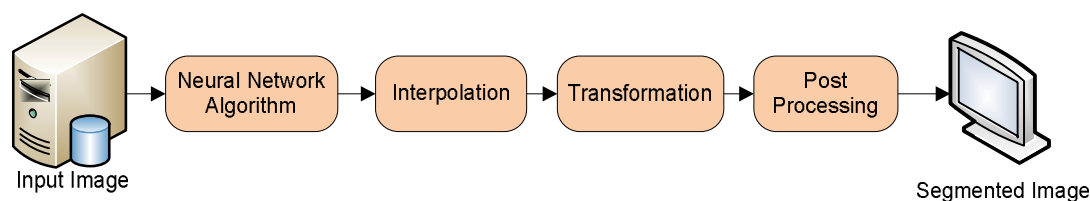


Figure 1: Block Diagram of the Image Segmentation

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirce.com

Vol. 5, Issue 3, March 2017

IV. EXPECTED RESULTS

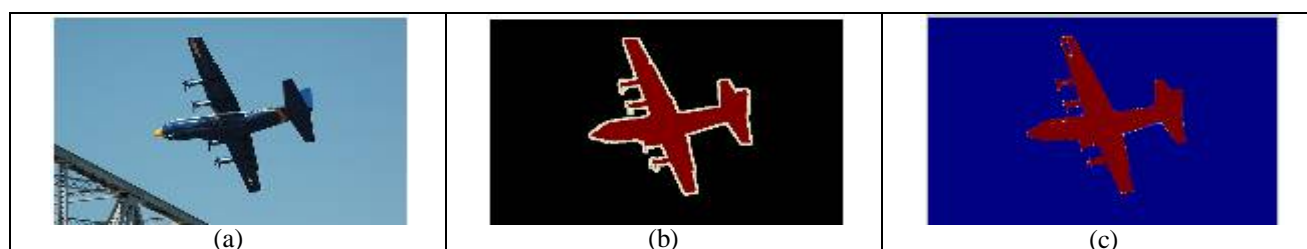


Figure 2: (a) Input image for the neural network technique; (b) Output the neural network; (c) Output the Transformation.

Table 2: Parameter comparison table for Expected

Paper	Algorithm	Accuracy
D. Marmanis et.al [2]	Deep Convolution Neural Network	80%
Danyal Maheshwari et.al [11]	Watershed transformation and Morphological operation	85%
Expected	DCNN, Watershed transformation and Morphological operation	86%

V. CONCLUSION

The survey on the segmentation of the image is done using the Deep Convolution Neural Network for the type of deep models in which trainable filters and local neighborhood pooling operations are applied alternating on the raw input images, resulting in a hierarchy of increasingly complex features. The information in the image is need for the training part of the DCNN and taken as inputs and then generated the segmentation maps as outputs. The output of the DCNN is is segmented using the segmentation algorithm and by this the higher accuracy can be obtained.

REFERENCES

- [1] Matthew Lai, "Deep Learning for Medical Image Segmentation", IEEE, pp. 1 – 7, 2015.
- [2] D. Marmanis, K. Schindler, J. D. Wegner, S. Galliani, M. Datcu, U.Stilla, "Classification with an edge: improving semantic image segmentation with boundary detection", IEEE, pp. 502 – 512, 2016.
- [3] George Papandreou, Liang-Chieh Chen, Kevin Murphy and Alan L. Yuille "Weakly- and Semi-Supervised Learning of a Deep Convolutional Network for Semantic Image Segmentation", IEEE, pp. 1742 – 1750, 2015.
- [4] Evan Shelhamer, Jonathan Long, and Trevor Darrell, "Fully Convolutional Networks for Semantic Segmentation", IEEE, pp. 1 – 1, 2016.
- [5] Tadej Vodopivec, Vincent Lepetit and Peter Peer, "Fine Hand Segmentation using Convolutional Neural Networks", IEEE, 2016.
- [6] Peng Gu, Won-Mean Lee, Marilyn A. Roubidoux, Jie Yuan, Xueding Wang and Paul L. Carson, "Automated 3D ultrasound image segmentation to aid breast cancer image interpretation", Elsevier, 2015.
- [7] Jean-Charles Bricola, Michel Bilodeau and Serge Beucher, "Morphological processing of stereoscopic image superimpositions for disparity map estimation", Elsevier, 2016.
- [8] Tara N. Sainath, Abdel-rahman Mohamed, Brian Kingsbury and Bhuvana Ramabhadran "Deep Convolutional Neural Networks For LVCSR", IEEE, pp. 315 - 320, 2013.
- [9] Isabella Nogue, Jianhua Yao, Daniel Mollura and Ronald M. Summers, "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning", IEEE, pp. 1285 – 1298, 2016.
- [10] Emmanuel Maggiori, Yuliya Tarabalka, Guillaume Charpiat and Pierre Alliez "Convolutional Neural Networks for Large-Scale Remote Sensing Image Classification", IEEE, Vol. 55, No. 2, 2016.
- [11] Danyal Maheshwari, Ali Akber Shah, Muhammad Zakir Shaikh, Bhawani Shankar Chowdhry and Sara Rahman Memon, "Extraction of Brain Tumour in MRI Images Using Marker Controlled Watershed Transform Technique in MATLAB", Journal Of Biomedical Engineering And Medi Cal Imaging, Vol. 2, No. 4, 2015.
- [12] Soumen Biswas and Dibyendu Ghoshal "Blood Cell Detection using Thresholding Estimation Based Watershed Transformation with Sobel Filter in Frequency Domain", Elsevier, 2016.