ORIGINAL ARTICLE

MMTDNN: Multi-View Massive Training Deep Neural Network for Segmentation and Detection of Abnormal Tissues in Medical Images

Hassan Khastavaneh * , **Hossein Ebrahimpour-Komleh**

Department of Computer Engineering, Faculty of Computer and Electrical Engineering, University of Kashan, Kashan, Iran

*Corresponding Author: Hassan Khastavaneh
Email: khastavaneh@hotmail.com

Received: 13 January 2020 / Accepted: 12 March 2020

Abstract

Purpose: Automated segmentation of abnormal tissues in medical images is considered as an essential part of those computer-aided detection and diagnosis systems which analyze medical images. However, automated segmentation of abnormalities is a challenging task due to the limitations of imaging technologies and complex structure of abnormalities, including low contrast between normal and abnormal tissues, shape diversity, appearance inhomogeneity, and the vague boundaries of abnormalities. Therefore, more intelligent segmentation techniques are required to tackle these challenges.

Materials and Methods: In this study, a method, which is called MMTDNN, is proposed to segment and detect medical image abnormalities. MMTDNN, as a multi-view learning machine, utilizes convolutional neural networks in a massive training strategy. Moreover, the proposed method has four phases of preprocessing, view generation, pixel-level segmentation, and post-processing. The International Symposium on Biomedical Imaging (ISBI)-2016 dataset is used for the evaluation of the proposed method.

Results: The performance of the proposed method has been evaluated on the task of skin lesion segmentation as one of the challenging applications of abnormal tissue segmentation. Both qualitative and quantitative results demonstrate outstanding performance. Meanwhile, the accuracy of 0.973, the Jaccard index of 0.876, and the Dice similarity coefficient of 0.931 have been achieved.

Conclusion: In conclusion, the experimental result demonstrates that the proposed method outperforms state-of-the-art methods of skin lesion segmentation.

Keywords: Medical Imaging; Abnormal Tissues Segmentation; Convolutional Neural Networks; Multi-View Learning; Artificial Neural Networks; Multi-View Massive Training Deep Neural Network.

DOI: https://doi.org/10.18502/fbt.v7i1.2722

1. Introduction

Medical image segmentation is the task of delineating structures of interest in medical images, either manually or automatically. Segmentation of medical images has broad applications, including locating tumors and pathologies, studying anatomical structures, and measuring tissue volume. Among these, delineation of tissue abnormalities is one of the major application areas of medical image segmentation. As monitoring and analysis of abnormal tissues are crucial for optimal treatment, accurate segmentation of these abnormalities is required. In contrast to the automated segmentation of abnormal tissues, manual segmentation is commonly time-consuming, cumbersome, error-prone, and subjective. Therefore, the automatic segmentation of these tissue abnormalities is preferred.

Automated segmentation of abnormal tissues in medical images is a challenging and sophisticated task in both Computer-Aided Detection (CADe) and Computer-Aided Diagnosis (CADx) systems [1]. There are many cases of complexity, which mainly roots in the complex structure of the body tissues, abnormalities, and also limitations of imaging technologies. These complexities include inhomogeneity, partial volume effect, the existence of noise and artifacts, low contrast between normal and abnormal tissues, shape diversity, and the fuzzy border of abnormalities. With these circumstances, sophisticated methods are required to tackle mentioned difficulties; and consequently, accurate segmentation of abnormalities in medical images.

Representation learning [2] methods are a candidate solution for challenging real-world problems, including segmentation of abnormal tissues in medical images and understanding their contents as well. The reason for the prosperity of representation learning methods is that they cover more general priors of real-world intelligence. These priors include smoothness, multiple explanatory factors, the sparsity of features, transfer learning, independence of features, natural clustering and distributed representation, semi-supervised learning, and hierarchical organization of features [3]. As describing the texture of medical images and segmentation of abnormalities based on the texture contents are among difficult tasks of computer vision, the power of representation learning is employed to address related issues.

There are many representation learning techniques which act as end-to-end learning machines capable of segmenting abnormalities in medical images. In such techniques, the features are learned directly from the training images as a part of the training process. These methods, which mainly root in the neural networks, are fallen into two main categories of Massive Training Artificial Neural Networks (MTANNs) and Convolutional Neural Networks (CNNs). With many similarities, the main difference between these two techniques is the place of the convolution operator. The convolution operator is outside of the network for the MTANN family while this operator is inside the network for CNNs [4].

Although MTANNs are firstly proposed for the reduction of false positives in computerized detection of lung nodules in CT images [5], they are capable to perform various image processing and pattern recognition tasks thanks to the ability to learn useful features from the training data. As Figure 1 depicts, both input and output of a typical MTANN are images. Training input images of MTANN are overlapped sub-regions of the original image, and the teacher image, which serves the network as desired output is an image as the same size as the training sub-regions. However, the main burden of learning is carried by the fully connected neural network component of MTANN. This component takes a gray level sub-image and produces a single output corresponding to the central pixel of that sub-image. Concerning the fact that the input of MTANN is a sub-image, in order to process and judge all pixels of the input image, the trained MTANN must slide over the input image. This sliding process is called convolution. In contrast to the CNN, where convolution is controlled inside the network structure, this process is controlled outside of the network structure in MTANN.

Some representative abnormal tissue segmentation methods that are based on MTANN are segmentation of Multiple Scleroses (MS) lesions [6], [7], detection of polyps in CT colonography [8], [9], and other tasks related to segmentation and detection of abnormal tissues [10], [11], [12].

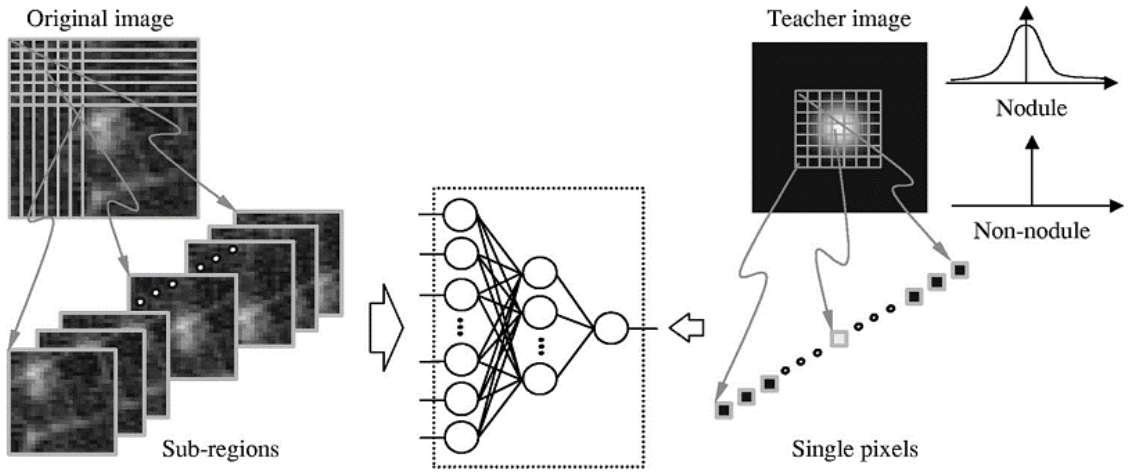

Figure 1. Architecture and training method of MTANN for detection of lung nodules in CT images [5]

The leading representative CNN-based studies for segmenting tissues include segmentation of brain tumors [13], [14], segmentation of bladder in CT images [15], skin lesion segmentation [16], [17], MS lesion segmentation [18], segmentation of white matter hyper-intensities [19], and segmentation of soft tissues in digital mammograms [20].

In conclusion, both methods based on MTANN and CNN have many similarities. More concretely, both network architectures attempt to learn discriminative features directly from the data. Although CNNs are very good for analysis of visual data because of their excellent representation capabilities, they are suffering from some limitations, including long training time and need of massive amounts of training data. In contrast, the methods based on MTANN need less training example which is very beneficial in the medical image analysis methods, where always proper training data are rare. In another perspective, MTANN may suffer from the fact that the data could not be represented as well as CNN. So, it is desired to have networks that gain benefits from both architectures. In other words, the networks with high representation capabilities and trainable based on a small amount of data are desired.

This study aims to propose a pixel-based method for the segmentation of abnormalities in medical images called Multi-view Massive Training Deep Neural Network (MMTDNN). The proposed method relies mostly on the multi-view capabilities of Convolutional Neural Networks (CNNs), which attempts to describe sub-images that surround all pixels in a typical medical image in different perspectives and scales [21].

The multi-view concept allows to various global, regional, and local aspects of images which clue the segmentation be considered. Having CNNs with multi-view capability brings some challenges to the training of the proposed method. In order to tackle these challenges, besides the proper design of network architecture, enormous amounts of data are needed. The massive training concept in the proposed method deals with the generation of too many sub-images from the currently accessible image data for network training.

The remainder of this paper is organized as follows. The proposed MMTDNN, as a segmentation method of abnormal tissues, is presented in section 2. Experimental results are presented and discussed in section 3. Finally, the paper is concluded in section 4.

2. Materials and Methods

2.1. Data-Set

In order to evaluate the performance of the proposed method, a benchmark and publicly available data-set of skin lesions known as ISBI-2016 [22] is used. Because of significant variations in the intensity, color, shape, size, and texture of both lesion and normal areas, the segmentation of these lesions is challenging. This data-set consists of 900 and 379 training and testing images, respectively. Both

training and testing images are provided with the ground truth generated by experts. Given the fact that these images are collected from various centers and devices; and the data-set contains lesions with various irregular shapes, diverse backgrounds, low contrast between lesion and non-lesion areas, inhomogeneous appearance, hair artifacts, color charts, and fuzzy borders, they are appropriate for robust evaluation of automated lesion segmentation methods.

2.2. Method

In this section, the proposed MMTDNN for the segmentation of abnormal tissues in medical images is explained. MMTDNN is considered as an extension of MTANN; in fact, the representation capability of MTANN is increased by replacing the fully connected layers with multiple convolutional layers. In other words, we add convolutional layers to the MTANN architecture to enable considering sub-images in their original two-dimensional form [23]. These convolutional layers allow generating better features than fully connected layers of the original MTANN. More concretely, convolution layers perform feature learning while fully connected layers perform the task of high-level reasoning [2]. In contrast to MTANN, which models abnormalities in local scale, the proposed method considers multiple views at multiple scales of the input image. In other words, the proposed method labels each pixel of the input image as either normal or abnormal by considering its local, regional, and global contextual information through local, regional, and global views, respectively. Indeed, the image patches, which surround a typical pixel in different views, are abstracted together using a customized CNN. Although the multi-view extension of MTANN leads a better description of image patches by utilizing convolutional layers inside the network, on the other hand, the power of MTANN utilizing outside convolution remains. Outside convolution allows generating multiple overlapped sub-regions of input image and model abnormalities with a small number of training images which is beneficiary for medical applications. The proposed MMTDNN uses the benefits of both CNN and MTANN simultaneously as two popular methods for the analysis of medical images.

As the block diagram of the proposed MMTDNN in Figure 2 demonstrates, it has four phases, namely preprocessing, view generation, pixel-level segmentation, and post-processing. The input of the proposed method is a raw two-dimensional medical image, and its output is a binary mask that corresponds to abnormal areas of the input image. The sections ahead shall explain the goal and functionality of all four main phases of the proposed method.

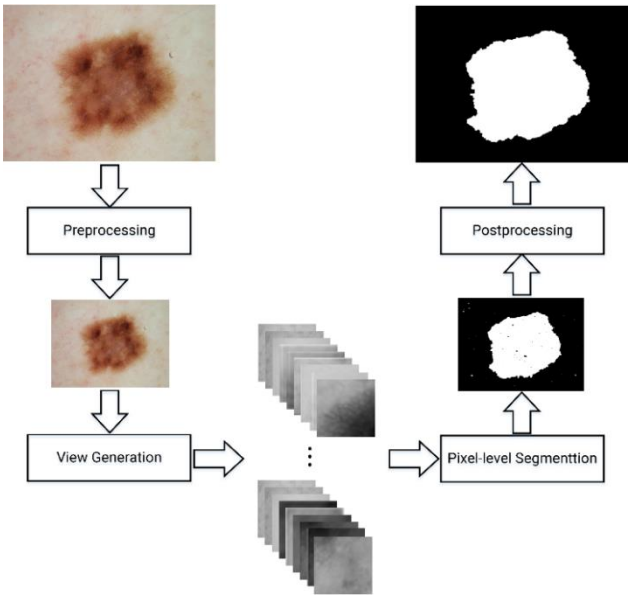

Figure 2. Block diagram of the proposed MMTDNN; input of MMTDNN is a medical image with a lesion area, and the output is a lesion binary mask

2.2.1. Preprocessing

In order to segment the input images efficiently, two preprocessing tasks of image resizing and intensity normalization are performed in the preprocessing phase. In the image resizing task, all the images are resized to 300 × 400 pixels by utilizing pixel-area-relation interpolation. Also, the intensity of each RGB channel of images is rescaled to the range [0,1] using min-max normalization.

2.2.2. View Generation

As mentioned previously, the proposed pixel-based segmentation method judges each pixel of the input image by considering that pixel in various views. The main reason for considering a pixel in multiple views is that the information from multiple scales is

complementary and leads to a more robust segmentation of abnormal tissues. In the proposed method, three views are supposed to impose complementary contextual information about each pixel in the segmentation pipeline. These views carry information in three levels of local, regional, and global. The local view is an n×n image patch centered on the pixel under analysis. The regional view is a 2n × 2n image path which is again centered on the pixel under analysis; this view allows having a broader view over the pixel under analysis. The global view is a 4n × 4n image path which is more extensive than both local and regional views.

In the process of view generation, the marginal pixels, which are far from the center, are challenging. In this situation, a typical view that surrounds a marginal pixel may fall outside of the image under consideration. In order to tackle this problem, a suitable padding strategy is employed. In the proposed method, the images are extended by replicating columns and rows to empty columns and rows in the views. Moreover, by knowing the fact that lesions are mainly placed in the center of input images, marginal pixels are not very informative in the process of learning.

After acquiring all three local, regional, and global views of a pixel, all of the image patches, which correspond to a view, are downscaled to have the size of n×n as the same size of the local view. These image patches are concatenated together to form a nine-channel image. This nine-channel image can be considered as a comprehensive tensor which contains information about a pixel. In other words, this nine-channel image or comprehensive tensor allows having more complementary information about a pixel. A set of tensors, each of which corresponds to a pixel of the preprocessed image are extracted in lexicographic order. Finally, each tensor, which is considered as an extension of a pixel, is passed to the next phase to compute the likelihood of that pixel belonging to the lesion area or not.

2.2.3. Pixel-Level Segmentation

In order to perform accurate pixel-level segmentation, all three local, regional, and global views, which are concatenated to form a comprehensive tensor, must be considered simultaneously. For this matter, a CNN capable of handling this comprehensive tensor is proposed. The input of this network is a comprehensive tensor with extended information about a pixel. In addition, the network output is an abnormality likelihood of the pixel under consideration. The network receives contextual information at multiple levels through the tensor and jointly exploits features at different scales. Therefore, this network is capable of abstracting input data; and consequently, producing a posterior probability for each tensor which describes a pixel. As a result, pixels with a high posterior probability likely belongs to the abnormal areas. After considering all pixels of the input image, a posterior probability map for that image is generated; later, this posterior probability map is processed to produce the final lesion mask.

The segmentation network must have a particular structure and utilize proper regularization techniques to be able to compute the likelihood properly. With these conditions, a deep enough network is needed to generate more representative features to deal with the significant variations of both normal and abnormal tissues in medical images. As it is depicted in Figure 3, the network consists of 13 layers. Moreover, the total number of trainable parameters in this network is 4,544,769, which leads the network optimization to be a hard task. In this architecture, nine convolutional layers carry the burden of feature learning. These features distinguish pixels of healthy tissues from the pixels belonging to abnormal tissues. The convolution kernel size for non-dimension-expansion layers is 3×3. Fully connected layers at the end of the network carry the task of high-level reasoning and computing the final likelihood according to the features learned in the previous layers. Since training of such deep networks is challenging, some techniques, including batch normalization [24], dropout [25], and residual learning [26] are utilized to facilitate network training and parameter optimization.

To conclude, the proposed pixel-based segmentation network makes use of multiple convolutional layers to abstract and extract discriminatory features for distinguishing abnormal pixels from the normal ones. After the pixel-level segmentation phase, a likelihood map of the input

Figure 3. The block diagram of the proposed pixel-based segmentation network

image is generated by putting all likelihood together in a 2D grid. From this likelihood map, the abnormal tissue mask is extracted after some post-processing tasks.

To conclude, the proposed pixel-based segmentation network makes use of multiple convolutional layers to abstract and extract discriminatory features for distinguishing abnormal pixels from the normal ones. After the pixel-level segmentation phase, a likelihood map of the input image is generated by putting all likelihood together in a 2D grid. From this likelihood map, the abnormal tissue mask is extracted after some post-processing tasks.

2.2.4. Post-Processing

In the post-processing phase of the proposed method, the lesion likelihood map from the previous phase is further processed to generate a binary lesion mask. For this matter, the output lesion mask is thresholded to generate the initial binary lesion mask. In order to perform binarization, the Otsu thresholding method is employed [27]. The initial lesion mask goes under the connected-component-analysis processes to select the largest connected component as the best candidate for lesion mask. Finally, this largest connected component goes under the whole-filling process to produce the final lesion mask.

3. Results and Discussion

After the successful implementation of the proposed segmentation method, its performance has been evaluated in the challenging task of skin lesion segmentation. In order to quantitatively measure the performance of the proposed segmentation method, a variety of standard measures are utilized [1]. These measures, which perform an evaluation at a pixel level, include Accuracy (ACC), Sensitivity (SEN), Specificity (SPE), Dice Similarity Coefficient (DSC), and Jaccard index (JAC).

Definition of ACC, SEN, SPE, DSC, and JAC are described by Equations 1 to 5, respectively. These measures are calculated based on four quantifiers of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). TP is the number of pixels that the proposed segmentation method correctly identified as belonging to a lesion area; TN is the number of pixels that the proposed segmentation method correctly identified as belonging to a normal area; FP is the number of pixels that the proposed segmentation method wrongly identified as belonging to a lesion area, and FN is the number of pixels that the proposed segmentation method wrongly identified as belonging to a normal area. Moreover, SEN and SPE quantify the performance of the proposed segmentation methods on identifying pixels belonging to the abnormal regions and the pixels belonging to the normal areas, respectively. Besides, DSC and JAC attempt to summarize SEN and SPE in one unique measure capable of reflecting both SEN and SPE.

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$SEN = \frac{TP}{TP+FN} \quad (2)$$

$$SPE = \frac{TN}{TN+FP} \quad (3)$$

$$DSC = \frac{2\times TP}{2\times TP+FN+FP} \quad (4)$$

$$JAC = \frac{TP}{TP+FN+FP} \quad (5)$$

The quantitative performance of the proposed method is measured for all of the images in the test-set. With this regard, all of the mentioned performance measures are calculated for each image individually by incorporating a ground-truth mask and the lesion mask produced by the proposed segmentation method.

Furthermore, alongside the segmentation network architecture, the view sizes, and the threshold value of the likelihood map are critical components of the proposed method to be validated for the best performance. In the paragraphs ahead, we will discuss how the parameters of the proposed method are selected and tuned; later, the results with the best settings are reported, discussed, and compared to the state-of-the-art.

3.1. Training of the Segmentation Network

The segmentation network, with the structure, explained in section 2.2.3, is the heart of MMTDNN. As a result, proper training of this network is mandatory for better performance. The input of the segmentation network is a tensor. As explained in section 2.2.2, each tensor provides information for the description of a typical pixel. Accordingly, for each image in the training set, a set of tensors is extracted to be used as training data of the segmentation network. Furthermore, 20 percent of the training images are selected as the validation set to monitor the training process. To set the network weights properly, Stochastic Gradient Descends (SGD) is employed as an optimizer [28]. Important parameters of this optimizer are the learning rate and the momentum which has been set to 0.01 and 0.9, respectively. SGD optimized the network in a total of 40 epochs; in the last 15 epochs, no more significant improvement on the accuracy of the validation set has occurred. Therefore, the training is ended at epoch 40 in its best situation. Continuing training for more epochs leads the training accuracy to improve more and the validation accuracy to decline. In other words, in the last 15 epochs, the accuracy of the validation become oscillatory and had started to reverse from its optimal condition. In this situation, where the accuracy of the training set continues to improve, the training process is stopped in its best point to prevent the network from overfitting.

3.2. The Effect of View Size on the Segmentation Network

The size of views is an important parameter that affects the performance of the segmentation network; and consequently, the performance of the proposed MMTDNN. To choose the best view size, the performance of the segmentation network under views of various sizes is evaluated. Meanwhile, the accuracy of the segmentation network for view sizes of 25×25, 31×31, and 35×35 are measured. As is summarized in Table 1, the best view size is 31×31. This size gives the highest possible amount of accuracy on the validation set. Performance of the best size is almost near the performance of the views of size 25×25 and far from the views of size 35×35; this means that the more extensive views cannot perform as accurately as small views.

Table 1. Effect of view size on the performance of the segmentation network

Window Size	25×25	31×31	35×35
Accuracy	0.961	**0.975**	0.952

3.3. Evaluation of the Proposed Method on the Test Set

After finding the best settings of different components of the proposed segmentation method, its performance has been evaluated on the test set. In this round of evaluation, all of the previously mentioned measures are calculated for each image in the test set. These results, which are averaged over all of the testing images, are summarized in Table 2.

In terms of accuracy, the proposed method performs satisfactorily; this means that the segmentation network correctly learns the patterns that are necessary for distinguishing lesion and non-lesion pixels. The high specificity of MMTDNN demonstrates that the proposed method can accurately identify nearly all normal pixels of the testing images as normal. In contrast to specificity, the sensitivity of MMTDNN is low; this means that the proposed method fails to correctly identify some of the pixels belonging to the lesion areas. Overall, the proposed method performs well in terms of quantitative measures, including ACC, SEN, SPE, DSC, and JAC. This satisfactory performance is the result of learning various patterns by the segmentation network. In other words, the segmentation network is trained with numerous profiles of both normality and abnormality thanks to the massive training strategy.

Table 2. Performance of the proposed MMTDNN on the test-set

Performance Measure	ACC	SEN	SPE	DSC	JAC
Value	0.973	0.912	0.986	0.931	0.876

Figure 4 demonstrates a set of challenging examples, in which the proposed automatic segmentation method of skin lesions performs well. In this set, Figures 4(a), 4(b), and 4(c) show examples with massive hair artifacts; Figures 4(g) and 4(h) depict examples with color charts; Figures 4(c) and 4(d) demonstrate examples of low contrast between the lesion and normal areas; Figures 4(i) and 4(j) represent examples of inhomogeneous appearance; Figure 4(k) and 4(l) display examples with fuzzy borders. Overall, the proposed method is almost robust to all of these challenging circumstances; consequently, it can accurately delineate nearly all lesion areas in these situations.

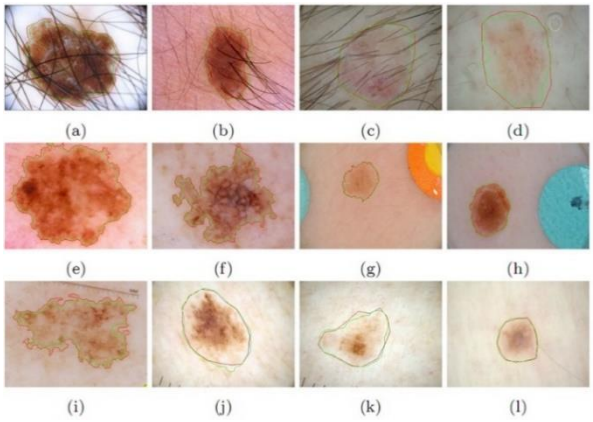

Figure 4. Examples that MMTDNN correctly segments the lesion areas. The red and green contours indicate the segmentation results of the ground truth and MMTDNN, respectively

Although the proposed method can achieve satisfactory performance in delineating lesion areas in examples with the previously mentioned challenging conditions, there are still some limitations. The proposed method, sometimes over-segment or under-segment the lesion areas. As Figure 5 depicts, usually, the cases with low contrast, unsharp borders, and irregular shapes suffer from the adverse phenomena of under/over-segmentation.

Figure 5. Examples that the proposed method over-segments or under-segments lesion areas. The red and green contours indicate the segmentation results of the ground truth and MMTDNN, respectively

The main reason for this shortcoming is that the proposed segmentation network cannot learn the discriminative patterns in these examples as well as other examples. Increasing the network capacity and proper post-processing techniques may alleviate these problems.

The main reason for this shortcoming is that the proposed segmentation network cannot learn the discriminative patterns in these examples as well as other examples. Increasing the network capacity and proper post-processing techniques may alleviate these problems.

3.4. Comparing the Results of MMTDNN with the Results of Other Methods

In this section, we shall compare the performance of the proposed method with some of the existing representative methods of skin lesion segmentation. For the sake of fair comparison, only those studies that use ISBI-2016 dataset in their evaluation procedure are considered. The results of this comparison in terms of quantitative measures are summarized in Table 3. As it is shown in this table, the proposed MMTDNN outperforms all of these methods of skin lesion segmentation.

In terms of ACC, DSC, JAC, and SPE, our method performs superior to other methods in this table. In other words, MMTDNN reduces the error rate (1-ACC) of the segmentation by 27 percent. As the segmentation of abnormal tissues is a class-imbalance problem, JAC can better describe the performance of MMTDNN. Therefore, we insist that MMTDNN improves the gap (1-JAC) to the best JAC value by 38 percent. Besides this excellent performance, the proposed method performs poorly in terms of SEN

Table 3. Comparison of MMTDNN results with other methods

Method	ACC	SEN	SPE	DSC	JAC
Fully CNN with Jaccard Distance [17]	0.963	**0.926**	0.971	0.922	0.861
CUMED [29]	0.949	0.911	0.957	0.897	0.829
Multi-stage Fully CNN [30]	0.955	**0.922**	0.965	0.912	0.846
MTANN [7]	0.861	0.790	0.847	0.713	0.580
MMTDNN	**0.973**	0.912	**0.986**	**0.931**	**0.876**

measure. In other words, MMTDNN is less sensitive than the two methods highlighted in Table 3. Lower sensitivity indicates that the proposed method performs worse than the two mentioned methods, to identify the pixels belonging to the lesion areas correctly. Overall, MMTDNN performs very satisfactorily in comparison to the methods that have been evaluated on the ISBI-2016 dataset.

This study aimed to increase the representation capability of MTANN by providing more powerful architecture based on CNNs. In order to demonstrate the superiority of MMTDNN over MTANN, we perform an experiment and provide the results in Table 3. As was expected, MTANN performs very poorly.

4. Conclusion

In this study, we have proposed and developed a pattern recognition technique mainly based on artificial neural networks, termed multi-view massive training deep neural network. The proposed method is, in fact, an extension of both MTANN and CNN in a way to better handling multi-view or tensor input. In other words, the proposed method attempts to gain the benefits of both CNN and MTANN. CNN increases representation capability than MTANN by automatically learning sophisticated and relevant features directly from the data.

In another perspective, the proposed method is suitable for cases that the training data are rare; in such situations, too many training samples are generated thanks to the massive training strategy. Also, a massive training strategy prevents the network from overfitting and leading the network to be more generalized. The key to this high generalizability might be due to the division of one image into a large number of overlapped sub-regions.

In this study, the performance of the proposed method has been evaluated on the task of automated skin lesion segmentation. As experimental results demonstrate, the proposed method outperforms state-of-the-art methods of skin lesion segmentation. Moreover, MMTDNN is robust to hair artifacts, shape irregularities, appearance inhomogeneity, and fuzzy boundaries. The main reason for the prosperity of MMTDNN is its high capability of learning discriminative features in the multi-view neighborhood of image pixels. Moreover, the proposed method is very promising to be applied to other tasks of abnormality detection and segmentation in medical images.

The performance of the proposed method can be improved by incorporating more views in the view generation process and also increasing the capacity of the model by creating a deeper segmentation network and adding more feature maps to the convolutional layers. Also, the proposed MMTDNN can be implemented in parallel to be run faster.

In our future works, we will increase the capacity of the segmentation network to reduce network error. Moreover, generating more informative and compliment views is another effort for improving performance of the proposed method.

References

1-H. Khastavaneh and H. Ebrahimpour-komleh, "Automated Segmentation of Abnormal Tissues in Medical Images," *J. Biomed. Phys. Eng.*, Jun. 2019.

2-H. Khastavaneh and H. Ebrahimpour-Komleh, "Representation Learning Techniques: An Overview," in *Data Science: From Research to Application*, Springer, 2020.

3- Y. Bengio, A. Courville, and P. Vincent, "Representation Learning: A Review and New Perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 35, No. 8, pp. 1798–1828, Aug. 2013.

4- K. Suzuki, "Overview of deep learning in medical imaging," *Radiol. Phys. Technol.*, Vol. 10, No. 3, pp. 257–273, 2017.

5- K. Suzuki, S. G. Armato, F. Li, S. Sone, and K. Doi, "Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose computed tomography," *Med. Phys.*, vol. 30, no. 7, pp. 1602–1617, Jun. 2003.

6- H. Khastavaneh and H. Haron, "A Conceptual Model for Segmentation of Multiple Scleroses Lesions in Magnetic Resonance Images Using Massive Training Artificial Neural Network," in *2014 5th International Conference on Intelligent Systems, Modelling and Simulation*, 2014, vol. 2015-Septe, pp. 273–278.

7- H. Khastavaneh and H. Ebrahimpour-Komleh, "Neural Network-Based Learning Kernel for Automatic Segmentation of Multiple Sclerosis Lesions on Magnetic Resonance Images.," *J. Biomed. Phys. Eng.*, vol. 7, no. 2, pp. 155–162, Jun. 2017.

8- K. Suzuki, H. Yoshida, J. Näppi, S. G. Armato, and A. H. Dachman, "Mixture of expert 3D massive-training ANNs for reduction of multiple types of false positives in CAD for detection of polyps in CT colonography," *Med. Phys.*, vol. 35, no. 2, pp. 694–703, Jan. 2008.

9- J.-W. Xu and K. Suzuki, "Massive-training support vector regression and Gaussian process for false-positive reduction in computer-aided detection of polyps in CT colonography," *Med. Phys.*, vol. 38, no. 4, pp. 1888–1902, Mar. 2011.

10- K. Suzuki, H. Yoshida, J. Näppi, and A. H. Dachman, "Massive-training artificial neural network (MTANN) for reduction of false positives in computer-aided detection of polyps: Suppression of rectal tubes," *Med. Phys.*, vol. 33, no. 10, pp. 3814–3824, Sep. 2006.

11- K. Suzuki, Jun Zhang, and Jianwu Xu, "Massive-Training Artificial Neural Network Coupled With Laplacian-Eigenfunction-Based Dimensionality Reduction for Computer-Aided Detection of Polyps in CT Colonography," *IEEE Trans. Med. Imaging*, vol. 29, no. 11, pp. 1907–1917, Nov. 2010.

12- K. Suzuki, H. Abe, H. MacMahon, and K. Doi, "Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (MTANN)," *IEEE Trans. Med. Imaging*, vol. 25, no. 4, pp. 406–416, Apr. 2006.

13- S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1240–1251, May 2016.

14- M. Havaei *et al.*, "Brain tumor segmentation with Deep Neural Networks," *Med. Image Anal.*, vol. 35, pp. 18–31, Jan. 2017.

15- K. H. Cha, L. Hadjiiski, R. K. Samala, H. Chan, E. M. Caoili, and R. H. Cohan, "Urinary bladder segmentation in CT urography using deep-learning convolutional neural network and level sets," *Med. Phys.*, vol. 43, no. 4, pp. 1882–1896, Mar. 2016.

16- M. H. Jafari, E. Nasr-Esfahani, N. Karimi, S. M. R. Soroushmehr, S. Samavi, and K. Najarian, "Extraction of skin lesions from non-dermoscopic images for surgical excision of melanoma," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 6, pp. 1021–1030, Jun. 2017.

17- Y. Yuan, M. Chao, and Y.-C. Lo, "Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks With Jaccard Distance," *IEEE Trans. Med. Imaging*, vol. 36, no. 9, pp. 1876–1886, Sep. 2017.

18- T. Brosch, L. Y. W. Tang, Y. Yoo, D. K. B. Li, A. Traboulsee, and R. Tam, "Deep 3D Convolutional Encoder Networks With Shortcuts for Multiscale Feature Integration Applied to Multiple Sclerosis Lesion Segmentation," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1229–1239, May 2016.

19- H. Li *et al.*, "Fully convolutional network ensembles for white matter hyperintensities segmentation in MR images," *Neuroimage*, vol. 183, no. June, pp. 650–665, Dec. 2018.

20- T. de Moor, A. Rodriguez-Ruiz, A. G. Mérida, R. Mann, and J. Teuwen, "Automated soft tissue lesion detection and segmentation in digital mammography using a u-net deep learning network," Feb. 2018.

21- H. Khastavaneh and H. Ebrahimpour-Komleh, "On Multi-view Interpretation of Convolutional Neural Networks," in *2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI)*, 2019, pp. 587–591.

22- D. Gutman *et al.*, "Skin Lesion Analysis toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC)," pp. 3–7, 2016.

23- V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning," *Arxiv*, pp. 1–28, 2016.

24- S. Ioffe and C. Szegedy, "Batch Normalization:

Accelerating Deep Network Training by Reducing Internal Covariate Shift," in *Proceedings of the 32nd International Conference on Machine Learning, PMLR*, 2015, vol. 37, pp. 448–456.

25- N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout : A Simple Way to Prevent Neural Networks from Overfitting," *J. Mach. Learn. Res.*, Vol. 15, no. 1, pp. 1929–1958, 2014.

26- K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

27- N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Trans. Syst. Man. Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.

28- L. Bottou, "Large-Scale Machine Learning with Stochastic Gradient Descent," in *Proceedings of COMPSTAT'2010*, Heidelberg: Physica-Verlag HD, 2010, pp. 177–186.

29- L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A. Heng, "Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks," *IEEE Trans. Med. Imaging*, vol. 36, no. 4, pp. 994–1004, Apr. 2017.

30- L. Bi, J. Kim, E. Ahn, A. Kumar, M. Fulham, and D. Feng, "Dermoscopic Image Segmentation via Multistage Fully Convolutional Networks," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 9, pp. 2065–2074, Sep. 2017.