A COMPARATIVE ANALYSIS ON DEEP LEARNING ALGORITHMS FOR LARGE OUTPUT SPACES

Bhimsen Moharana[1], **Dr. Jitendra Sheetlani**[2], **Dr. J. P. Patra**[3]
[1]Ph.D Scholar, Dept. of Computer Science, Sri Satya Sai University ofTechnology and MedicalSciences, Sehore, MP
[2]Professor, Dept. of Computer Science, Sri Satya Sai University of Technologyand MedicalSciences, Sehore, MP
[3]Professor, Dept. of CSE, Shri Shankaracharya Institute of Professional Managementand Technology, Raipur, Chhattisgarh
bhimasen.moharana@gmail.com, dr.jsheetalni@gmail.com, jp.patra@ssipmt.com

Abstract

Now a days, as the large amount of annotated medical data has been growing quickly, and giving more attention for the deep learning-based approaches and having a lot of achievement in the medical segmentation field, such as CAD and other things. A hierarchical representation of data in medical image recognition problems is able to learn by using Deep learning when it is used in biologically-inspired architectures. This helps it learn how to distinguish between different image types. In other words, if the discriminative info is only found in small parts of the image, an existing classic deep learning framework may still have problems finding them without local-level annotations. In this paper, we show how to use "a Novel multi-Phasebased deep learning framework" to find local discriminative details for medical image segmentation that can be found in large-scale output space.

Keywords: Discriminative local Data discovery, Multi-Phase,CNN

I.Introduction

In medical image analysis, which uses a lot of output space, the main goal is to show and extract clinical info. This is the foundation for a lot of complex frameworks. A lot of algorithms with automatic or semi-automatic have been made in this group [1–4]. When clinicians or researchers look at medical images, they can use these tools to help them understand and evaluate them in a wide range of ways, from simple tasks like locating anatomical landmarks and dividing an organ

into parts [5,6] to more complex computer-aided diagnosis (CAD) systems [10–12]. Because different organ systems have very different characteristics, medical image analysis models are usually trained or built for specific anatomies so they can use the right info from previous research.

Deep learning [13] methods have been used in a wide range of applications, including object recognition [15], natural language processing,signal processing [14], and more. They are attracted to a deep network architecture because it can automatically learn new features without learning to do anything. The deep network architecture is made up of many layers of activation functions that are simple but not linear. These activation functions change the input data into different Phase of featuresdepiction, from the level of low to high. The network can figure out how to make these kinds of representations from a lot of training data on its own. In [15] Image recognition, which is one of the veryhopeful applications. Deep learning acquires how to represent visual features in a way that is clear and useful. Deep learning, on the other hand, must be able to help with the large amount of medical image segmentation, because it has a lot of data items to work with. If you want to compare the progress of deep learning-based algorithms for things like medical image recognition, this is the paper for you. In this paper, we'll look at a medical image recognition algorithm that uses multiple phases and multiple instances of deep learning.

II.Methodology

A novel multi-phase based deep learning is used in this study to "discover" discriminative and non-informative local regions for exact image segmentation without the need for time span consuming local manual annotations. The classifier is also learned in the framework. Multi-instance learning is used to learn a CNN in the first step. This helps the system find the patches that are most discriminative in the area. Specifically, each image is broken up into a number of local areas. The deep network then gets a group of labelled images (bags), each of which has a lot of local patches in it (instances). The CNN’s loss function is chosen so that as long as one local patch (instance) is correctly classified, the labelling of the corresponding image slice (bag) is thought to be correct, even if it is not. When the pre-trained CNN sees the discriminative local patches, it will be more likely to respond to them than other parts of the image. Pre-trained

CNNs show us which local patches are discriminative and which ones aren't. These patches are used to help the pre-trained CNN get even better in the nextphase of my learning framework, which is called the "boosting phase."When the run-time (dynamic level), a sliding window method is used to put on the boosted CNN to the target image. In this way, the CNN classifies a body part by concentrating on the utmostdistinguishing local info and discarding the less important local areas.

A training set T= {X_m, m=1, ……, M}with matching labels l_m. Everytaken training image sets, X_m, is separated into a group of local patches declared as$\mathcal{L}(\mathbf{X}_m) = \{\mathbf{x}_{mn}, n = 1, \ldots, N\}$ These areturn out to be the basic training samples of the CNN and they have the same labels as the original images. i.e., all $\mathbf{x}_{mn} \in \mathcal{L}(\mathbf{X}_m)$part the same label l_m When the construction of CNN is still the similar as the normal one, the loss function is

$$L_2(\mathbf{W}) = \sum_{\mathbf{X}_m \in \mathcal{T}} -\log(\max_{\mathbf{x}_{mn} \in \mathcal{L}(\mathbf{X}_m)} \mathbf{P}(l_m|\mathbf{x}_{mn}; \mathbf{W})), \quad (1)$$

Here $\mathbf{P}(l_m|\mathbf{x}_{mn}; \mathbf{W})$is the probability that the local patch $\mathbf{x}_{mn}$, is properlydevided as l_m using CNN coefficients W.

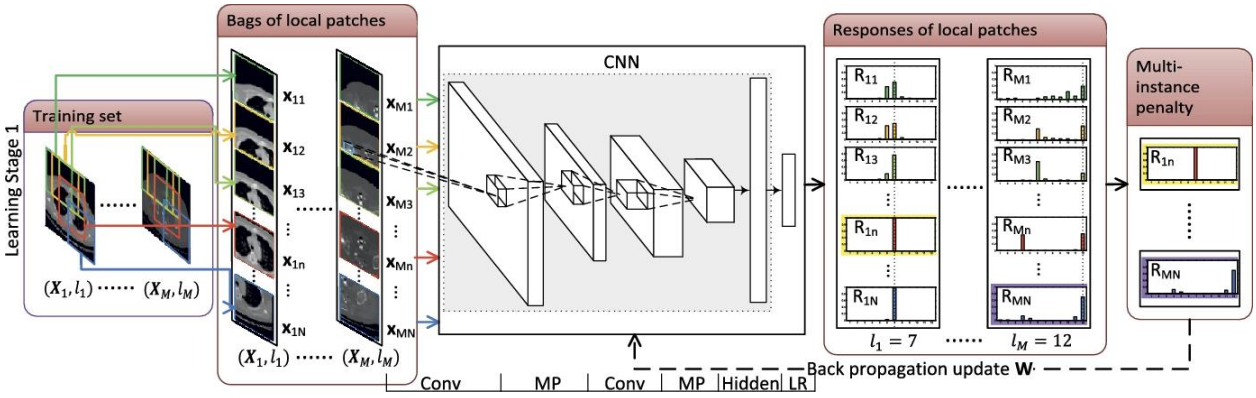

FIGURE 1.Design of the pre-training phase.

We build a sample to show this findingcapability, in figure 2.A depicts the random placement and combination of 4 categories of geometry elements, namely circles,squares, diamonds, and trianglesto produce binary images of two different sorts. When diamond and circle are permitted

to seem in any classes, square and triangle are strictly kept by Class 2 and 1, correspondingly. Figure 2.B illustrates the exposed discriminative patches (including square or triangle) for the image segmentation job in toy example. This is in line with the fact that "triangle" and "square" are the only ways to tell these two classes apart. Thisdemonstrates that present algorithm is able to determine the key local patches without classification. Of course, this problem will become trivial if we have prior knowledge of the exclusionary subspaces and design particular classifiers on them. However, in real-world identification tasks, it is not easy to find out the most discriminative local patches for diverse classes. Annotating local patches and training local classifiers can be time consuming even with ad hoc knowledge. The solution consequently becomes non-scalable.

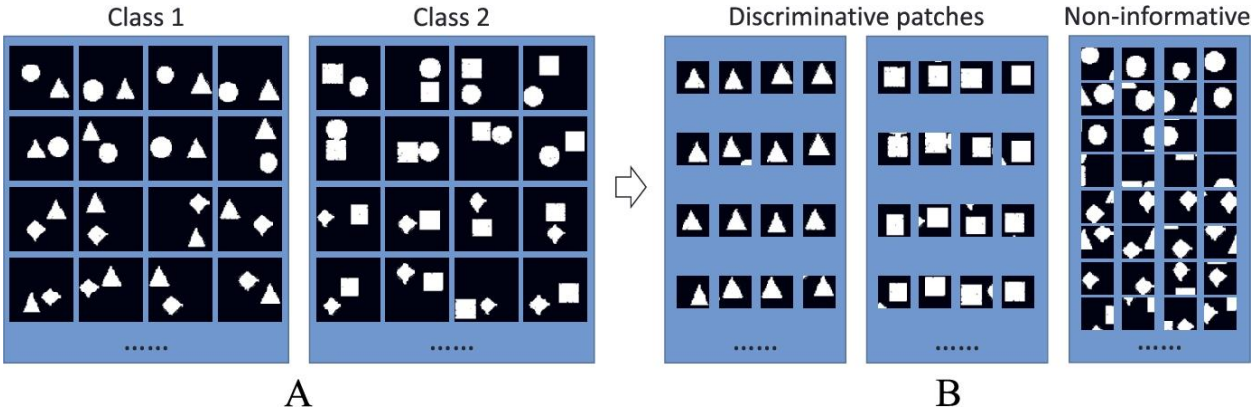

FIGURE 2. A mock (sample)sample. (A) Artificialpictureswith two classes. (B) Selections by the pretrained CNN model of discriminative yet illuminating local patch data

To ensure that the learnt CNN would have sustained high responses on exclusionary local patches, a spatially consistency component is additionally included into loss function as

$$L_3(\mathbf{W}) = \sum_{\mathbf{X}_m \in \mathcal{T}} -\log\Big(\max_{\mathbf{x}_{mn} \in \mathcal{L}(\mathbf{X}_m)} \sum_{\mathbf{x} \in \mathfrak{N}(\mathbf{x}_{mn})} \mathbf{P}(l_m|\mathbf{x};\mathbf{W})\Big). \quad (2)$$

Here, $\mathfrak{N}(\mathbf{x}_{mn})$, declared the local-patches in the nearest region of $\mathbf{x}_{mn}$. At theexecution period, the enhanced CNN is utilised for body-part identification in a sliding window method. The test image X is divided into local patches of N overlapping by the sliding window $\mathcal{L}(\mathbf{X}) = \{\mathbf{x}_n, n = 1, \ldots, N\}$. For each local patch $\mathbf{x}_n$, the boosted CNN outputs a response vector

with, $K+1$, components $\{\mathbf{P}(k|\mathbf{x}_n; \mathbf{W}^{opt})|k = 1, \ldots, K+1\}$, where $\mathbf{W}^{opt}$ denotes the optimal coefficients of Eq. (3). The class of the local patch $\mathbf{x}_n$ is then determined as,

$$c(\mathbf{x}_n) = \underset{k\in\{1,\ldots,K+1\}}{\operatorname{argmax}} \mathbf{P}(k|\mathbf{x}_n; \mathbf{W}^{opt}). \quad (3)$$

Meanwhile the class $K+1$ isan intentionally built semi ones, local patches belongs to this class must be omitted in body segment identification. The maximum discriminative patch $\mathbf{x}_{n*}$ in the picture is chosen as being the most twitchy properly labelled one with the exception of the non-informative patches:

$$\mathbf{x}_{n*} = \underset{\mathbf{x}_n\in\mathcal{L}(\mathbf{X});c(\mathbf{x}_n)\neq K+1}{\operatorname{argmax}} \mathbf{P}(c(\mathbf{x}_n)|\mathbf{x}_n; \mathbf{W}^{opt}). \quad (4)$$

A tough decision of labeled fusion is integrating the probability for each class in the region of the most exclusionary patch:

$$C(\mathbf{X}) = \underset{k\in\{1,\ldots,K\}}{\operatorname{argmax}} \sum_{\mathbf{x}_n\in\mathfrak{N}(\mathbf{x}_{n*})} \mathbf{P}(k|\mathbf{x}_n; \mathbf{W}^{opt}). \quad (5)$$

III.Result

Four distinct kinds of CNN are compared: (a) (SCNN)standard CNN, as shown in Fig. 1, trained on the whole picture; (b) (PCNN)local patch-based CNN without boost, i.e., the CNN trained by well before phase only; (c) (BCNN1)local patch-based CNN boosted without increase in non class; (d) (BCNN2)local patch-based CNN boosted with both discriminative and non-informative patches. Method I depict typical CNN learning (using features extracted from complete image) (by utilizing features extracting from entirepicture). Techniques (b) and (c) are two different versions of our suggested technique (d), which are offered to validate the impacts of every component of our strategy. There are 36 patches retrieved from each 60×60 image with a sliding window with 6-pixel step size.The patch size for all patch-based CNNs is 30×30.All CNNs employ the same middle architecture: 300 terminals with one hidden layer,a 2×2 kernel with one max-pooling layer,10 5×5 filters with one convolutional layer, and for the output response, it is followed by an LR layer.

PCNN outperforms SCNN by 16% when local discriminative info is used, as seen in Table 1. It suggests that typical CNN would not fully uncover and study the discriminative local patches, "square" and "triangle". As demonstrated in Fig. 2, the CNN with multi-instance learning effectively finds the most discriminative and non-informative local patches. As a result of overfitting on discriminative patches (because BCNN1's parameters are started by those of PCNN and improved by training with the extracted discriminative patches exclusively), BCNN1 is poorer than PCNN (because the parameters of BCNN1 are initialised by those of PCNN, and refined by training with the extracted discriminative patches only). BCNN2 is the best performing BCNN1 with discriminative and non-informative patches training.

Table 1: Segmentation accuracies (%) on mock data set

	Recall			Precision			F1		
Class	1	2	Total	1	2	Total	1	2	Total
SCNN	84.2	82.4	83.3	82.7	83.9	83.3	83.5	83.2	83.3
PCNN	99.6	99.7	99.7	99.7	99.6	99.7	99.7	99.7	99.7
BCNN1	98.4	99.7	99.1	99.7	98.4	99.1	99.0	99.1	99.1
BCNN2	100	100	100	100	99.9	100	100	100	100

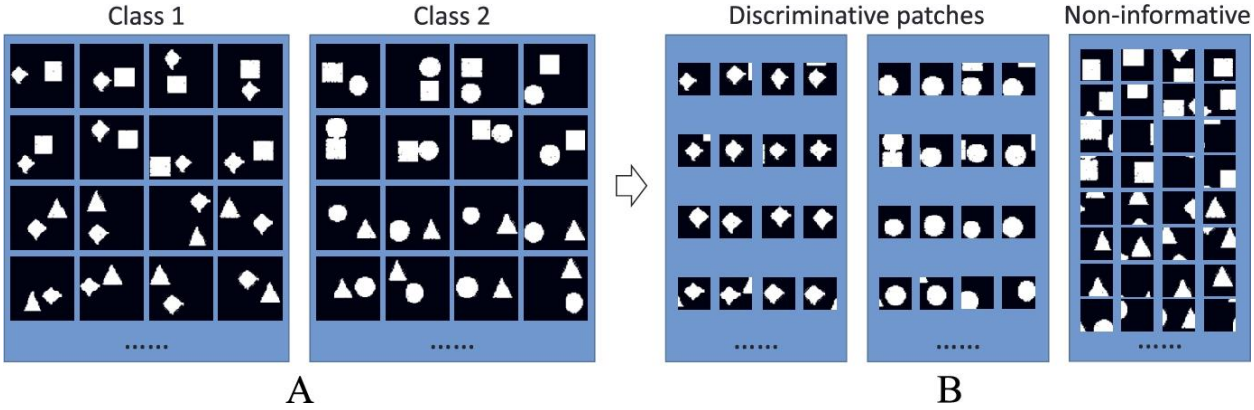

FIGURE 3. The second artificial sample. (A) Mockpictures of two classes separated by diamond and circle. It is crucial to note that we used the identical samples as in Fig. 3, however re-labeled the photos into two classes based on various principles. (B) The exclusionary and non-informative local patches found by the pre-trained CNN model.

So I relabelled the simulated data using diamond and circle as distinguishing components in Class 1 and 2. In other words, while the dataset is identical, the local patches used to distinguish

the two classes are distinct. In real-world problems, the datasets are identical, but the classification goal is altered. After the first learning phase, the trained CNN was extracted for local patches. Again, the retrieved local patches have the diamond and circle features. To demonstrate our multi-instance CNN learning's discriminative local regions for classification problems, we used this result.

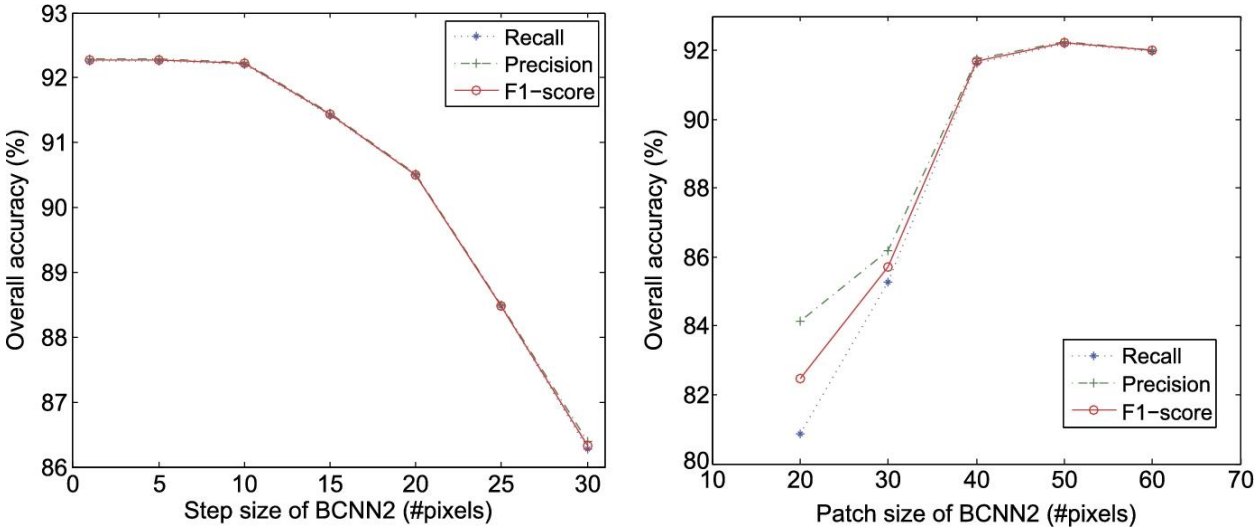

Figure 4.Performance assessments on the relevance of variables in BCNN2. (Left) Segmentation accuracies vs. iterations of sliding window. (Right) Segmentation accuracies vs. patch size.

As shown in Fig. 4, (right)the influence of patch size to the segmentation accuracy is also explored. We can observe from the plot that I the patch size will not be too short to capture the unlabeled data (size 20 or 30); (ii) the efficiency is not very sensitive to the local patch size once it is large enough to accommodate discriminative info (sizes from 40 to 60 in present work) (sizes from 40 to 60 in this task).As among the critical variables in BCNN2 algorithm, step size of sliding window testing is explored referring to the accuracies. The running times for step sizes 1, 5, 10, 15, 20, 25, and 30 pixels are 541.1, 30.6, 11.7, 5.3, 5.2, 3.6, and 3.4 ms per image, accordingly. Regarding the remaining time complexity and precision, step size 10 or 15 is an acceptable choice in this assignment.

IV.Conclusion

I created a novelmulti-phase based deep learning framework for big scale medical image segmentation. Using a noveldeep learning technique, it identifies the useful and useless local patches, and then applies our multi-phase learning strategies to train CNN to recognize images. From the authentications on mock (sample) dataset and a large-scale CT dataset, we can saw clear gains compared with the other deep learning methods. It is demonstrated that the success of the suggested algorithm against the standard CNN does not derive from more enhanced training samples (seems the outcomes of our novel one). Thismay helpful in the medical image preparation by gating the applicable auto-algorithms before being loaded for manual reading. In this method, the automatic and meaningful outcomes can be showed quickly in the reading room to speed-up radiologists' reading process. It is worth emphasizing that since no manual annotations are necessary to designate these local patches, our solution becomes incredibly scalable. It might be beneficial to combine algorithms like multi-scale convolution or multi-scale image patch to locate and detect various discriminative local sections in multiple categories.

V.References

1. T. McInerney, D. Terzopoulos, Deformable models in medical image analysis: a survey, Med. Image Anal. 1996;1(2):91–108.
2. J.A. Maintz, M.A. Viergever, A survey of medical image registration, Med. Image Anal. 1998;2(1):1–36
3. D.L. Pham, C. Xu, J.L. Prince, Current methods in medical image segmentation, Annu. Rev. Biomed. Eng. 2000;2(1):315–337.
4. D.L. Hill, P.G. Batchelor, M. Holden, D.J. Hawkes, Medical image registration, Phys. Med. Biol. 2001;46(3):R1.
5. M. Betke, H. Hong, D. Thomas, C. Prince, J.P. Ko, Landmark detection in the chest and registration of lung surfaces with an application to nodule registration, Med. Image Anal. 2003;7(3):265–281.
6. Y. Zheng, M. John, R. Liao, J. Boese, U. Kirschstein, B. Georgescu, S.K. Zhou, J. Kempfert, T. Walther, G. Brockmann, et al., Automatic aorta segmentation and valve landmark detection in C-arm CT: application to aortic valve implantation, In: Medical Image Computing and Computer-Assisted Intervention. Springer; 2010:476–483.

7. D. Shen, S. Moffat, S.M. Resnick, C. Davatzikos, Measuring size and shape of the hippocampus in MR images using a deformable shape model, NeuroImage 2002;15(2):422–434.
8. H. Ling, S.K. Zhou, Y. Zheng, B. Georgescu, M. Suehling, D. Comaniciu, Hierarchical, learning-based automatic liver segmentation, In: IEEE International Conference on Computer Vision and Pattern Recognition. IEEE; 2008:1–8.
9. C. Li, R. Huang, Z. Ding, J.C. Gatenby, D.N. Metaxas, J.C. Gore, A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI, IEEE Trans. Image Process. 2011;20(7):2007–2016.
10. R. Bellotti, F. De Carlo, G. Gargano, S. Tangaro, D. Cascio, E. Catanzariti, P. Cerello, S.C. Cheran, P. Delogu, I. De Mitri, et al., A CAD system for nodule detection in low-dose lung CTs based on region growing and a new active contour model, Med. Phys. 2007;34(12):4901–4910.
11. L.A. Meinel, A.H. Stolpen, K.S. Berbaum, L.L. Fajardo, J.M. Reinhardt, Breast MRI lesion classification: improved performance of human readers with a backpropagation neural network computer-aided diagnosis (CAD) system, J. Magn. Reson. Imaging 2007;25(1):89–95.
12. K. Doi, Current status and future potential of computer-aided diagnosis in medical imaging, Br. J. Radiol. 2005;78(suppl_1):s1–s19.
13. Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 2015;521(7553):436–444.
14. G. Hinton, L. Deng, D. Yu, G.E. Dahl, A.r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T.N. Sainath, et al., Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups, IEEE Signal Process. Mag. 2012;29(6):82–97.
15. A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, In: Advances in Neural Info Processing Systems. 2012:1097–1105.