ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 2, February 2019

An Improved Anatomically Constrained Neural Networks (ACNNS) : Application To Cardiac Image Enhancement And Segmentation Using Fuzzy Based Super Pixel Model

M.Divya M.E[#1], M.Nanthini Muthuvalli[#2], M.Rathi Priya[#3],

Assistant Professor, Department of CSE, ULTRA College of Engineering & Technology, Madurai, Tamilnadu, India[#1]

UG Scholar, Department of CSE, ULTRA College of Engineering & Technology, Madurai, Tamilnadu, India[#2,#3]

ABSTRACT: Incorporation of prior knowledge about organ shape and location is key to improve performance of image analysis approaches. In particular, priors can be useful in cases where images are corrupted and contain artefacts due to limitations in image acquisition. The highly constrained nature of anatomical objects can be well captured with learning based techniques. However, in most recent and promising techniques such as CNN based segmentation it is not obvious how to incorporate such prior knowledge. State-of-the-art methods operate as pixel-wise classifiers where the training objectives do not incorporate the structure and inter-dependencies of the output. To overcome this limitation, we propose a
generic training strategy that incorporates anatomical prior knowledge into CNNs through a new regularisation model, which is trained end-to-end. The new framework encourages models to follow the global anatomical properties of the underlying anatomy (e.g. shape, label structure) via learned non-linear representations of the shape. We show that the proposed approach can be easily adapted to different analysis tasks (e.g. image enhancement, segmentation) and improve the prediction accuracy of the state-of-the-art models. The applicability of our approach is shown on multi-modal cardiac datasets and public benchmarks. Additionally, we demonstrate how the learned deep models of 3D shapes can be interpreted and used as biomarkers for classification of cardiac pathologies

I. INTRODUCTION

This chapter provides basic introduction about digital image processing and segmentation techniques involved in this research work. The field of digital image processing refers to processing digital images by means of a digital computer. Note that a digital image is composed of a finite number of elements, each of which has a particular location and value. These elements are referred to as picture elements, image elements, pels and pixels. Pixel is the term most widely used to denote the elements of a digital image. The areas of application of digital image processing are so varied that some form of organization is desirable in attempting to capture the breadth of this field. One of the simplest ways to develop a basic understanding of the extent of image processing applications is to categorize images according to their source (e.g.,visual,X-ray,and so on),Gonzalez et al (2000) comment in their work. AlZubi et al (2010) comment their work by Shadi AlZubi et al (2010) the use of 3D image processing has been increased especially for medical applications; this leads to increase the qualified radiologists' number who navigate, view, analyze, segment, and interpret medical images.

II. EXISITNG APPROACH

In the next section, we briefly summarise the state-of-theart methodology for image segmentation (SEG) and superresolution (SR), which is based on convolutional neural networks (CNNs). We then present a novel methodology

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

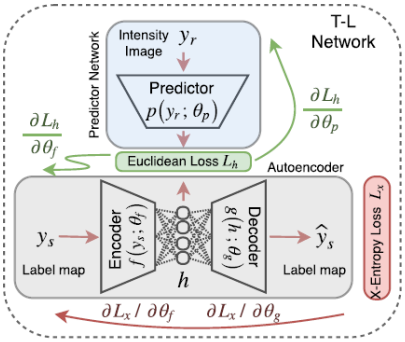
International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijareeie.com

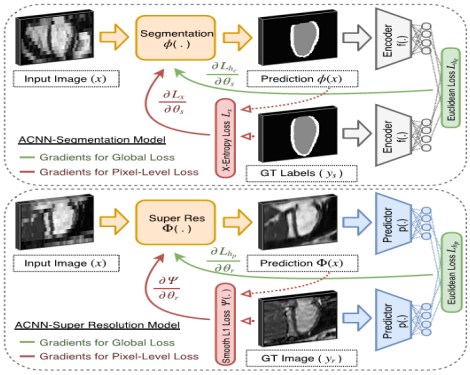
Vol. 8, Issue 3, March 2019

that extends these CNN models with a global training objective to constrain the output space by imposing anatomical shape priors. For this, we propose a new regularization network that is based on the T-L architecture which was used in computer graphics [19] to 3D render objects from natural images. Medical Image Segmentation With CNN Models Let ys = {yi}i∈S be an image of class labels representing different tissue types with yi∈ L = {1, 2,...C}. Furthermore let x = {xi ∈R,i∈ S} be the observed intensity image. The aim of image segmentation is to estimate ys having observed x. In CNN based segmentation models [22], [29], [38], this task is performed by learning a discriminative function that models the underlying conditional probability distribution P(ys|x). The estimation of class densities P(ys|x) consists in assigning to each xi the probability of belonging to each of the C classes, yielding C sets of class feature maps fc that are extracted through learnt non-linear functions. The final decision for class labels is then made by applying softmax to the extracted class feature maps, in the case of crossentropy $L_x = -\sum_{c=1}^{C} \sum_{i \in S} \log\left(\frac{e^{f(c,i)}}{\sum_j e^{f(j,i)}}\right)$ these featuremaps correspond to log likelihood values. As in the U-Net [38] and DeepMedic [22] models, we learnthe mapping between intensities and labelsφ(x): X→Lbyoptimising the average cross-entropy loss of each classLx = Cc=1L(x,c) using stochastic gradient descent. As shownin Fig. 2, the mapping function φis computed by passingthe input image through a series of convolution layers andrectified linear units across different image scales to enlargethe model's receptive field. The presented model is composed of two parts: feature extraction (analysis) similar to aVGG-Net and reconstruction(synthesis) as in the case ofa 3D U-Net [38]. However, in contrast to existing approaches,we aim for sub-pixel segmentation accuracy by trainingup-sampling layers with high-resolution ground-truth maps.This enables 3D analysis of the underlying anatomy in case ofthick slice 2D image stack acquisitions such as cine cardiacMR imaging.



Block diagram of the stacked convolutional autoencoder (AE) network (in grey), which is trained with segmentation labels. The AE model is coupled with a predictor network (in blue) to obtain a compact nonlinear representation that can be extracted from both intensity and segmentation images. The whole model is named as T-L network.

.



ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijareeie.com

Vol. 8, Issue 3, March 2019

Block diagram of the stacked convolutional autoencoder (AE) network (in grey), which is trained with segmentation labels. The AE modelis coupled with a predictor network (in blue) to obtain a compact nonlinear representation that can be extracted from both intensity andsegmentation images. The whole model is named as T-L network.

III. PROPOSED APPROACH

Each layer output in a convnet is a three-dimensional array of size h×w×d, where h and w are spatial dimensions, and d is the feature or channel dimension. The first layer is the image, with pixel size h × w, and d channels. Locations in higher layers correspond to the locations in the image they are path-connected to, which are called their receptive fields. Convnets are inherently translation invariant. Their basic components (convolution, pooling, and activation functions) operate on local input regions, and depend only on relative spatial coordinates. Writing xij for the data vector at location (i, j) in a particular layer, and yij for the following layer, these functions compute outputs yij by

$$\mathbf{y}_{ij} = f_{ks}\left(\{\mathbf{x}_{si+\delta i,sj+\delta j}\}_{0\le\delta i,\delta j<k}\right)$$

where k is called the kernel size, s is the stride or subsampling factor, and fks determines the layer type: a matrix multiplication for convolution or average pooling, a spatial max for max pooling, or an elementwise nonlinearity for an activation function, and so on for other types of layers. This functional form is maintained under composition, with kernel size and stride obeying the transformation rule

$$f_{ks} \circ g_{k's'} = (f \circ g)_{k'+(k-1)s',ss'}.$$

While a general net computes a general nonlinear function, a net with only layers of this form computes a nonlinear filter, which we call a deep filter or fully convolutional network. An FCN naturally operates on an input of any size, and produces an output of corresponding (possibly resampled) spatial dimensions

A real-valued loss function composed with an FCN defines a task. If the loss function is a sum over the spatial dimensions of the final layer, `(x; θ) = P ij ` 0 (xij ; θ), its parameter gradient will be a sum over the parameter gradients of each of its spatial components. Thus stochastic gradient descent on ` computed on whole images will be the same as stochastic gradient descent on ` 0 , taking all of the final layer receptive fields as a minibatch. When these receptive fields overlap significantly, both feedforward computation and backpropagation are much more efficient when computed layer-by-layer over an entire image instead of independently patch-by-patch.

Consider a layer (convolution or pooling) with input stride s, and a subsequent convolution layer with filter weights fij (eliding the irrelevant feature dimensions). Setting the earlier layer's input stride to one upsamples its output by a factor of s. However, convolving the original filter with the upsampled output does not produce the same result as shift-and-stitch, because the original filter only sees a reduced portion of its (now upsampled) input. To produce the same result, dilate (or "rarefy") the filter by forming

$$f'_{ij} = \begin{cases} f_{i/s,j/s} & \text{if } s \text{ divides both } i \text{ and } j; \\ 0 & \text{otherwise,} \end{cases}$$

.

(with i and j zero-based). Reproducing the full net output of shift-and-stitch involves repeating this filter enlargement layer-by-layer until all subsampling is removed. (In practice, this can be done efficiently by processing subsampled versions of the upsampled input.) Simply decreasing subsampling within a net is a tradeoff: the filters see finer information, but have smaller receptive fields and take longer to compute. This dilation trick is another kind of tradeoff: the output is denser without decreasing the receptive field sizes of the filters, but the filters are prohibited from accessing information at a finer scale than their original design. Although we have done preliminary experiments with dilation, we do not use it in our model. We find learning through upsampling, as described in the next section, to be effective and efficient, especially when combined with the skip layer fusion described later on. For further detail regarding dilation, refer to the dilated FCN

We define a skip architecture to extend FCN-VGG16 to a three-stream net with eight pixel stride shown in Figure 3. Adding a skip from pool4 halves the stride by scoring from this stride sixteen layer. The 2× interpolation layer of the skip is initialized to bilinear interpolation, but is not fixed so that it can be learned as described in Section 3.3. We call this twostream net FCN-16s, and likewise define FCN-8s by adding a further skip from pool3 to make stride eight predictions. (Note that predicting at stride eight does not significantly limit the maximum achievable mean IU; see

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

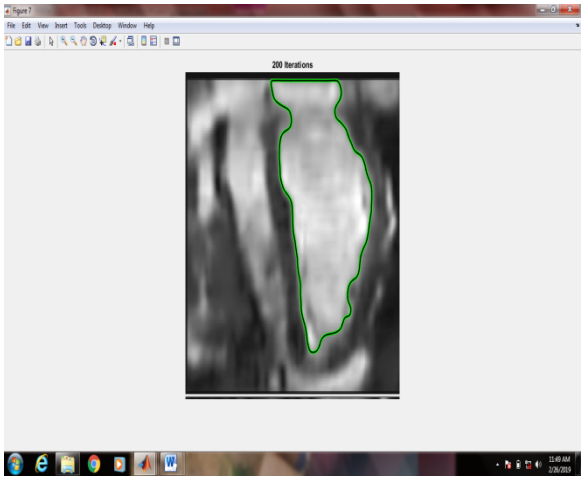
Website: www.ijareeie.com

Vol. 8, Issue 3, March 2019

Section 6.3.) We experiment with both staged training and all-at-once training. In the staged version, we learn the single-stream FCN-32s, then upgrade to the two-stream FCN-16s and continue learning, and finally upgrade to the three-stream FCN-8s and finish learning. At each stage the net is learned end-to-end, initialized with the parameters of the earlier net. The learning rate is dropped 100× from FCN-32s to FCN- 16s and 100× more from FCN-16s to FCN-8s, which we found to be necessary for continued improvements

We fine-tune all layers by backpropagation through the whole net. Fine-tuning the output classifier alone yields only 73% of the full fine-tuning performance as compared in Table 3. Fine-tuning in stages takes 36 hours on a single GPU. Learning FCN-8s all-at-once takes half the time to reach comparable accuracy. Training from scratch gives substantially lower accuracy.

IV. RESULT



V. CONCLUSION

We propose a method of segmentation with improved fcn alexnet propagation that is based on originally designed neural networks. By taking the contextual input into account, the spatial consistency of segmentation is enforced. Also, we conduct thorough and unprecedented testing to evaluate the generalization ability of our model and achieve performance better than or comparable to the state-of-the-art. Furthermore, With this new training objective, at testing time NNs make predictions that are in agreement with the learnt shape models of the underlying anatomy, which are referred as image priors

REFERENCES

[1] S. Barma, B.-W. Chen, W. Ji, F. Jiang, and J.-Fa Wang, "Measurement of duration, energy of instantaneous frequencies, and splits of subcomponents of the second heart sound," *IEEE Trans. Instrum. Meas.*, vol . 64, no. 7, pp. 1958-1967, July 2015.
[2] B.Karnath and W. Thornton, "Auscultation of the heart," *Hospital Physician*, vol. 38, no. 9, pp. 39-43, 2002.
[3] S. Barma, B.-W. Chen, K. L. Man, and J.-F. Wang, "Quantitative measurement of split of the second heart sound (S2)," *IEEE/ACM Trans. Comput. Biology and Bioinformatics*, vol. 12, no. 4, pp. 851-860, July-August 2015.
[4] Springer et al. "Logistic regression-HSMM-based heart sound segmentation," *IEEETrans. Biomed. Eng.* vol. 63, no. 4, pp. 822-32, April 2016.
[5] Tien-En Chen et al., "S1 and S2 heart sound recognition using deep neural networks,"*IEEE Trans. Biomed. Eng.*, vol. 64, no. 2, pp. 372-380, February 2017.
[6] C.Papadaniil and L. Hadjileontiadis, "Efficient heart sound segmentation and extraction using ensemble empirical mode decomposition and kurtosis features," *IEEE*
[7] J. Biomed.and Health Informatics, vol. 18, no.4, pp. 1138-1152, July 2014.J. Herzig, A. Bickel, A. Eitan, and N. Intrator, "Monitoring cardiac stress using features extracted from S1 heart sounds," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 4, pp. 1169-1178, April 2015.
[8] X.-Y. Zhang, E. MacPherson, and Y.-T. Zhang, "Relations between the timing of the second heart sound and aortic blood pressure," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 4, pp. 1291-1297, April 2008.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijareeie.com

Vol. 8, Issue 3, March 2019

[10] J. Xu, L.-G. Durand, and P. Pibarot,"Nonlinear transient chirp signalmodeling of the aortic and pulmonary components of the second heart sound," *IEEE Trans. Biomed.Eng.*, vol. 47, no. 10, pp. 1328-1335, October 2000.
[11] S. Xiao et al., "A relative value method for measuring and evaluating cardiac reserve,"*Biomed. Eng. Online*, vol. 6, pp. 1-6, December 2002.
[12] C. Kwak and O.W. Kwon, "Cardiac disorder classification by heart sound signals using murmur likelihood and hidden Markov model state likelihood," *IET Signal Processing*, vol. 6, no. 4, pp. 326-334, June 2012.
[13] S. Sun et al., "Segmentation-based heart sound feature extraction combined with classifier models for a VSD diagnosis system," *Expert Sys. Appl.*, vol. 41, no. 4, pp. 1769-1780, March 2014.
[14] Felner JM. The Second Heart Sound. In: Walker HK, Hall WD, Hurst JW
editors. Clinical Methods: The History, Physical, and Laboratory Examinations. 3rd edition. Boston: Butterworths; 1990. Chapter 23. Availablefrom:http://www.ncbi.nlm.nih.gov/books/NBK341
[9] M. Rangayyan, "Biomedical Signal Analysis: A Case-Study Approach," *Wiley-IEEE Press*, 2001.
Abbas K. Abbas and Rasha Bassam, "Phonocardiography Signal Processing," *Morgan &Claypool Publishers*, 2009.
[15] A. Ravin, "Auscultation of the Heart," 3rd edition, Chicago, Year Book Medical Pub-lishers, 1977.
[16] V. Nivitha Varghees and K. I. Ramachandran, "Heart murmur detection and classification using wavelet transform and Hilbert phase envelope," *IEEE 21st National Conference onCommunications*, pp. 1-6, February 2015.
[17] R. M. Rangayyan and R. J. Lehner, "Phonocardiogram Signal Analysis: A Review", *CRCCritical Reviews in Biomedical Engineering*, vo1. 15, no. 3, pp. 211-236, February 1987.
[18] S. Santos, P. Carvalho, R. P. Paiva, and J. Henriques, "Detection of the S2 split using the Hilbert and wavelet transforms," in *Congresso de M´etodos Num´ericos em Engenharia*, pp. 1-11, January 2011.
[19] P. S. Vikhe, N. S. Nehe, and V. R. Thool, "Heart sound abnormality detection using short time Fourier transform and continuous wavelet transform," in *Proc. IEEE 2nd Int.Conf. Emerging Trends in Engineering and Technology*, pp. 50-54, December 2009.
[20] P. S. Vikhe, S. T. Hamde, and N. S. Nehe, "Wavelet transform based abnormality analysis of heart sound," in *Proc. Int. Conf. Advances in Computing, Control, and Telecommuni-cation Technologies*, pp. 367-371, December 2009. Vivek Nigam and Roland Priemer, "A dynamic method to estimate the time split between the A2 and P2 components of the S2 heart sound," *Physiol. Meas.*, vol. 27, no. 7, pp. 553-567, July 2006.