Convolutional Neural Network for Kidney and kidney Tumor segmentation *

Vikas Kumar Anand[1][0000−0002−7109−1638], Pranav Aurangabadkar, Mahendra Khened[1], and Ganapathy Krishnamurthi[1][0000−0002−9262−7569]

Department of Engineering Design, Indian Institute of Technology Madras, Chennai 600036, India
gankrish@iitm.ac.in

Abstract. In this work, we have attempted to develop algorithms for automatic segmentation of kidney and kidney tumous from CT images. We have exploited encoder decoder architecture of fully convolutional neural network.Pre-processesing steps involves slice extraction, data standardization and Hounsfield unit windowing. The proposed network has been trained on CT images of kidney and kidney tumors with their ground truth. Weighted combination of focal loss and dice loss has been minimized using Adam as optimizer. Dice coefficient of 94.68% and 94.51% has been achieved for kidney and kidney tumor segmentation respectively.

Keywords: Deep Learning · CNN · Segmentation · Computed Tomography.

1 Introduction

Kidney Tumors are result of out of control growth of malignant kidney cells. Generally, Kidney tumors are detected in early stage. However tumors may become large before detection. Abdominal imaging is common procedure for many different medical disorders such as hypertension, diabetes and, disease related to kidney. Computed tomography (CT) helps in detection of abnormalities in kidney. kidney abnormalities be a simple renal cyst or benign renal lesions or malignant renal cell. Renal cyst does not require treatment or follow-up, but in a notable proportion of cases masses of benign and malignant cells require additional procedures or interventions. Contrast enhanced CT shows morphology of kidney tumors. Tumor morphology plays important role in lesion's diagnosis and treatment. To establish quantitative relationship between tumor morphology and clinical output, it requires manual analysis of images. This task becomes very laborious due to manual quantification of large number of images for each case. Solution to this problem is automatic semantic segmentation of kidney and kidney tumors.

In this work is an attempt to develop an algorithms for fully automatic segmentation of kidney and kidney tumor using CT images. Our methods involves

---

convolutional neural network (CNN)[8] which takes CT images as input and ground truth generated on corresponding CT image as output.

CNNs very popular algorithms for wide variety of pattern recognition tasks. Most common applications of CNNs are image classification [7, 14, 2] and semantic segmentation using fully convolutional networks (FCN) [9]. CNNs have also been applied to medical image segmentation and classification [1, 10]. In this paper, we propose a CNN based architecture for segmentation of kidney and kidney tumors from CT images of the abdomen. Our network is inspired by from DenseNets [4] and U-Net[13].

2   Related Work

3   Materials and Method

3.1   Data pre-processing:

Each CT volume has fixed image size of 512 with different numbers of slices. 2-D slices are extracted from the CT volumes. Approximately 46000 2-D slices have been extracted from 210 cases. Hounsfield windowing (20-190) has been performed to enhance the contrast of kidney and kidney tumor tissues. This is followed by data normalization and data augmentation.

Data Normalization Volumes has been normalized to have zero mean and unit variance using Eq.1.

$$X_{norm} = \frac{X - \mu}{\sigma} \tag{1}$$

where X is the data, μ and σ are mean and global standard deviation associated with X.

3.2   Proposed Network

For segmentation of kidney and kidney tumor from CT images, we have explored encoder-decoder architecture. The encoder of the segmentation network is inspired by Densenet-121 architecture and has dense connectivity pattern. Where as the decoder of the network consists of up sampling modules and convolutional layers. Feature maps of decoder is concatenated with feature maps decoder. the architecture of the network is illustrated in Fig.1 The input to the network is 2-D CT slices of abdomen. There are 64 7×7 kernels are available in the first convolutional layer. Batch normalization [5] and a non linearity layer (ReLU)[12] has been applied to resultant feature of the first layer. To reduce the spatial dimension of feature maps max-pooling layer is added with kernel size 3×3 and strides of 2. Max-pooling layer is followed by series of dense block and transition block. Each dense block has different numbers of convolutional layers, here each layer receives inputs from all previous convolutional layers. Each layer of dense block which is shown in Fig.3 has batch normalization, ReLU,
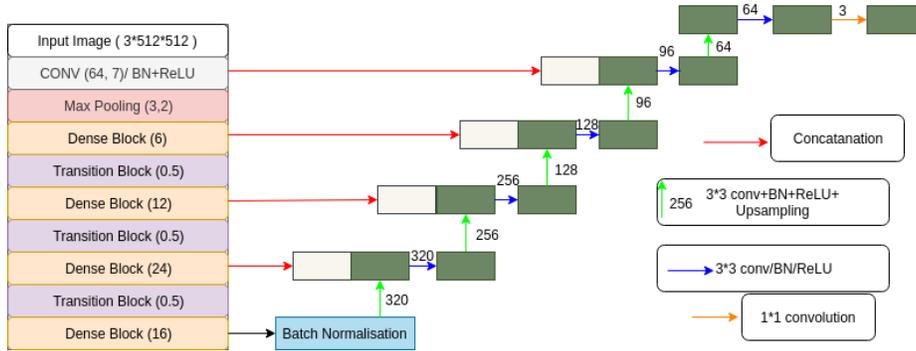
Fig. 1: Proposed network

convolution, and Dropout operation in series. This network has 4 dense block with $6, 12, 24,$ and 16 convolutional layer respectively. Transition block is shown in Fig. 4 and has batch normalization, ReLU, 1×1 convolution, dropout and max-polling layer operation in series.
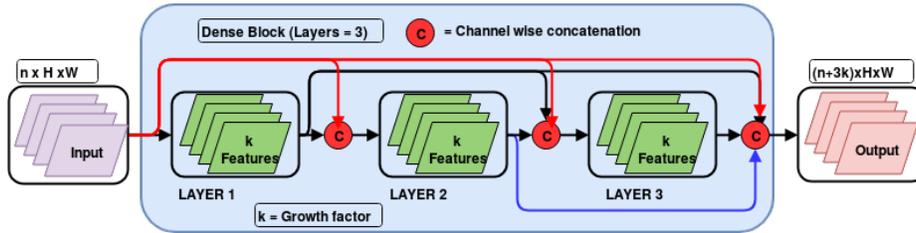

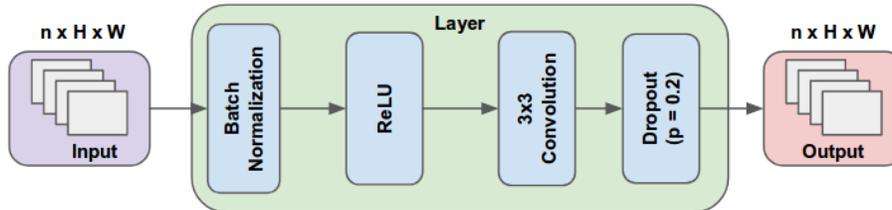Fig. 2: Dense block with 3 layers
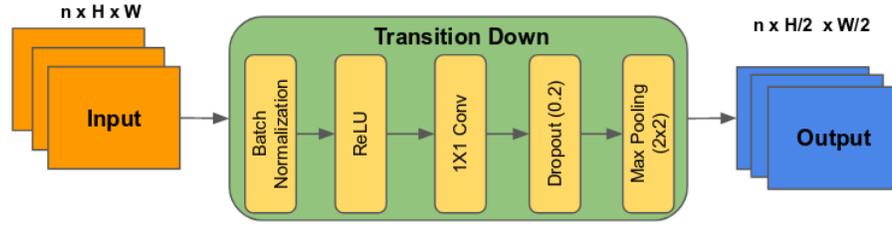

Fig. 3: Layer of dense block

Fig. 4: Transition block

3.3   Loss Function

Lesions are represented by a minuscule proportion of voxels in a medical volume by thereby leading to class imbalance. This issue was circumvented by training the network to minimize a hybrid loss function. The hybrid cost function comprised of weighted cross entropy & dice loss [11].

The dice co-efficient is an overlap metric used for assessing the quality of segmentation maps. The dice coefficient between two binary volumes can be written as:

$$DICE = \frac{2\sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2} \tag{2}$$

where the sums run over the N voxels, of the predicted binary segmentation volume $p_i \in P$ and the ground truth binary volume $g_i \in G$

The parameters of the network was optimized so as to minimize the $total\_loss$, Equation. (3).

$$total\_loss = \lambda(focal\_loss) + \gamma(dice\_loss\_bg) + \delta(dice\_loss\_kidney) + \theta(dice\_loss\_tumor) \tag{3}$$

where λ, δ and γ are empirically assigned weights to individual losses, fg and bg represent foreground voxels which corresponded to lesion regions and background voxels which corresponded to non-lesion regions respectively. In this work we set $\gamma = 0.25$, $\delta = 0.25$ and $\lambda = 0.50$.

Training The proposed model has been trained on a batch size of 4 for 60 epochs using ADAM [6] as the optimizer. The given dataset is spilted into training,testing and validation sets and percentage of split is 70%, 20% and 10% respectively. In encoding layer of network, we have used weights of DenseNet-121 which is pre-trained on ImageNet dataset. Decoder layer of network is randomly initialized and fine tuned with training the network where as encoder layer weights are frozen. During training of our network, the model has been saved on every epoch and the best model selection criteria is based on the model which gave the highest dice score on the validation set.

4 Experimental Setup and Result

4.1 Data sets and Evaluation Criteria

Dataset: For each case, CT images of abdomen has been acquired in late arterial contrast phase that show full information about abdomen. Manual delineation of kidney and kidney tumor has been performed on axial plane, and series has been regularly subsampled in the longitudinal direction to keep approximate 50 slices per case that have kidney and tumor information. Acquisition process of CT images and Ground truth generation has been explained by challenge organizer in [3]. The challenge dataset has total 300 cases, out of which 210 cases has been made available for training phase and rest 90 cases has been made available for testing phase.

Evaluation Criteria: Average Score S, for the 90 test cases, has been calculated for each team. The challenge organizer has the following criteria for computation of average score S,

$$S = \frac{1}{90} \sum_{i=0}^{89} \frac{1}{2} \left(\frac{2 * n_{t,tp}^{(i)}}{2 * n_{t,tp}^{(i)} + n_{t,fp}^{(i)} + n_{t,fn}^{(i)}} + \frac{2 * n_{k,tp}^{(i)}}{2 * n_{k,tp}^{(i)} + n_{k,fp}^{(i)} + n_{k,fn}^{(i)}} \right)$$

where, $n_{t,tp}^{(i)}$ = True positive for tumor, $n_{t,fp}^{(i)}$ = False positive for tumor, $n_{t,fn}^{(i)}$ = False negative for tumor, $n_{k,tp}^{(i)}$ = True positive for kidney and tumor, $n_{k,fp}^{(i)}$ = False positive for kidney and tumor, $n_{k,fn}^{(i)}$ = False negative for kidney and tumor

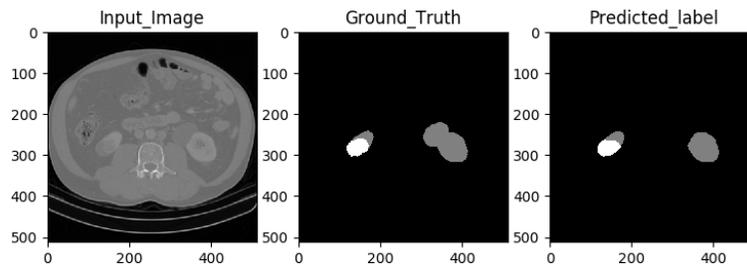### 4.2 Experimental Results

Fig. 5: Prediction made by our model.

Our proposed model able to achieve dice similarity of 0.9454615 and 0.88882035 for kidney and kidney tumor on held out test data.

5   Conclusion

Algorithms for automatic segmentation of kidney and kidney tumor segmentation has been developed. Segmentation of kidney and tumor plays important role in diagnosis.

6   Acknowledgement

We would like to thank Intel AI DevCloud for providing computational resource.

References

1. Ciresan, D., Giusti, A., Gambardella, L.M., Schmidhuber, J.: Deep neural networks segment neuronal membranes in electron microscopy images. In: Advances in neural information processing systems. pp. 2843–2851 (2012)
2. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
3. Heller, N., Sathianathen, N., Kalapara, A., Walczak, E., Moore, K., Kaluzniak, H., Rosenberg, J., Blake, P., Rengel, Z., Oestreich, M., et al.: The kits19 challenge data: 300 kidney tumor cases with clinical context, ct semantic segmentations, and surgical outcomes. arXiv preprint arXiv:1904.00445 (2019)
4. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: CVPR. vol. 1, p. 3 (2017)
5. Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 (2015)
6. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
7. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems. pp. 1097–1105 (2012)
8. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE **86**(11), 2278–2324 (1998)
9. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 3431–3440 (2015)
10. Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., et al.: The multimodal brain tumor image segmentation benchmark (brats). IEEE transactions on medical imaging **34**(10), 1993 (2015)
11. Milletari, F., Navab, N., Ahmadi, S.A.: V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: 3D Vision (3DV), 2016 Fourth International Conference on. pp. 565–571. IEEE (2016)
12. Nair, V., Hinton, G.E.: Rectified linear units improve restricted boltzmann machines. In: Proceedings of the 27th international conference on machine learning (ICML-10). pp. 807–814 (2010)
13. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241. Springer (2015)

14. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)