Brain Tumor Segmentation using Dense Fully Convolutional Neural Network

Mazhar Shaikh, Ganesh Anand, Gagan Acharya, Abhijit Amrutkar, Varghese Alex, and Ganapathy Krishnamurthi

Medical Imaging and Reconstruction Lab, Department of Engineering Design, Indian Institute of Technology Madras, Chennai, India
gankrish@iitm.ac.in

Abstract. Manual segmentation of brain tumor is often time consuming and the performance of the segmentation varies based on the operators experience. This leads to the requisition of a fully automatic method for brain tumor segmentation. In this paper, we propose the usage of the 100 layer Tiramisu [6] architecture for the segmentation of brain tumor from multi modal MR images, which is evolved by integrating a densely connected fully convolutional neural network (FCNN), followed by post-processing using a Dense Conditional Random Field (DCRF). The network consists of blocks of densely connected layers, transition down layers in down-sampling path and transition up layers in up-sampling path. The method was tested on dataset provided by Multi modal Brain Tumor Segmentation Challenge (BraTS) 2017. The training data is composed of 210 high-grade brain tumor and 74 low-grade brain tumor cases. The proposed network achieves a mean whole tumor, tumor core & active tumor dice score of 0.87, 0.68 & 0.65. respectively on the BraTS '17 validation set and 0.83, 0.65 & 0.65 on the Brats '17 test set.

Keywords: Fully convolutional neural networks, multi modal MRI segmentation, conditional random fields, Tiramisu

1 Introduction

Accurate tumor segmentation is crucial for treatment and survival prediction of cancer patients. Segmentation of the gliomas from MR images is the preliminary step for treatment and surgical planning. Considering manual segmentation of gliomas being tedious and often results in inter rater variability, we adopt a 103 layer deep fully convolution neural network (FCNN) for automatic segmentation of gliomas. Originally proposed in [6], the Tiramisu model has significantly outperformed most state-of-the-art techniques on the CamVid and Gatech datasets. The architecture uses a characteristic upsampling path to overcome the problem of memory demand as a result of feature map explosion and allows the network to be very deep.

In this paper, we demonstrate the adaptation of this model for slice based biomedical image segmentation purposes. The network is further supplemented

by post processing techniques like CRF and connected components analysis to remove false positives generated by the network.

2 Related Work

Most of the work on brain tumor segmentation using convolutional neural networks(CNN) can be divided on the basis of the input and the corresponding output. While, there are models that input 2D/3D patch and output the center voxel class (Patch based), we also have networks that input the entire slice/volume and give a segmentation map having the same size as the input (Slice based). Both these techniques have their own benefits and demerits.

2.1 Patch Based Techniques

Patch Based Techniques such as by Pereira et. al. [10] make use of a CNN which takes a patch from the image and predicts the class for the center voxel. These techniques allow for using information from the immediate surrounding region for prediction. The main drawback of such techniques is that, being local, the predictions for voxels are independent of each other and as such there can be inconsistencies between nearby predictions which need to be corrected with post-processing. However, patch based models can address the issue of class imbalance by selective sampling of patches from the MRI data during training, like in [14]. Extending patch based techniques which use 2D image patches to 3D volumes (*DeepMedic* [7]) involves predicting the classes for a subset of the volume given as input to a 3D CNN.

2.2 Slice based techniques

Slice based techniques refer to networks which predict the classes for all the pixels or voxels for a given input. These techniques generally use convolutional layers along with downsampling to capture high level information about the input in a downsampling path and then use transposed convolutions or unparametrized upsampling methods such as bilinear upsampling in an upsampling path to predict the classes for the whole input. They make use of skip connections [12] to preserve low level feature information in the upsampling path for sharper predictions. FCNNs are commonly used as they offer many advantages. Prediction of classes for the whole input makes prediction quick. The predictions make use of more information than patch based techniques as the inputs are usually larger. FCNNs for 3D volumes [9] and Unet-3D [13] offer better performance but have greater computational and memory requirements. The inherent problem in any of these techniques is handling class imbalance.

3 Materials & Method

3.1 Data

The images used to train and validate this model were obtained from the BraTS 2017 challenge dataset [1], [2]. The training dataset consisted of multi modal MR images of 284 patients, with 210 patients from the high grade gliomas category (HGG) and 74 patients from the low grade gliomas (LGG) category. The following MRI modalities were provided for each patient: T2-weighted fluid attenuated inversion recovery (FLAIR), T1-weighted (T1), T1-weighted contrast-enhanced (T1ce), and T2-weighted (T2). The provided images were co-registered to the same anatomical template, interpolated to the same resolution (1 mm^3) and skull-stripped. The image dimension is 240×240×155, with 155 being the number of slices in the axial direction. Manually annotated ground truth segmentations were provided for three classes : GD-enhancing tumor (ET label 4), the peritumoral edema (ED label 2), and the necrotic and non-enhancing tumor (NCR/NET label 1). The network was trained on slices extracted from the axial plane.

3.2 Pre-processing

Multi modal scans can vary between patients depending on several factors including the instrument used, image acquisition axis, etc. In order to account for the patient-to-patient variation in the MR images, we adopted z-score normalization where we subtract the mean and divide by the standard deviation of the entire volume for each of the four channels of an individual's scan. Histogram matching was tried out but abandoned due to reduced performance.

3.3 Densely connected FCNN model

Our segmentation technique is based on the One Hundred Layers Tiramisu model proposed for semantic segmentation originally by Simon Jgou et al [6]. Like most state-of-the-art models, the Tiramisu model involves a down-sampling path and up-sampling path, where a single slice of the brain is provided as input at the beginning of the model and class-wise probabilities for every pixel is output at the end of the up-sampling path.

The Tiramisu model ,shown in Fig. (1), consists of dense blocks(DB) used in DenseNet [5], which are made up of repeated Batch Normalization layers, ReLU, 3×3 convolutions and small skip connections. The dense blocks are paired with transition down layers(TD) in the down-sampling path and transition up layers(TU) in the up-sampling path with long skip connections. Skip Connections here refer to concatenation of the layers. The transition down layer consists of 1×1 convolutions followed by 2×2 max-pool layer with stride 2. Transition up layer is composed of 3×3 transpose convolutions with stride 2. The various blocks used in the model as well as the number of layers per dense block are shown in Fig. (2).

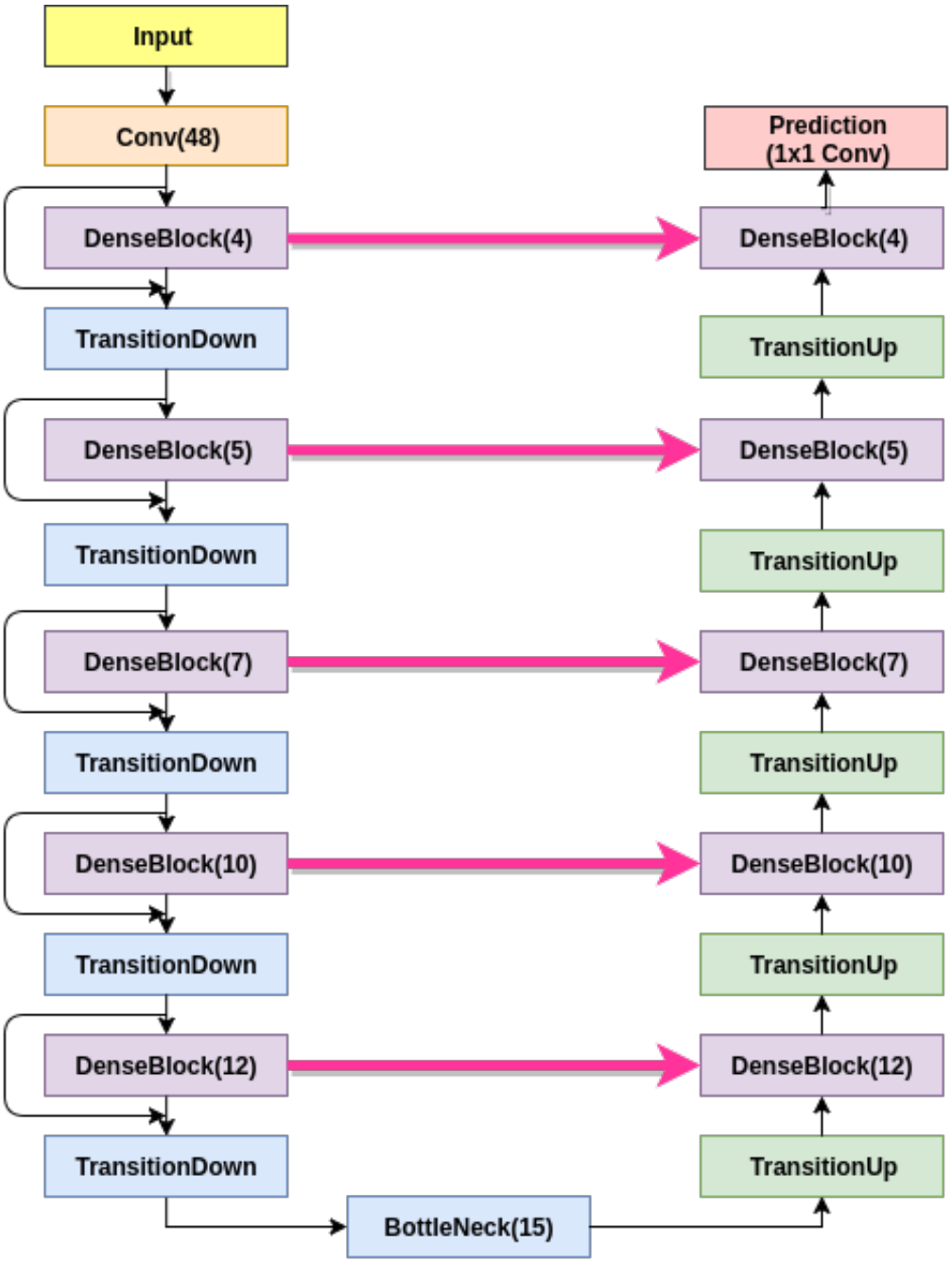
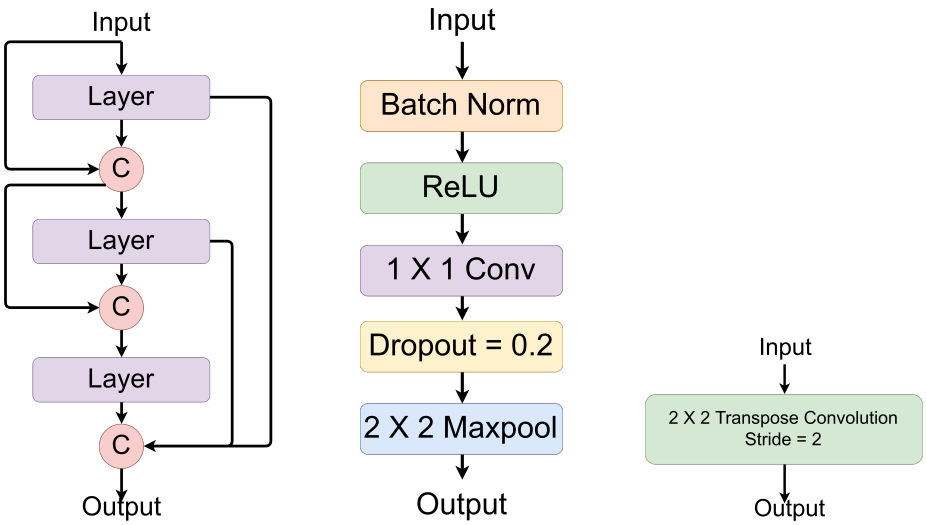


Fig. 1: Architecture of the adopted network



(a) (b) (c)

Fig. 2: Blocks used in the model. a) Dense Block. b) Transition Down. c) Transition Up

3.4 Post-processing using Dense-CRFs and Connected Components Analysis

To smoothen the segmentation predicted by the above model, we used fully connected conditional random fields with Gaussian edge potentials as proposed by Krähenbühl et al [8]. The unary potentials used by the CRF was computed using the predicted softmax probabilities. Further, the MR brain slice (all four modalities) along with the computed unary potentials was used for inferring the pixel labels. pydensecrf [1] repository was used for this purpose.

The false positives in the prediction were further reduced by using connected component analysis, wherein only the largest component of the segmentation predicted is retained. The post processing improved the mean dice score by 1%.

4 Implementation

The network was developed using TensorFlow[15] framework. Nvidia Titan X Maxwell GPU was used for training the model. The model was trained separately using cross entropy cost function, soft dice loss and cross entropy with soft dice score (1) as a regularizer. We observed that the soft dice loss gave a higher mean dice score and hence proceeded to train the network using the same. The soft dice score for binary classification [9] is given as

[1] pydensecrf : https://github.com/lucasb-eyer/pydensecrf

$$D = \frac{2\sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2} \tag{1}$$

ADAM[11] optimizer was used with $\beta_1 = 0.99$ and $\beta_2 = 0.995$ and an exponential decay rate of 0.9 every 1000 minibatches. Because of memory limitations, the batch size used during the training phase was 6 slices.

Class imbalance. As the employed network is slice based, the data is prone to have class imbalance owing to the fact that tumor comprises only a small fraction of the slice. In order to account for this imbalance in the data, we assign the following class weights in the loss function; 1 for the background class and 100 for the other classes. Also, only those image slices with tumor pixels in them were passed for training. We acknowledge that this is a run-of-the-mill technique and could be handled more effectively using patch based techniques.

5 Results

The performance of the proposed technique on the local HGG test data (n=21) is shown in Table (1) & Fig. (3). On the local test HGG data, the network achieved as mean whole tumor, tumor core and active tumor dice score of 0.84, 0.83, 0.80 respectively. The proposed post processing technique (CRF+ connected component analysis) yield a 1 % improvement in the whole tumor dice score, 1% in tumor core and 0.5 % in active tumor respectively.

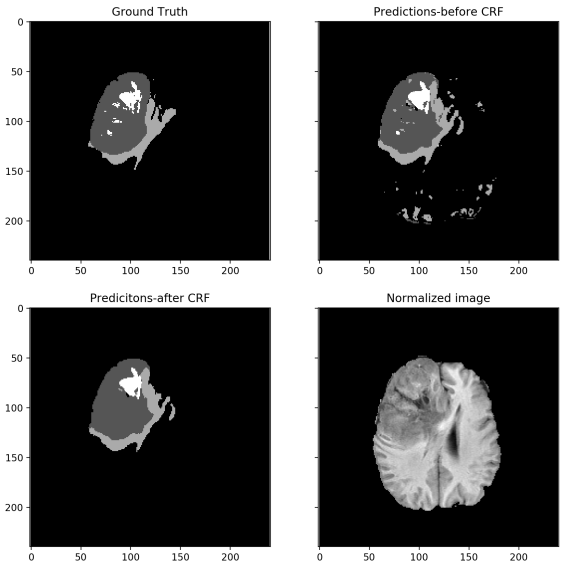
On the local LGG test data (n=8), The performance of the proposed technique is shown in Table (2). Compared to HGG, the model under performs on Tumor core segmentation, while maintaining good performance on the whole tumor segmentation.

Table 1: Results of local test HGG data

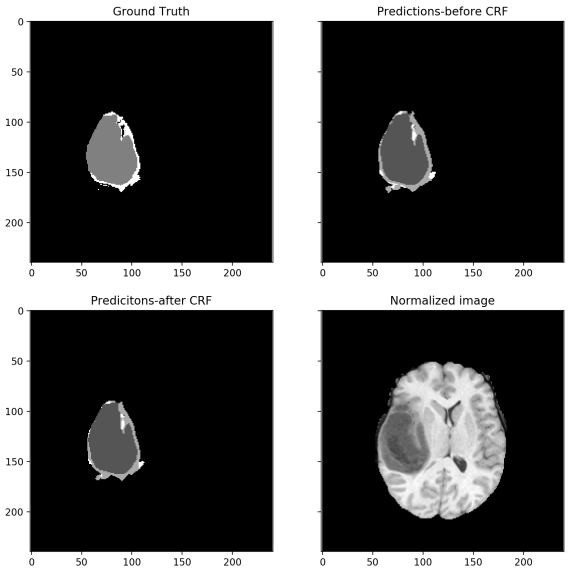
	Whole Tumor	Tumor Core	Active Tumor
Mean	0.84	0.83	0.80
Std Deviation	0.16	0.18	0.14
Median	0.89	0.87	0.84

Table 2: Results of local test LGG data

	Whole Tumor	Tumor Core
Mean	0.82	0.43
Std Deviation	0.11	0.29
Median	0.85	0.44

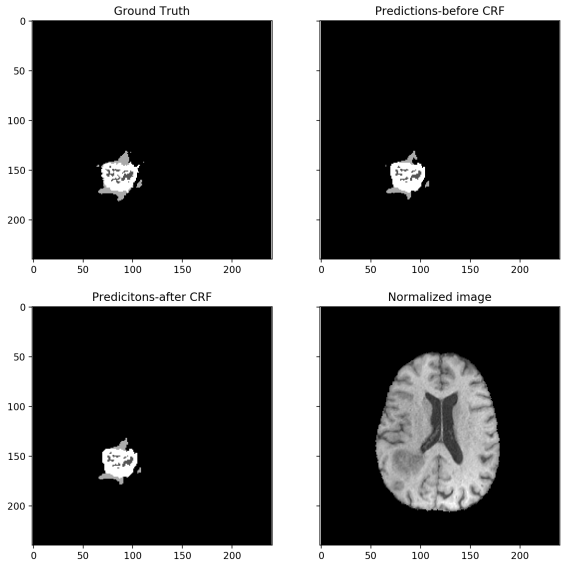


(a)

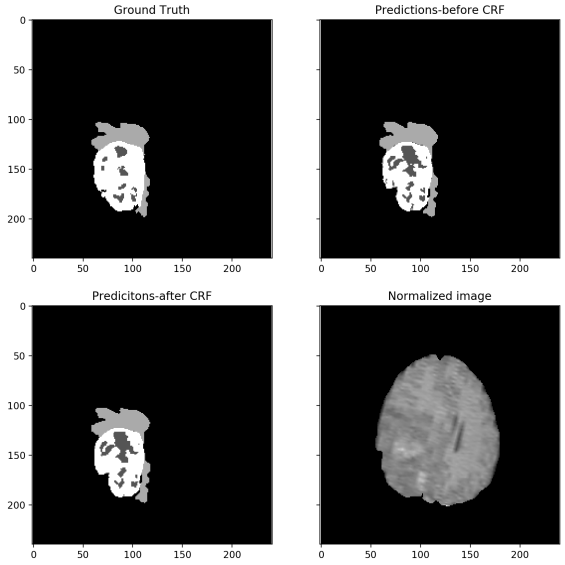


(b)

Fig. 3: Results of the proposed network on local test data. For each subfigure (Left to Right, Top to Bottom), Ground truth, Prediction(before post-processing), Prediction(after post-processing) , and the normalized FLAIR image slice, in that order.



(c)



(d)

Fig. 3: (contd.)

The performance of the proposed technique on the BraTS 2017 validation data (mixture of HGG and LGG) in given in Table (3). For whole tumor segmentation, the network maintains its performance on the validation data. However, a dip in performance was observed in the tumor core and active tumor regions on the validation set when compared to the local test data. The poor performance of the proposed technique on LGG tumor core segmentation negatively skews the performance statistics of our method on the validation data.

Table 3: Results of BraTS 2017 validation data (number of cases = 46)

	Whole Tumor	Tumor Core	Active Tumor
Mean	0.87	0.68	0.65
Std Deviation	0.11	0.34	0.32
Median	0.91	0.82	0.78

Table 4: Results of BraTS 2017 test data (number of cases = 147)

	Whole Tumor	Tumor Core	Active Tumor
Mean	0.83	0.65	0.65
Std Deviation	0.17	0.33	0.31
Median	0.89	0.84	0.78

In order to reduce the false negative occurrences in the predictions caused by the class imbalance, we utilized hard negative mining i.e repeated retraining of the network with slices that have large false negative outputs. However, it was observed that this did not yield significant improvement in the local validation set.

6 Conclusion

In this paper,we propose an automatic technique to segment gliomas from MR scans.

- A 103 layer deep network was implemented for segmentation of the gliomas from MR scans.
- A single network was used for the segmentation task, irrespective of the grade of the glioma.
- The Dense CRF improved the performance of the network in all compartments. Further connected component analysis removed grainy segments from the prediction[2].

[2] While reducing false positives, this technique will however result in removal of smaller tumor clusters

– The proposed network completes the entire pipeline (preprocessing, prediction & post processing) under 30 seconds.

As visible from the results the network performs well on Whole Tumor but faces difficulty distinguishing between classes and as such has lower Tumor Core and Active Tumor dice scores. This work only uses a 2D FCNN model. 2D models have to make predictions for a region with less information than 3D FCNNs. Even with the limitations in the batch size that could be trained, the network achieves good results.

Going forward 3D FCNNs will be used to improve performance. An exploration of the design space for 3D FCNNs will be done to find architectures which improve performance from 2D while still being able to run on a single GPU.

References

1. Bakas S, Akbari H, Sotiras A, Bilello M, Rozycki M, Kirby JS, Freymann JB, Farahani K, Davatzikos C. "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features", Nature Scientific Data, (2017) [In Press]
2. Menze BH et al. "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)", IEEE Transactions on Medical Imaging 34(10), 1993-2024 (2015)
3. Bakas S, Akbari H, Sotiras A, Bilello M, Rozycki M, Kirby J, Freymann J, Farahani K, Davatzikos C. "Segmentation Labels and Radiomic Features for the Pre-operative Scans of the TCGA-GBM collection", The Cancer Imaging Archive, 2017. DOI: 10.7937/K9/TCIA.2017.KLXWJJ1Q
4. Bakas S, Akbari H, Sotiras A, Bilello M, Rozycki M, Kirby J, Freymann J, Farahani K, Davatzikos C. "Segmentation Labels and Radiomic Features for the Pre-operative Scans of the TCGA-LGG collection", The Cancer Imaging Archive, 2017. DOI: 10.7937/K9/TCIA.2017.GJQ7R0EF
5. Huang G, Liu Z, Weinberger KQ, van der Maaten L. "Densely connected convolutional networks. arXiv preprint arXiv:1608.06993. 2016 Aug 25.
6. Jégou S, Drozdzal M, Vazquez D, Romero A, Bengio Y. The one hundred layers tiramisu: Fully convolutional DenseNets for semantic segmentation. arXiv preprint arXiv:1611.09326. 2016 Nov 28.
7. K. Kamnitsas, C. Ledig, V. F. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, et al., "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," Medical Image Analysis, vol. 36, pp. 61-78, 2017
8. Krhenbhl, Philipp, and Vladlen Koltun. "Efficient inference in fully connected crfs with gaussian edge potentials." Advances in neural information processing systems. 2011.
9. Fausto Milletari, Nassir Navab, Seyed-Ahmad Ahmadi. "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation." arXiv preprint arXiv:1606.04797. 15 Jun 2016
10. S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Deep convolutional neural networks for the segmentatin of gliomas in multi-sequence MRI," Proceedings MICCAI BraTS (Brain Tumor Segmentation Challenge), pp. 52-55, 2015
11. Diederik P. Kingma, Jimmy Ba. "Adam: A Method for Stochastic Optimization." arXiv:1412.6980. 30 Jan 2017

12. Ronneberger, O., Fischer, P., Brox, T.: "U-net: Convolutional networks for biomedical image segmentation". In: Medical Image Computing and Computer-Assisted InterventionMICCAI 2015, pp. 234241. Springer (2015)
13. zgn iek, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, Olaf Ronneberger, "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation" arXiv:1502.03240
14. Xiaomei Zhao et al., "A deep learning model integrating FCNNs and CRFs for brain tumor segmentation"
15. Martn Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Rafal Jozefowicz, Yangqing Jia, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Man, Mike Schuster, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Vigas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. "TensorFlow: Large-scale machine learning on heterogeneous systems, 2015". Software available from tensorflow.org.