




Automatic disease stage classification of glioblastoma multiforme histopathological images using deep convolutional neural network

Asami Yonekura¹ · Hiroharu Kawanaka¹ · V. B. Surya Prasath^{2,3,4}  · Bruce J. Aronow^{2,3} · Haruhiko Takase¹

Received: 20 December 2017 / Revised: 28 May 2018 / Accepted: 17 June 2018 / Published online: 25 June 2018
© Korean Society of Medical and Biological Engineering and Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

In the field of computational histopathology, computer-assisted diagnosis systems are important in obtaining patient-specific diagnosis for various diseases and help precision medicine. Therefore, many studies on automatic analysis methods for digital pathology images have been reported. In this work, we discuss an automatic feature extraction and disease stage classification method for glioblastoma multiforme (GBM) histopathological images. In this paper, we use deep convolutional neural networks (Deep CNNs) to acquire feature descriptors and a classification scheme simultaneously. Further, comparisons with other popular CNNs objectively as well as quantitatively in this challenging classification problem is undertaken. The experiments using Glioma images from The Cancer Genome Atlas shows that we obtain 96.5% average classification accuracy for our network and for higher cross validation folds other networks perform similarly with a higher accuracy of 98.0%. Deep CNNs could extract significant features from the GBM histopathology images with high accuracy. Overall, the disease stage classification of GBM from histopathological images with deep CNNs is very promising and with the availability of large scale histopathological image data the deep CNNs are well suited in tackling this challenging problem.

Keywords Deep learning · Histopathology · Image analysis · Glioblastoma multiforme · Classification · Convolutional neural network

✉ V. B. Surya Prasath
surya.prasath@cchmc.org; prasatsa@uc.edu

Asami Yonekura
417m246@m.mie-u.ac.jp

Hiroharu Kawanaka
kawanaka@elec.mie-u.ac.jp

Bruce J. Aronow
bruce.aronow@cchmc.org

Haruhiko Takase
takase@elec.mie-u.ac.jp

- ¹ Graduate School of Engineering, Mie University, 1577 Kurima-machiya, Tsu, Mie 514-8507, Japan
- ² Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, USA
- ³ Department of Biomedical Informatics, College of Medicine, University of Cincinnati, Cincinnati, OH 45267, USA
- ⁴ Department of Electrical Engineering and Computer Science, University of Cincinnati, Cincinnati, OH 45221, USA

1 Introduction

This paper discusses a new scheme for histopathological image analysis to acquire efficient feature descriptors and a classification scheme at the same time. Recent studies in histopathology have made promising developments for analysis that would enable precision medicine and patient-specific diagnosis [1–3]. In typical clinical practices, medical doctors or pathologists manually analyze histopathological images leading to a diagnosis.

We have, however, some problems as follows. The number of tissue images for diagnosis gives large burdens to pathologists in the case of manual analysis. For instance, they have time pressure in the case of diagnosis during surgery, because they have to analyze a plenty of histopathological images and make a diagnosis quickly. Also, evaluation criteria heavily depend on the experiences and subjectiveness of each medical doctor or pathologist. The result of the analysis will not be quantitative.

Table 1 Detailed layer information for our deep CNN configuration

Layer	Type	Filter size	# of Feature maps	Spatial size
0	Input		1	1000 × 1000
1	Resize		1	224 × 224
2	Conv1 + ReLU	5 × 5	32	224 × 224
3	Pool1	2 × 2	32	112 × 112
4	Conv2 + ReLU	5 × 5	32	112 × 112
5	Pool2	2 × 2	32	56 × 56
6	Conv3 + ReLU	5 × 5	48	56 × 56
7	Pool3	2 × 2	48	28 × 28
8	Conv4 + ReLU	5 × 5	48	28 × 28
9	Pool4	2 × 2	48	14 × 14
10	Conv5 + ReLU	5 × 5	64	14 × 14
11	Conv6 + ReLU	5 × 5	64	14 × 14
12	Pool6	2 × 2	64	7 × 7
13	Conv7 + ReLU	5 × 5	128	7 × 7
14	FC1 + ReLU + Dropout		1024	1 × 1
15	FC2 + Softmax		2	1 × 1
16	Output			

For these problems, recent research works on Biomedical Informatics use computer vision techniques for computational pathology [4–10]. These papers aim to realize quick, efficient and quantitative analysis, and some computer vision techniques is used to detect tumors from histopathological images. These approaches require a strong feature extraction and classification techniques to determine a patient’s diagnosis. In other words, we need to discover strong feature descriptors for given images and have to develop the best suited classifier at the same time. It is a difficult task since there are many feature descriptors that needs to be tested and their numerous available combinations.

In this work, we propose a disease stage classification method with Deep Learning techniques for Glioma histopathological images. This paper employs deep convolutional neural network (Deep CNN) as a Deep Learning model, and construct the data set using Glioma histopathological images that consists of Hematoxylin and Eosin (H&E) stained color images for experimental materials. As a result of our evaluations with deep convolutional neural network experiments, the final classification accuracy of 96.5% was obtained. We further conducted comparisons with other popular deep learning pipelines, and with higher cross validation folds other networks perform similarly with a highest obtained accuracy of 98.0%.

The rest of this letter is organized as follows. Section 2 reviews previous works in this area, Sect. 3 provides the details of the glioma histopathology imagery and Sect. 4 provides the experimental results with deep CNN. Section 5 concludes the paper.

2 Related works

There are many prior works with respect to histopathological image analysis [11–16]. For instance, Fu et al. [12] proposed a nuclei segmentation method and applied CNN to fluorescence microscopy images. Huang et al. [12] proposed a neuronal outlines segmentation method and applied CNN to growth cones images. These authors used CNN techniques as segmentation to analyze histopathological images. Lim et al. [14] discussed diabetic macular edema and used 100 fundus images for evaluation experiments. The method achieved the classification accuracy of 80%. AbuHassan et al. [15] discussed tuberculosis disease and used tuberculosis (TB) images for evaluation experiments. The method achieved the classification accuracy of 97%, and this value was enough good for practical use. These methods, however, need to be followed by human correction of imperfections, that is, these approaches require not only customization of the algorithm but also human intervention. Also, these reports only applied the developed techniques to some specific diseases like breast cancer, esophagitis, but not other tissues.

Previously, we focused on Glioma and discussed feature descriptors for the given images. Tamaki et al. [7] discussed feature descriptors to distinguish disease from healthy cases. In the literature, they used Glioma histopathology images (archived in Allen Brain Institute), and determined significant feature descriptors to distinguish disease from healthy cases. Fukuma et al. [10] discussed feature descriptors and a classification scheme for disease stage classification. In the method, they determined feature descriptors with statistical tests and constructed a

scheme for disease stage classification using Support Vector Machines (SVM) and Random Forests (RF). Glioma is one of the most malignant tumors occurring in the brain. The prognosis of Glioma is usually quite poor in clinical practice. This is the reason for the authors to focus on Glioma, however, these works just focused on feature descriptors, and it is hard to know whether these descriptors are perfect for histopathological image analysis or not. Other feature descriptors may or may not be required for more advanced analysis. Also, appropriate feature descriptors should be used for analysis simultaneously because the combination of the feature descriptors is essential for accurate classification. The combination number of feature descriptors are enormous. The previous studies did not discuss this issue, and the technique is required to acquire not only sufficient feature descriptors but also powerful classifier at the same time.

3 Experimental material

Generally, Glioma can be categorized into four grades based on their disease stage. For example, Glioma of Grade-1 has a slight illness, and Grade-4 is so serious and has a poor prognosis. In particular, the average life expectancy of Grade-4 is 18 approximately months. Also, Glioma is infiltration growth, therefore it is quite difficult to remove all of them by surgery. That is the reason why disease stage classification of Glioma is required for effective treatments.

The histopathological images are dominated by regions with many cell nuclei and much cytoplasm. Some images contain other structures or tissue types, such as vessels, blood cells. In the imagery, the objects stained deep purple are cell nuclei, the objects stained bright red are blood cells, and the light colored regions are primarily cytoplasm.

The images used in this study were obtained from the publicly available The Cancer Genome Atlas (TCGA) [17] database. This database has many histopathological images such as Glioma, Breast Cancer, Esophagitis and so on. TCGA also contains two types of images, i.e., Lower-Grade Glioma (LGG) and Glioblastoma multiforme (GBM). LGG includes images of Grades 1 and 2, GBM includes images of Grades 3 and 4. In this study, these two distinct type images were used as the experimental imagery for automatic analysis. Figure 1 shows example images of Glioma. In these images, the nuclei were stained deep purple, and other tissues were stained pale purple and red by Haematoxylin and Eosin (H&E) staining. Image features about the nuclei appear different based on the grade, and features of the organization are different from each other. Generally, these differences are believed to result from the disease progression, including changes in gene and protein expression.

Also, the sizes of given images were large, and it might be unsuited for processing. For this reason, we divided the original images to patched images whose sizes were 1000×1000 pixels, and then the patched images including sufficient nuclei were used as materials. In this research work, 100 tissue images were downloaded in each category

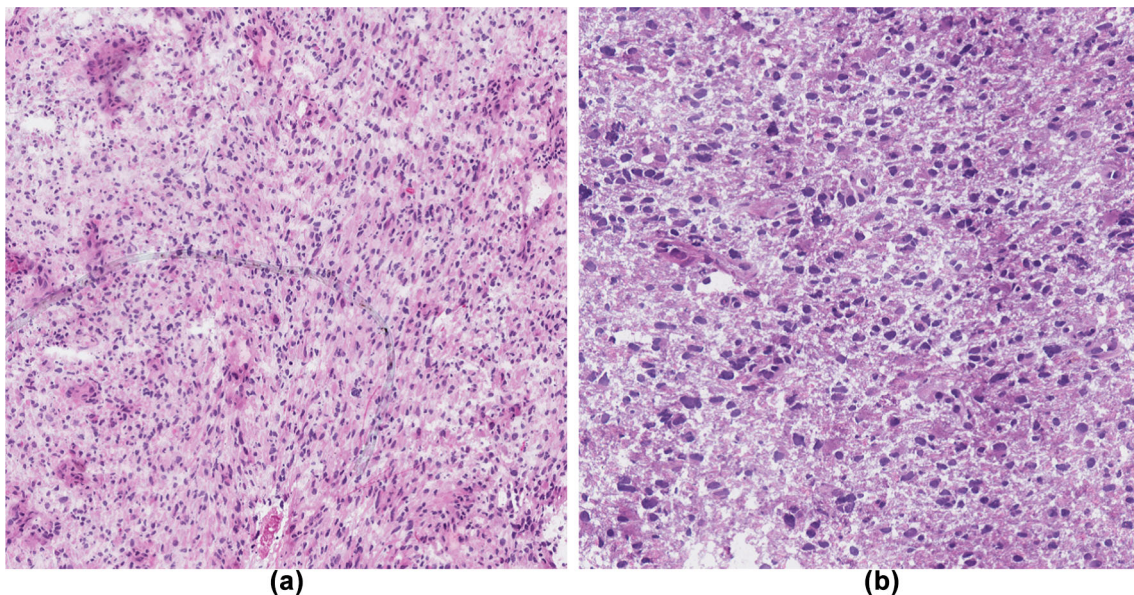


Fig. 1 An example of Histopathological images of Glioma from The Cancer Genome Atlas (TCGA). **a** Glioblastoma multiforme (GBM), **b** lower-grade glioma (LGG). Note the distinct distribution of nuclei

between low grade versus high grade. The shape, distribution, and morphology of the nuclei features can be used to predict the disease stage classification. **a** HGG, **b** LGG

and 100 patched images are obtained from a single image. In this work, 10,000 patched images were used in each class as experimental material. In the previous work, the authors used 10 image in each category [18], and the obtained results did not show a clear distinction between the disease stages. One explanation could be artifacts in the image, in the different situation due to differences in the level of the H&E staining. Therefore, in this work we used 100 LGG images and 100 GBM images respectively to avoid this problem. Further, We divided the training and test datasets in each fold based on single images. We did not use the patch level due to the fact that the obtained accuracy can be unreasonably high due to strong correlation between the same subject patch images.

4 Experimental method and results

4.1 Overview of deep convolutional neural network

In this paper, we utilized a Deep Learning method for the histopathological image analysis to acquire efficient feature descriptors and a classification scheme at the same time. This paper employs a Deep Learning method for this problem. By using Deep Learning method, feature descriptors are extracted from the data (image, sound, etc.) automatically, and we can not only extract feature descriptors from the given images but also construct a classifier for the problem automatically at the same time. Deep learning is currently popular in the field of computer vision and pattern recognition [19, 20] and in particular for computer assisted biomedical image analysis [21, 22]. Also, it showed outstanding performance in solving various biomedical image analysis problems [23, 24]. One of the popular Deep Learning techniques is Deep Convolutional Neural Network (Deep CNN), which is increasingly applied in various image analysis problems. There have been recent applications of these Deep CNN to adapt to biomedical applications [25, 26, 27]. In this study, we

Table 2 Experimental environment

Operating system	Ubuntu 14.04 LTS 64-bit
Memory	32GB
CPU	Core i7 6865K 6-core
GPU	NVIDIA GeForce GTX 1080Ti 11GB
HDD	2TB
Language	Python
Tool	TensorFlow

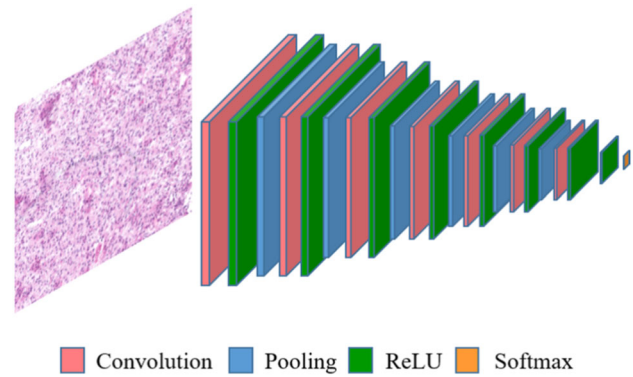


Fig. 2 Architecture of our Deep CNN configuration for Glioma histopathological images. The approach comprises of 7 convolution, 8 ReLU, 6 pooling, 1 softmax layers. For more information on our network, see Table 1

utilize Deep CNN for Glioma histopathological image to learn and extract features and check the classification accuracy in disease stage classification.

In the typical classification and pattern recognition problems, Deep CNN takes an image as input and produces a probability map as output. It basically performs multiple operations through hidden layers to produce some high level features that can potentially represent the target classes. The important operations that usually occurs are convolution, max pooling and rectified linear units (ReLU). Convolution is responsible for convolving regions with weights represented as filters to produce activation maps with discriminative features. The low-level feature maps usually have low level features such as curves or lines, while at the end of the network, the high-level representation can be achieved. Pooling performs down sampling to collect or combine low-level features in specified area to gain large invariance and sub-region summarization. ReLU is a kind of non-linearity which is useful in solving the gradient vanishing problem caused by the sigmoid function. Also, it produces a sparse representation that makes the network fast to learn. Softmax classification is used to discriminate between classes depending upon the probabilities produced from the fully connected layer.

4.2 Our network architecture

In our work, we utilized Deep CNN for the purpose of disease stage classification from Glioma histopathological images. Our adapted Deep CNN receives 1000×1000 patches of the Glioma images, and outputs two classes for GBM or LGG. This network has been built on top of the existing TensorFlow [28]. Experimental environment is shown in Table 2. Originally, we tried to adopt a network that has been used for the previous study, which consists of 3 convolution layers, 3 tanh and 3 pooling layers [29].

However, we found from our experiments that this network was not deep enough to capture the properties for our complex Glioma histopathological images. To obtain better segmentations, we built a new CNN architecture which is more deeper. Our new Deep CNN configuration mainly consists of 7 convolution layers, 8 ReLU and 6 pooling layers as shown in Fig. 2, with the spatial sizes and feature maps depth of the corresponding layers are given Table 1. The filter size is 5×5 for convolutional layers and 3×3 for pooling layers. Then, we constructed Dropout layer that rate was 0.5. The last softmax layer is to discriminate between the two classes to produce a patch which is either GBM or LGG.

4.3 Other network architecture

To check the performance of our architecture, we compared with some of the most successful convolutional neural networks. In this paper, we tried the comparison with following architectures.

- LeNet [30]
- ZFNet [31]
- VGGNet [32]

LeNet is the early and classic CNN model that Yann LeCun invented in 1998. It consists of 2 convolution layers, 2 pooling layers, and 3 fully connected layers. ZFNet is the winning model of ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2013. It is a CNN model that improved AlexNet which is the winning model of ILSVRC 2012 [33]. It consists of 5 convolution layers, 3 pooling layers, and 3 fully connected layers. The design concept of the basic architecture follows Neocognitron and LeNet. VGGNet was proposed by a research group at Oxford and it was the runners-up of the ILSVRC 2014 competition. 2–4 convolution layers with small filters are stacked in succession and after that, make it half in size

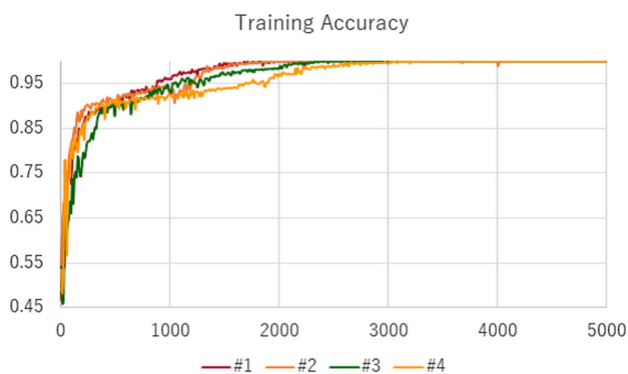


Fig. 3 Training accuracy result of our Deep CNN configuration

Table 3 Classification results of our network for each cross validation

# of Fold	Classification accuracy (%)
1	98.4
2	96.1
3	95.8
4	95.8
Average	96.5

with a pooling layer. It is characterized by the structure that repeats these process.

4.4 Experimental results

4.4.1 Our network architecture

Our network was trained for 5000 steps with a mini batch size of 100 using fourfold cross validation. Average of 4 values produced by each fold is the final classification accuracy. The training and testing took about a couple of days by each fold. As a result of this study, the final classification accuracy was 96.5 % with Fig. 3 shows the train result and Table 3 shows the classification accuracy. We note that the previous classification accuracy obtained using gray scale images was 87.2%. Compared to the previous result, we could improve the accuracy by using color images instead of gray scaling images. This result shows that Glioma histopathological images have meaningful color information. In the future, other feature descriptors would be required to discover new subtypes using the obtained feature matrices. This experimental results found that both of efficient feature descriptors and a classification scheme acquired at the same time.

4.4.2 Other network architecture

Other network were trained for 100 (LeNet), 1000 (ZFNet) and 2000 (VGGNet) steps with a mini batch size of 100 using fourfold cross validation. Average of 4 values produced by each fold is the final classification accuracy. The training and testing took about 20 min (LeNet), 7 h (ZFNet) and a couple of days (VGGNet) by each fold. As a result of this work, the final classification accuracy was 79.4% (LeNet), 95.2% (ZFNet) and 97.0% (VGGNet) with Fig. 4 shows each train result and Table 4 shows each classification accuracy. Actually, our Architecture was based on LeNet. But from this result, VGGNet have a potential for improvement classification accuracy of Glioma. Therefore, we may get higher classification

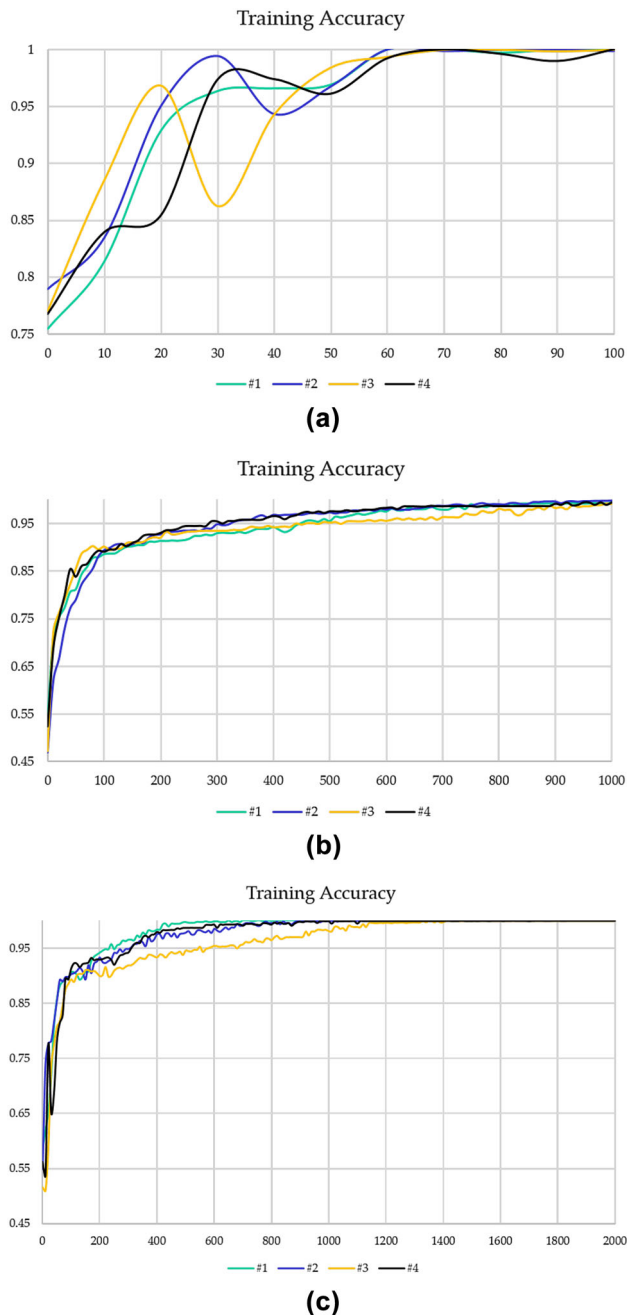


Fig. 4 Training accuracy result of other deep CNN configuration. **a** LeNet, **b** ZFNet, **c** VGGNet

accuracy if we add VGGNet component to our architecture. Of course, other CNN architecture may have a potential of more higher accuracy than VGGNet. Therefore, in the future, we will try many other CNN architectures to improve our architecture.

Table 4 Classification results of other network for each cross validation

# of Fold	Classification accuracy (%)		
	LeNet	ZFNet	VGGNet
1	78.5	94.4	97.4
2	78.1	94.7	95.2
3	82.8	95.9	97.4
4	78.2	95.7	98.0
Average	79.4	95.2	97.0

5 Conclusion

In this paper, the authors proposed a disease stage classification method with Deep Learning for Glioma histopathological images. In particular, Deep Convolutional Neural Network (Deep CNN) was employed as a Deep Learning model. By using the proposed method, we obtained a final average classification accuracy of 96.5%. Currently, we are investigating to discover new subtypes using the obtained feature matrices and confirm the relationships between the disease stage and the result of gene expression analyses. We believe that the proposed deep CNN method would be of help in the field of medical decision analysis of brain tumors.

Compliance with ethical standards

Conflict of interest The author declares that there is no conflict of interest.

Ethical statement This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Bilgin C, Demir C, Nagi C, Yener B. Cell-graph mining for breast tissue modeling and classification? In: IEEE annual international conference on engineering in medicine and biology society (EMBC); 2007. pp. 5311–5314.
2. Sertel O, Kong J, Catalyurek UV, Lozanski G, Saltz JH, Gurcan MN. Histopathological image analysis using model-based intermediate representations and color texture: Follicular lymphoma grading. *J Signal Process Syst.* 2009;55:169.
3. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: a review. *IEEE Rev Biomed Eng.* 2009;2:47–171.
4. Marchevsky AM, Wick MR. Evidence-based medicine, medical decision analysis, and pathology. *Hum Pathol.* 2004;35(10):1179–88.
5. Boucheron LE. Object-and spatial-level quantitative analysis of multispectral histopathology images for detection and

- characterization of cancer. Ph. D. dissertation, University of California, Santa Barbara, CA; 2008.
6. Rodenacker K, Bengtsson E. A feature set for cytometry on digitized microscopic images. *Anal Cell Pathol.* 2003;25:1–36.
 7. Tamaki K, Fukuma K, Kawanaka H, Takase H, Tsuruoka S, Aronow BJ, Chaganti S. Comparative study on feature descriptors for brain image analysis. In: *Proceedings of the IEEE international conference on and advanced intelligent systems (ISIS)*, 2014; pp. 679–682.
 8. Prasath VBS, Fukuma K, Aronow BJ, Kawanaka H. Cell nuclei segmentation in glioma histopathology images with color decomposition based active contours. In: *IEEE international conference on bioinformatics and biomedicine (BIBM)*, 2015; pp. 1734–1736.
 9. Fukuma K, Prasath VBS, Kawanaka H, Aronow BJ, Takase H. A study on nuclei segmentation, feature extraction and disease stage classification for human brain histopathological images. In: *20th international conference on knowledge based and intelligent information and engineering systems (KES)*, 2016; pp. 1202–1210.
 10. Fukuma K, Kawanaka H, Prasath VBS, Aronow BJ, Takase H. A study on feature extraction and disease stage classification for glioma pathology images. In: *IEEE international conference on fuzzy systems (FUZZ-IEEE)*. 2016; pp. 2150–2156.
 11. Huang JY, Hughes NJ, Goodhill GJ. Segmenting neuronal growth cones using deep convolutional neural networks. In: *IEEE international conference on digital image computing: techniques and applications (DICTA)*; 2016.
 12. Fu C, Ho D J, Han S, Salama P, Dunn KW, Delp EJ. Nuclei segmentation of fluorescence microscopy images using convolutional neural networks. In: *IEEE 14th international symposium on biomedical imaging (ISBI)*; 2017.
 13. Astrom K, Heyden A. Semantic segmentation of microscopic images of H&E stained prostatic tissue using CNN. In: *IEEE international joint conference on neural networks (IJCNN)*; 2017.
 14. Shi P, Zhong J, Huang R, Lin J. Automated quantitative image analysis of hematoxylin–eosin staining slides in lymphoma based on hierarchical K means clustering. In: *IEEE 8th international conference on information technology in medicine and education (ITME)*; 2016.
 15. Lim ST, Ahmed MK, Lim SL. Automatic classification of diabetic macular edema using a modified completed Local Binary Pattern (CLBP). In: *IEEE international conference on signal and image processing applications (ICSIPA)*; 2017.
 16. AbuHassan KJ, Bakhori NM, Kusnin N. Automatic diagnosis of tuberculosis disease based on Plasmonic ELISA and color-based image classification. In: *IEEE 39th annual international conference on engineering in medicine and biology society (EMBC)*; 2017.
 17. NIH The Cancer Genome Atlas (TCGA). <https://cancergenome.nih.gov/>. Accessed 28 May 2018.
 18. Yonekura A, Kawanaka H, Prasath VBS, Aronow BJ, Takase H. Glioblastoma multiforme tissue histopathology images based disease stage classification with deep CNN. In: *The 6th international conference on informatics, electronics & vision (ICIEV)*; 2017.
 19. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: *IEEE CVPR*; 2015. pp. 3431–3440.
 20. Ciresan D C, Giusti A, Gambardella L M, Schmidhuber J. Mitosis detection in breast cancer histology images with deep neural networks. In: *MICCAI*; 2013. pp. 411–418.
 21. Iqbal S, Khan MU, Saba T, Rehman A. Computer-assisted brain tumor type discrimination using magnetic resonance imaging features. *Biomed Eng Lett.* 2018;8(1):5–28.
 22. Prasath VBS. Deep learning based computer-aided diagnosis for neuroimaging data: focused review and future potential. *Neuroimmunol Neuroinflammation.* 2018;5:1.
 23. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: *MICCAI*; 2015. pp. 234–241.
 24. Greenspan H, van Ginneken B, Summers RM. Deep learning in medical imaging: overview and future promise of an exciting new technique. *IEEE Trans Med Imaging.* 2016;35(5):1153–9.
 25. Mansour RF. Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy. *Biomed Eng Lett.* 2018;8(1):41–57.
 26. Kassim YM, Prasath VBS, Glinskii OV, Glinsky VV, Huxley VH, Palaniappan K. Microvasculature segmentation of arterioles using deep CNN. In: *IEEE international conference on image processing (ICIP)*; 2017. pp. 580–584.
 27. Yonekura A, Kawanaka H, Prasath VBS, Aronow BJ, Takase H. Disease stage classification for Glioblastoma Multiforme histopathological images using deep convolutional neural network. In: *Joint 17th world congress of international fuzzy systems association and 9th international conference on soft computing and intelligent systems (IFSAS-SCIS)*; 2017.
 28. TensorFlow. <https://www.tensorflow.org/>. Accessed 28 May 2018.
 29. Yonekura A, Kawanaka H, Prasath VBS, Aronow BJ, Takase H. Improving the generalization of disease stage classification with deep CNN for glioma histopathological images. In: *International workshop on deep learning in bioinformatics, biomedicine, and healthcare informatics (DLB2H)*; 2017. pp. 1222–1226
 30. Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. In: *Proceedings of the IEEE*; 1998. pp. 2278–2324.
 31. Zeiler M, Fergus R. Visualizing and understanding convolutional networks. In: *Proceedings of ECCV*; 2014.
 32. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: *Proceedings of ICLR*; 2015.
 33. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: *Proceedings of NIPS*; 2012. pp. 1097–1105.