Glioblastoma Multiforme Classification On High Resolution Histology Image Using Deep Spatial Fusion Network[*]

P. Sobana Sumi [**][1] and Radhakrishnan Delhibabu[1,2]

[1] School of Computer Science and Engg.,
Vellore Institute of Technology, Vellore, India
sobanasumi.p2018@vitstudent.ac.in
[2] Modeling Evolutionary Algorithms Simulation and Artificial Intelligence,
Faculty of Electrical & Electronics Engineering, Ton Due Thang University,
Ho Chi Minh City, Vietnam
r.delhibabu@vit.ac.in & radhakrishnandelhibabu@tdtu.edu.vn

Abstract. Brain tumor is a growth of abnormal cells in brain, which can be cancerous or non-cancerous. The Brain tumor have scarce symptoms so it is very difficult to classify. Diagnosing brain tumor with histology images will efficiently helps us to classify brain tumor types. Sometimes, histology based image analysis is not accepted due to its variations in morphological features. Deep learning CNN models helps to overcome this problem by feature extraction and classification. Here proposed a method to classify high resolution histology image. InceptionResNetV2 is an CNN model, which is adopted to extract hierarchical features without any loss of data. Next generated deep spatial fusion network to extract spatial features found in between patches and to predict correct features from unpredictable discriminative features. 10-fold cross-validation is performed on the histology image. This achieves 95.6 percent accuracy on 4-class classification (benign, malignant, Glioblastoma, Oligodendroglioma). Also obtained 99.1 percent accuracy and 99.6 percent AUC on 2-way classification (necrosis and non-necrosis).

Keywords: Glioblastoma Multiforme · Deep spatial fusion network · InceptionResNetV2 · classification · patches · CNN

1 Introduction

Cancer tumor anywhere in the body spreads to the brain or it starts in the brain. A brain tumor can be normal (benign) or cancerous (malignant) based on its characteristics, it is normaly found in children and adults. The brain tumors are differentiated in to two types as Low Grade Gliomas (LGG) and High Grade Gliomas (HGG). Grade 1 and grade 2 are defined as LGG, grade3 and

---

[**] Corresponding author

grade 4 are defined as HGG. Astrocytomas are grade 1 and grade 2 level tumor, Oligodendroglioma is grade 3 level tumor and Glioblastoma Multiforme is grade 4 level tumor. Children are affected by Astrocytoma, Ependymoma and Medulloblastoma. Adults suffer from Astrocytoma, Oligodendrogliomas, Meningioma, Glioblastoma and Schwannoma. Gliobastoma Multiforme (GBM) is the matured stage of brain tumor with scarce symptoms so it is very hard to classify. Diagnosing this kind of tumors at the right time helps to increase patient survival. Histology images are obtained through biopsy. The biopsy is a process of taking tissue from the tumor and that tumor tissue is analyzed under electron microscope. Histology images have to be analyzed in large numbers with multiple staining to diagnose a single case, which causes time consuming problem. Sometimes this kind of histology image analysis is not accepted due to large variations in pathological features. At the initial stage to detect tumors in histology images Computer Assisted Diagnosis (CAD) systems are used. In this technique, images are scanned at first next the digital images are processed and analyzed by visual feature extraction of machine learning technique. Color difference occurs due to various scanners, staining procedures, different age patients and due to tissue thickness. Color normalization helps to classify colors among samples [1]. CAD works well in breast cancer images, not in brain tumor.

Xu et al. [2] proposed deep activation features for large scale histology images. CNN is used for feature extraction and the extracted features are passed into SVM to classify the patches of necrosis and non necrosis area. This method achieved 90 percent accuracy on small datasets by classification and segmentation process. Fukuma et al. [3] have explored feature extraction and disease stage classification for glioma histology images. Here tumors are distinguished as LGG and GBM by using significant features as object level features and spatial arrangement features to classify disease stage. These features are evaluated by K-S test and this obtained result is classified by using SVM classifier. Classification accuracy is low when compared to other non significant features. In the work of automated discrimination of low and high grade glioma, Mousavi et al. [4] proposed about pseudopalisiding necrosis area that is detected by cell segmentation and by cell count profile. Microvascular proliferation detection (MVP) is detected by spatial and morphological feature extraction. Finally, the hierarchical decision is made through a decision tree. MVP detection accuracy is less when compared to necrosis because of its structural complexity. Macyszyn et al. [5] examined about multidimensional pattern classification method. Here SVM classifier is used to classify patient survival in short (6 months) and long terms (18 months). Two cross accuracy for short term survival and three cross accuracy for the long term survival is used. Here MRI images are analyzed to predict survival. Barkerc et al. [6] explored coarse to fine method to analyze the characters of pathology images. Spatial features like shape, color and texture are extracted from tiled region and passed into clustering to have better classification. K-means is used for clustering and PCA is used to reduce data dimensionality and classification complexity. Powell et al. [7] examined low grade gliomas using a bag of words approach. The edge detection algorithm is used

for nuclear segmentation on hematoxylin stain and eosin stains. Threshold is assigned using global value. K mean is used for feature extraction and SVM is used to classify overall patient survival in short and long terms. Xu et al. [8] suggested CNN architecture AlexNet for classification. Achieved good result than previous methods. Here image analysis is done with a limited number of images where CNN works well on a large dataset. Yonekura et al. [9] proposed CNN architecture with deep networks. This network consists of three convolution layers, three pooling layers, and three ReLU. Classification accuracy is low when compared with other network but the work mainly focuses on disease stage classification. Yonekura et al. [10] suggested deep CNN architecture of LeNet network to extract features and to classify disease stage. Classification accuracy is low when compared with other networks.

Classifying high resolution histology image is a major problem. CNN can extract unpredictable discriminative features but training CNN directly with high resolution image is computationaly high. Mostly, histology image is found with unpredictable discriminative features which causes a challenge in patch based CNN classification. So, to solve this problem, InceptionResNetV2 architecture is adopted for hierarchical feature extraction and deep spatial fusion network is used to predict spatial features found in between patches. This proposed system gives better accuracy.

The paper is organised as follows. Section 2 introduces the studied problem. Section 3 describes two freely available databases with high-resolution brain tumor histology images. Section 5 briefly describes types of neural networks used in computer vision and image analysis and peculiarities of histology image. Section 4 introduces the used architecture of deep neural networks. Section 6 provides readers with the essence of the proposed solution, while Section 7 explains training process aspects. Section 8 is devoted to machine experiments and results discussion. Finally, Section 9 concludes the paper.

2   Problem Statement

Histology image diagnosing is completely a human factor process. Sometimes this kind of analysis is disagreed due to various differences in morphological features. To diagnose a single tissue and to classify its disease stage, tissue has to be analyzed under various magnification factors. Several tissues have to be analyzed by a pathologist to conclude. In some cases, both surgery and tissue diagnosing has to be done at the same time, this leads to time consuming problems. Lack of images (datasets) and images with good quality are rare. Large datasets should be processed with deep learning technique. Many diseases are not diagnosed due to the lack of training data. Histology images are diagnosed as per H and E stains only. Imbalanced datasets can be solved by data augmentation and some other molecular features can be raised to diagnose histology images other than H and E stains.

3   Dataset

TCGA and TCIA are most popular databases from where these high-resolution brain tumor histology images are taken. TCGA database consists of 2034 high resolution brain tumor histology images with the size of 2048*1536 and TCIA database consist of 2005 images. Each histology images are taken by biopsy process to maintain its molecular composition and its original structure. Each dataset is H and E stained microscopic histology image, with various magnification factors like 40X, 100X, 200X and 400X. These datasets are found under various classifications as benign, malignant, Astrocytoma and Oligodendroglioma. Both datasets consist of this 4 class labels evenly. Astrocytoma and Oligodendroglioma are malignant type tumors. To avoid data imbalance and overfitting, Tensorflow perform data augmentation by rotating, saturation adjustment etc. Normalization is done with an interval of [-1, 1], before the augmentation process to reduce variance that occurs through H and E stains [14]. To have a better classification four types of magnification images are given in the ratio of 7:3 for the training and testing process. Tumor classification works mainly focus on benign and malignant binary classification.

4   Architecture

The proposed architecture is shown in figure 1. A high-resolution brain tumor histology image is given as input. Unpredictable discriminative features are present sparsely on the entire image, which denote that it is not necessary for all patches to be consistent with image-wise labels. To model this fact and also to have good image-wise prediction, a spatial fusion network has been proposed. First, the adapted InceptionResNetv2 is trained to extract hierarchical discriminative features and predict probabilistic values of different cancer type for local image patches. Compared to VGG [11], InceptionResNetV2 (INRV2) performs well because of its shortcut connection network. Skip connection structure of INRV2 helps to reduce several problems with training deep neural networks that occur while performing backpropagation and improve feature extraction. Secondly, a deep spatial fusion network is specially designed to learn spatial relationship between patches, it gets input from the spatial feature map. Patch-wise probability vector is taken as a base unit for spatial feature maps to have better usage. The fusion model learns to adjust the bias of patch-wise predictions and tends to have efficient image-wise prediction, compared to typical fusion methods.

5   Related Theory

5.1   Convolutional Neural Network

CNN is a category of the neural network, which shows its effective result in image classification, image recognition etc. Its operations are convolution, non
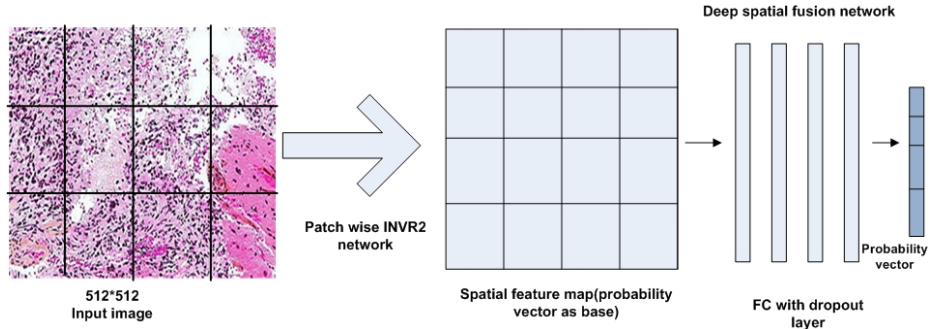
Fig. 1. Proposed architecture of Spatial fusion network. 512*512 size pixel is given as input to deep patch process. Spatial fusion network is trained by INRV2. The probabilistic vector is given as base to spatial feature map and processed with deep spatial fusion network. The dropout layer is added to avoid overfitting and increase robustness.

linearity (ReLU), pooling or subsampling, classification (fully connected layer). Extracting features from the input image is the major work of the convolution part; the image convolves with a filter and gives the feature map. A rectified linear unit is a non-linear operation, which performs element-wise operation and replaces all negative values found in feature map with zero value. Pooling helps to reduce the dimensionality of each feature map without any loss in information. A fully connected layer is a multilayer perceptron which uses softmax activation function in the output layer. Convolutional and pooling layers give high level features as the output. The fully connected layer uses these high level features to classify the input image.

5.2   InceptionResNetV2

InceptionResNetV2 is a CNN network, which is trained on more than one million images from ImageNet database. It consists of 164 deep layer network, it can classify images into 1000 object categories – variety of animals, birds, box, pencil, etc. Through this, the network has learned several features from various images. ResNet skips its connections with no loss of information. Training phase with this network is much faster and produces better accuracy than Inception network. INRV2 achieves 19.9 percent on top 1 error and 4.9 percent on top 5 error. In the proposed work this network consist of 24 layers and 4 blocks.

5.3   Histology Image Analysis

The brain tumor histology image is analyzed only by H and E stained slides. Malignancy state can be determined by the presence or absence of certain histological features such as presence and absence of necrosis, mitotically active cells, nuclear atypia, microvascular proliferation (enlarged blood vessels).
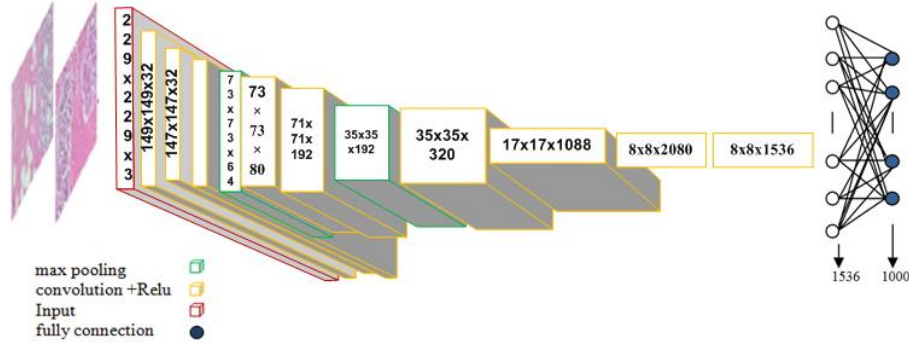
Fig. 2. InceptionResNetV2 network with 1536 feature vector.

These H and E stains are universally used for histological tissue examination. Classification and grading of brain tumor can be done by including other molecular information along with the histology image based information. In our proposed method deep patch based process and deep spatial fusion network is used to classify high resolution histology image.

6   Proposed Work

6.1   Patch-wise INVR2

Instead of adapting normal feedforward CNN our proposed architecture used InceptionResNetV2.

Compared to other CNN architectures INRV2 effectively reduces difficulties in training the deep network using shortcut connections and by residual learning. Though it skip connections there is no loss of information by non-linear functions. Extracted hierarchical features from low level to high level are combined to make final predictions, where as discriminative features are distributed in the image from cellular to tissue level. Input layer receives normalized image patches with the size of 512*512, which is sampled from whole histology image. Depth of the network is 24 layers with 4 block units for exploring region patterns in a different scale. 19*19 to 43*43, 51*51 to 99*99, 115*115 to 211*211 and 243*243 to 435*435 pixel are the size of four block groups size. This pixel size responds to region patterns in nuclei, structure organization and tissue organization.

6.2   Deep Spatial Fusion Network

The main purpose of the fusion model is to predict the image-wise label $\hat{z}$ among Y classes $C = \{C_1, C_2, \ldots, C_Y\}$, given all patch-wise feature maps F as the
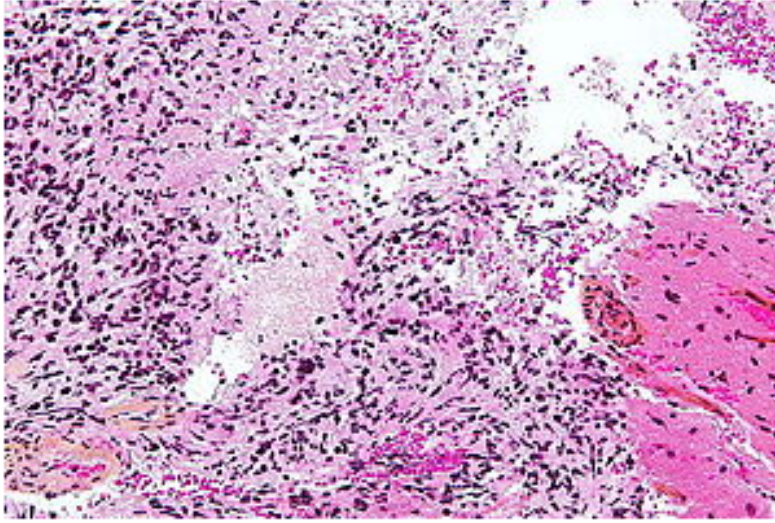
Fig. 3. Hematoxylin stains are acidic molecules, shades of blue. Eosin stains are basic materials shades of red, pink and orange.

input to the proposed INRV2 network. This image-wise classification prediction is defined as MAP [12] as follows

$$\hat{z} = arg \max_{z \in C} P(z|F).$$

If the entire high resolution image is divided into $M * N$ patches, then all the patch wise probability maps are arranged in spatial order, as

$$F = \begin{bmatrix} F_{11} & F_{12} & F_{13} & \dots & F_{1N} \\ F_{21} & F_{22} & F_{23} & \dots & F_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ F_{M1} & F_{M2} & F_{M3} & \dots & F_{MN} \end{bmatrix}.$$

Here, deep neural network (DNN) is used to utilize the spatial relationship between patches as shown in figure 5. Proposed fusion model contains 4 fully connected layers, each follows by ReLU activation function [13]. During image wise training, multilayer perceptron converts the spatial distribution of local probability maps into global class probability vectors. One dropout layer is added in before each hidden layer to avoid overfitting and to increase robustness. Also, a dropout layer is inserted between the flattened probabilistic vector and first hidden layer. By dropping out half of the probability maps, this models learns image wise prediction with half information of patches and also minimize the cross entropy loss in training.
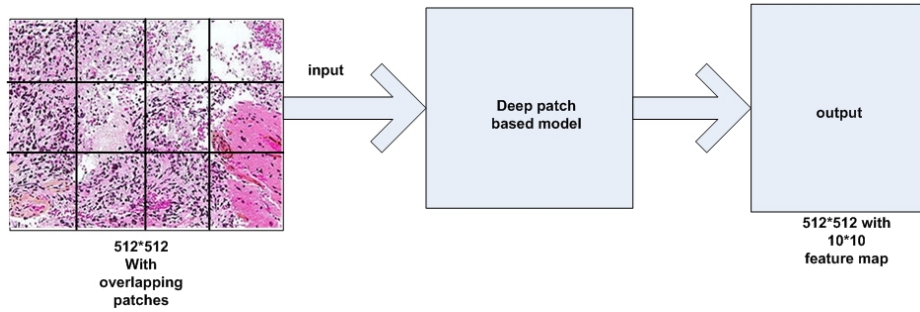
Fig. 4. Training of deep patch model. Given 512*512 size pixel as input to deep patch process, after training it consist of 512*512 patch with 10*10 feature map and 4 probabilistic values.
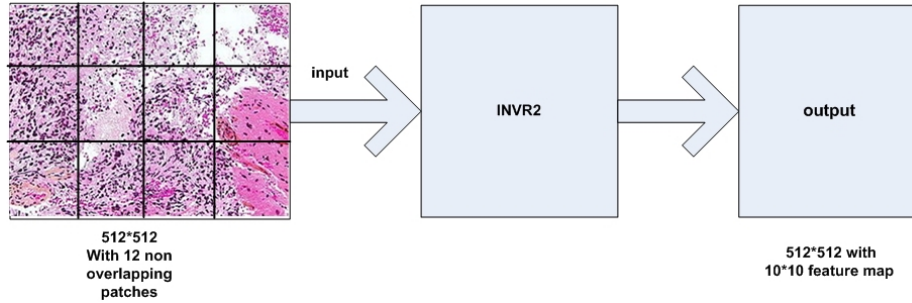

Fig. 5. Training of spatial fusion network. 512*512 size pixel is divided into 12 patches, each patch is given separately as input to InceptionResNetV2, after training it consist of 512*512 size patch with 10*10 feature map and with 4 probabilistic values.

7 Network Training

512*512 size pixel with overlapping patches extracted from the high resolution image and given as input for deep patch based model shown in figure 4. We assume that patch labels are consistent with image ground truth values because patch-wise labels are not given in the training dataset. Bias may suffer during patch based training and reduce patch-wise classification accuracy. By the second stage under supervised learning of image labels, bias get reduced during image based training. To avoid imbalanced dataset and overfitting, augmentation is done by random rotation, by making a change in contrast, brightness and by horizontal flipping etc. This generate 202,174 patches from the TCGA database and 140,099 patches from TCIA database. InceptionResNetV2 is trained on 32 sized mini batch to minimize the cost function of cross entropy using Adam optimization [15] with the learning rate of 10^{-5} for each 50 epoch. Now this

patch wise network after training, encodes the 512*512 patch to 10*10 feature map and to 4 class probabilistic values.

Data augmentation process is performed, to train spatial fusion network and generated 5,890 high resolution images from the TCGA training dataset and 3,998 images from TCIA training dataset. Each high resolution image is divided into 12 non overlapping patches with 512*512 size pixels after augmentation. Each of these patch is given individually as an input to InceptionResNetV2 and output 512 feature map of size 10*10 and with a class probability vector of size 1*4. Probabilistic vectors of patches in the current image are then combined into a probabilistic map following their own spatial order which is given as input for spatial fusion network. This generated probability map can be viewed as high level feature map that encodes all the patch wise discriminative features and the image wise spatial context features. After that it is analyzed with image ground truth values. Spatial fusion network weights are learned by using mini batch gradient descent with a batch size of 32 with Adam optimization. To minimize cross entropy loss function during training, the spatial fusion model encodes the biased probabilistic map into k class vector as per approximate image ground truth (k=4). By using the spacial context aware features hidden in probabilistic map, image based classification accuracy can be improved.

8   Experimental Results

We evaluated the performance of the patch based InceptionResNetV2 and then focused on the effectiveness of spatial fusion network by having multiple experiment comparison on the two dataset. In the first experiment (Baseline) on TCIA dataset is a baseline method [16], which is based on patch based plain CNN along with multiple vote based fusion method. Second (Residual and Vote) on TCIA dataset, replaced plain CNN with patch wise residual network. Our proposed work is examined on both dataset and shown in Table 1. All these methods are evaluated with ten fold cross validation. On the TCIA dataset, proposed method

Table 1. Comparision result on TCIA and TCGA dataset

Datasets	Methods	4 class ACC	2 class ACC	STD
TCIA	Baseline	0.778	0.844	-
TCIA	Residual and Vote	0.816	0.850	-
TCIA	InceptionResNetV2 and Spatial Network	0.868	0.891	-
TCGA	CNN and GDT	0.872	0.938	0.026
TCGA	InceptionResNetV2 and Spatial Network	0.956	0.991	0.022

achieve the accuracy of 86.8 percent for 4 class classification, which outperforms on baseline method [16] by 8.4 percent. ResNet and vote method brings improvement by 4.9 percent. As a comparison, on TCGA dataset CNN with GDT [17] is performed, which adopt several CNN network (ResNet50, InceptionV3, VGG16)

by using ensemble model and extracted features with gradient boost tree classifier. On TCGA dataset, our proposed method achieve 95.6 percent accuracy on 4 class classification and 99.1 percent accuracy, 99.6 percent AUC on 2 class classification (necrosis and non necrosis).
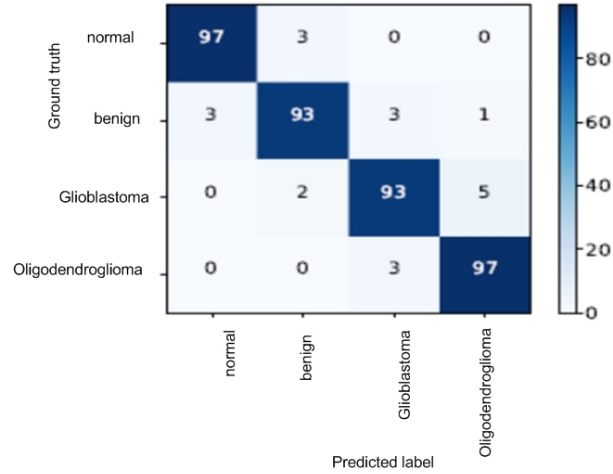

Fig. 6. Confusion matrix (without normalization). 10 fold cross-validation process on 4 class classification with 400 high resolution histology image.

Performance of classification in terms of ROC (receiver operating characteristic curve) and confusion matrix is shown in figure 6 and figure 7. By PyTorch library on NIVIDIA 1080Ti GPU, this performance is achieved. It took around 80ms to classify single high resolution histology image.

9   Conclusion

The proposed method of this paper describes a deep spatial fusion network that handles the complex building of discriminative features over patches. Also learns to adjust bias on patch wise prediction in high resolution histology image. Patch wise InceptionResNetV2 is adopted to extract features from the cellular level to tissue level. This method is exposed to analyze the spatial relationship between patches. Compared to the previous experiments of CNN using various architecture, our proposed method gives better performance. Further, this work can be extended with some other networks that efficiently analyzes more types of malignant tumors other than Glioblastoma and Oligodendroglioma.
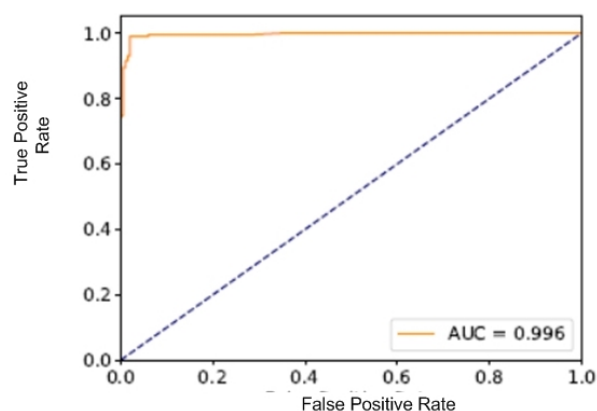
Fig. 7. 2-way classification as necrosis and non necrosis in terms of ROC.

References

[1] H.O. Lyon, A.P. De Leenheer, R.W. Horobin, W.E. Lambert, E.K.W. Schulte, B. Van Liedekerke, D.H. Wittekind, *Standardization of reagents and methods used in cytological and histological practice with emphasis on dyes, stains and chromogenic reagents, Histochem.* J. 26 (7) (1994) 533–544.

[2] Yan Xu, Zhipeng Jia, Yuqing Ai, Fang Zhang, Maode Lai, Eric I-Chao Chang, *Deep Convolutional Activation Features For Large Scale Brain Tumor Histopathology Image Classification And Segmentation*, 2015 IEEE ICASSP

[3] Kiichi Fukuma, Hiroharu Kawanaka, Surya Prasath, Bruce J. Aronow and Haruhiko Takase, *Feature Extraction and Disease Stage Classification for Glioma Histopathology Images*, 2015 17th International Conference on E-health Networking, Application and Services (HealthCom)

[4] Hojjat Seyed Mousavi, Vishal Monga, Ganesh Rao, Arvind U. K. Rao, *Automated discrimination of lower and higher grade gliomas based on histopathological image analysis*, J Pathol Inform 2015, 1-15

[5] Luke Macyszyn, Hamed Akbari, Jared M. Pisapia, Xiao Da, Mark Attiah, Vadim Pigrish,Yingtao Bi, Sharmistha Pal, Ramana V. Davuluri, Laura Roccograndi, Nadia Dahmane, Maria Martinez-Lage, George Biros, Ronald L. Wolf, Michel Bilello, Donald M. O'Rourke, and Christos Davatzikos,*Imaging patterns predict patient survival and molecular subtype in glioblastoma via machine learning techniques*, Neuro-Oncology 2015

[6] Jocelyn Barkerc, Assaf Hoogia, Adrien Depeursingea,b, Daniel L. Rubin,*Automated classification of brain tumor type in whole-slide digital pathology images using local representative tiles*, Elsevier, 2015.

[7] Reid Trenton Powell, Adriana Olar, Shivali Narang, Ganesh Rao, Erik Sulman, Gregory N. Fuller, Arvind Rao, *Identification of Histological Correlates of Overall Survival in Lower Grade Gliomas Using a Bag–of–words Paradigm: A Preliminary Analysis Based on Hematoxylin and Eosin Stained Slides from the Lower Grade Glioma Cohort of The Cancer Genome Atlas*, 2017 Journal of Pathology Informatics

[8] Yan Xu, Zhipeng Jia, Liang-Bo Wang, Yuqing Ai, Fang Zhang, Maode Lai and EricI-Chao Chang, *Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features*, 2017 BMC Bioinformatics.

[9] 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.

[10] Asami Yonekura, Hiroharu Kawanaka,V. B. Surya Prasath, Bruce J. Aronow, Haruhiko Takase, *Automatic disease stage classification of glioblastoma multiforme histopathological images using deep convolutional neural network*, Korean Society of Medical and Biological Engineering and Springer-Verlag GmbH Germany, part of Springer Nature 2018

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

[12] Greig, D.M., Porteous, B.T., Seheult, A.H.: *Exact maximum a posteriori estimation for binary images.* Journal of the Royal Statistical Society. Series B (Methodological) pp. 271–279 (1989)

[13] 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)

[14] Macenko, M., Niethammer, M., Marron, J., Borland, D., Woosley, J.T., Guan, X., Schmitt, C., Thomas, N.E.: *A method for normalizing histology slides for quantitative analysis.* In: Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on. pp. 1107–1110. IEEE (2009)

[15] Kingma, D.P., Ba, J.:*Adam: A method for stochastic optimization.* arXiv preprint arXiv:1412.6980 (2014)

[16] Araujo, T., Aresta, G., Castro, E., Rouco, J., Aguiar, P., Eloy, C., Polonia, A., Campilho, A.:*Classification of breast cancer histology images using convolutional neural networks.* PloS one 12(6), e0177544 (2017)

[17] Rakhlin, A., Shvets, A., Iglovikov, V., Kalinin, A.A.: *Deep convolutional neural networks for breast cancer histology image analysis.* arXiv preprint arXiv:1802.00752 (2018)