Electronic Letters on Science & Engineering 16(2) (2020)
Available online at dergipark.org.tr/else

Grading Brain Histopathological Images Using Deep Residual Networks and Support Vector Machine

Ahmet Haşim YURTTAKAL[1*], Hasan ERBAY[2], Recep Sinan ARSLAN[3]

[1]Department of Computer Technologies, Bozok University, Yozgat, Turkey
0000-0001-5170-6466, ahmet.yurttakal@bozok.edu.tr
[2]Department of Computer Engineering, University of Turkish Aeronautical Association, Ankara, Turkey
0000-0002-7555-541X, herbay@thk.edu.tr
[3]Department of Computer Engineering, Bozok University, Yozgat, Turkey
0000-0002-3028-0416, r.sinan.arslan@bozok.edu.tr

Abstract: Brain cancer is a deadly disease that occurs due to tumour formation in the brain. It can cause weakness in the arms and legs, speech and vision disorders, extremely severe headaches, and symptoms such as vomiting. It is generally classified in four grades. The first and second grades are considered as "low grade", that is, "benign", and the third and fourth grade are considered as "high grade", that is, "malignant". The early grading of the tumour is important for the treatment procedures. Grading brain tumours based on histopathological images is a tiring process that requires expertise. On the other hand, deep learning algorithms are frequently used in computer-aided diagnostic systems. In this study, automatic grading of 1133x40 brain histopathologic images belonging to four phases was performed. First of all, features have been extracted from the latest technology pre-trained Residual networks with the ResNet50 and ResNet101 models. Then, the hyper parameters were optimized by Bayesian Optimization and classified by the Support Vector Machine (SVM) algorithm. 80% of the data set is reserved for training and 20% for testing. When evaluated in terms of multiple classification problems, it is reached a high accuracy rate of 80.09% in ResNet50, while reaching 100% high recall value in Grade I detection in Resnet101.

Keywords: Residual Networks, Support Vector Machine, Histopathology, Grading.

Derin Artık Ağlar ve Destek Vektör Makineleri Kullanılarak Beyin Histopatolojik Görüntülerinin Evrelenmesi

Özet: Beyin kanseri, beyinde tümör oluşumuna bağlı olarak ortaya çıkan ölümcül bir hastalıktır. Kol ve bacaklarda güçsüzlük, konuşma ve görme bozuklukları, aşırı şiddetli baş ağrıları ve kusma gibi semptomlara neden olabilir. Genelde dört sınıfa ayrılır. Birinci ve ikinci sınıflar "düşük dereceli", yani "iyi huylu", üçüncü ve dördüncü sınıflar "yüksek dereceli", yani "kötü huylu" olarak değerlendirilir. Tedavi prosedürleri için tümörün erken evrelenmesi önemlidir. Beyin tümörlerinin histopatolojik görüntülere göre derecelendirilmesi, uzmanlık gerektiren yorucu bir süreçtir. Öte yandan, derin öğrenme algoritmaları bilgisayar destekli teşhis sistemlerinde sıklıkla kullanılmaktadır. Bu çalışmada, dört faza ait 1133x40 beyin histopatolojik görüntülerinin otomatik derecelendirilmesi yapılmıştır. Öncelikle, ResNet50 ve ResNet101 modelleri ile en son teknoloji önceden eğitilmiş Artık ağlardan özellikler çıkarılmıştır. Ardından, hiper parametreler Bayesian Optimizasyonu ile optimize edilerek, Destek Vektör Makinesi (SVM) algoritması ile sınıflandırılmıştır. Veri setinin %80'i eğitim ve %20'si test için ayrılmıştır. Çoklu sınıflandırma problemleri açısından değerlendirildiğinde Resnet50'de %80.09 gibi yüksek bir doğruluk oranına ulaşılırken, Resnet101'de Grade I tespitinde %100 yüksek duyarlılık değerine ulaşılmaktadır.

Anahtar Kelimeler:Artık Ağlar, Destek Vektör Makineleri, Histopatoloji, Evreleme.

Reference to this paper should be made as follows
A.H. Yurttakal, H. Erbay, R.S. Arslan 'Grading Brain Histopathological Images Using Deep Residual Networks and Support Vector Machine', Elec Lett Sci Eng , vol. 16(2) , (2020), 78-84

I. INTRODUCTION

The tumor is an abnormal growth of tissues in a particular area as a result of excessive division of cells in the body. Tumours that occur in the brain tissue are called as brain tumours. They are collected in two main classes: gliomas and non-gliomas. Gliomas originate from glia cells, which are cells that protect and support nerve cells [1]. Astrocytoma is the type of glioma and develops from star-shaped cells called astrocytes. High-grade astrocytoma's grow faster than low-grade ones and are often treated with a combination of surgery, radiation, and chemotherapy. Early staging of the tumor is very important to determine the treatment processes [2].

Generally, brain tumours are divided into 4 stages. Grade-1 tumours are considered as the most benign tumours, while Grade-4 are considered as the most malign ones. Histopathological staging of brain tumours, changes in cell morphology and tissue structure are performed by observing under light microscopy in routine dyes such as Hematoxylin-Eosin (H&E). This process is tiring and difficult [3].

The computer is a machine that can make long and very complex operations at a great speed, make decisions based on logical connections and executes transactions. Increased processing capacities in these days make it possible for data scientists to use the computer in many areas of life, including complex procedures such as medical diagnostics, speech recognition, image processing. Here, deep learning algorithms have achieved high success rate in the classification of histopathological images. Coudray et al. [4] predicted mutation by classifying images of non-small cell lung cancer histopathology using the inception v3 deep learning model. Khan et al. [5] used three deep learning architecture (GoogleNet, VGGNet and ResNet) for the classification of breast histopathological images. In another study, Nirschl et al. [6] performed segmentation on cardiac histopathology images.

In this study, features were extracted from the brain histopathological images of 4 different grades using residual networks such as Resnet50, Resnet101. Then obtained features were classified with Support Vector Machine (SVM) algorithm. Some of the important contributions of the study to the literature are comparing the performances of residual networks, using hybrid of Residual networks and SVM algorithm, and classifying difficult images such as brain histopathology of 4 different grades.

In the second part, dataset and proposed algorithm are presented. The results obtained are discussed in the third section. Conclusion and suggestions are given in the last section.

II. MATERIAL AND METHODS

A. Dataset

Medisp Histology Image Collection Library (HICL) dataset was used in this study [7]. In the original dataset, images of brain histopathology consist of 93 brain cancer cases. The raw clinical pictures were collected at Patras University Hospital in Greece. There are 2548 H&E stained images of different degrees in x20 and x40, including 840 low-grade cases, 1608 high-grade cases, and 100 uncertain cases. The light microscope imaging system includes a LEICA DM 2500 microscope connected to a LEICA DFC 420C camera (Leica Microsystems GmbH) [8]. Here, a total of 1133 different images of Grade I 132, Grade II 210, Grade II 434 and Grade IV 357 belonging to x40 degrees were classified.

B. Proposed Method

Deep convolutional neural network (CNN) has achieved higher success rate than traditional handmade feature extraction methods. AlexNet has learned complex features with 8 layers, 5 of which are convolution [9]. Theoretically, it is thought that as the number of layers of the model increases, the performance will rise. However, it has been experienced that this is not really the case. He et al. [10] showed that the error rate of a 56-layer CNN model was higher than that of a

20-layer model. For this reason, residual network structure was proposed. Fig. 1 shows the residual block structure.

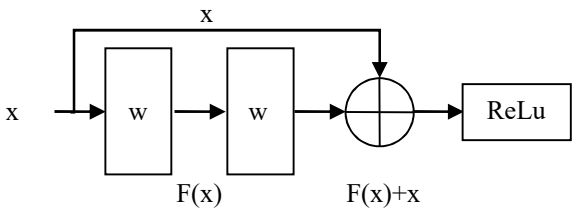


Fig. 1 Residual block

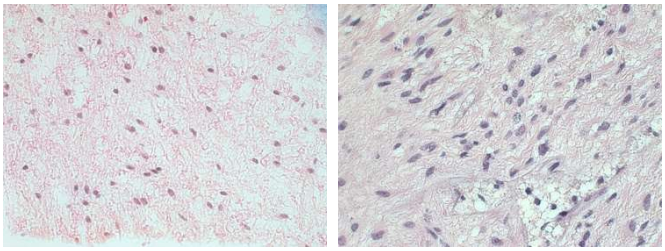
Residual networks learned feature representations on more than 1 million images of 1000 different classes in the ImageNet database [11]. ResNet is the abbreviation of Residual network and it is a network used for computer vision problems. The ResNet network uses redundant connections where gradients can flow directly to prevent gradients from becoming zero after the chain rule is applied. ResNet 101 and Resnet 50 are convolutional neural networks and consist of 101 and 50 deep layers, respectively. The dimensions of the image in the input layer are 224x224x3. Details of ResNet50 and ResNet101 network structures used in this study are shown in Table 1 .

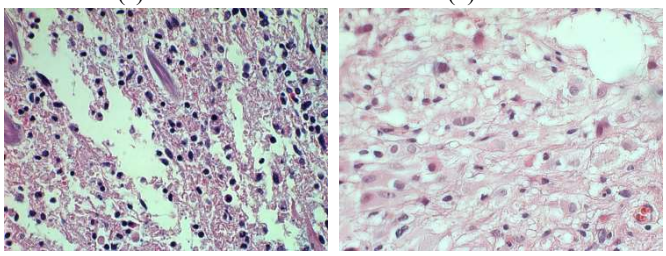
Table1 ResNet50 and ResNet101 Architecture [10]

Layer Name	**Output Size**	**ResNet50**	**ResNet101**
Conv1	112x112	3x3 max pool stride 2	7x7, 64, stride 2
Conv2	56x56	$\begin{bmatrix} 1x1,64 \\ 3x3,64 \\ 1x1,256 \end{bmatrix}$ x3	3x3 max pool stride2 $\begin{bmatrix} 1x1,64 \\ 3x3,64 \\ 1x1,256 \end{bmatrix}$ x3
Conv3	28x28	$\begin{bmatrix} 1x1,128 \\ 3x3,128 \\ 1x1,512 \end{bmatrix}$ x4	$\begin{bmatrix} 1x1,128 \\ 3x3,128 \\ 1x1,512 \end{bmatrix}$ x4
Conv4	14x14	$\begin{bmatrix} 1x1,256 \\ 3x3,256 \\ 1x1,1024 \end{bmatrix}$ x6	$\begin{bmatrix} 1x1,256 \\ 3x3,256 \\ 1x1,1024 \end{bmatrix}$ x23
Conv5	7x7	$\begin{bmatrix} 1x1,512 \\ 3x3,512 \\ 1x1,2048 \end{bmatrix}$ x3	$\begin{bmatrix} 1x1,512 \\ 3x3,512 \\ 1x1,2048 \end{bmatrix}$ x3
	1x1	Average pool, 1000-D Fully Connected, Softmax	Average pool, 1000-D Fully Connected, Softmax

III. EXPERIMENTAL RESULTS

Grading process was carried out in Matlab environment. The dimensions of the images are 1728x1296. Fig. 2 shows sample images from raw dataset.





(a) Grade I (b) Grade II

(c) Grade III (d) Grade IV

Fig. 2 Sample images

Randomly 20% of dataset was reserved for testing. Table 2 shows training and testing images statistics.

Table 2 Training and testing statistics

Groups	Grade I	Grade II	Grade III	Grade IV	Total
Training	106	168	347	286	907
Testing	26	42	87	71	226
Total	132	210	434	357	1133

The images were automatically resized to the network. Later the feature extraction phase was done. This process is an effective method to determine the representation power of pre-trained residual networks. It can be used as an effective method, especially when GPU capacity is limited. The deeper layer contains high-level features. Therefore, the features obtained from the fully connected layer were classified with the SVM algorithm. Using automatic hyper-parameter optimization, hyper-parameters were found to minimize the loss of 5 fold cross validation. Fig. 3 shows ResNet50 hyper-parameter optimization process. The blue line shows the minimum observed objective value in each iteration. The green line shows the estimated minimum objective value in each iteration. The estimation of minimum objective functions shows the difference between estimated minimum objective and the real minimum one. It is seen that the estimated objective value is very close to the observed objective value in each iteration. As a result, the optimization performed was found to be successful.

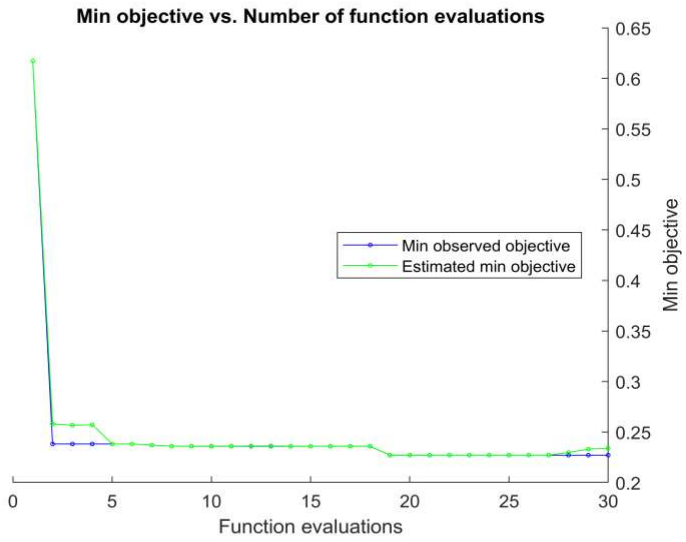


Fig. 3 ResNet50 hyper-parameter optimization process

Bayesian optimization was used while optimizing hyper-parameters. Table 3 shows optimization results. Observed Objective Value, Estimated Objective Value, Evaluation Time are given as a result of the bayesian optimization of the proposed models using ResNet50 and ResNet101. As ResNet101 is a deeper model, evaluation time is more. In both models, the observed and estimated values as a result of optimization are close.

Table 3 Optimization results

Models	**Observed Objective Function Value**	**Estimated Objective Function Value**	**Function Evaluation Time**
ResNet50	0.22712	0.23396	0.98599
ResNet101	0.2194	0.22142	77.7459

Classification success of properties on test data demonstrated by confusion matrices. Fig. 4 shows ResNet50 confusion matrix.

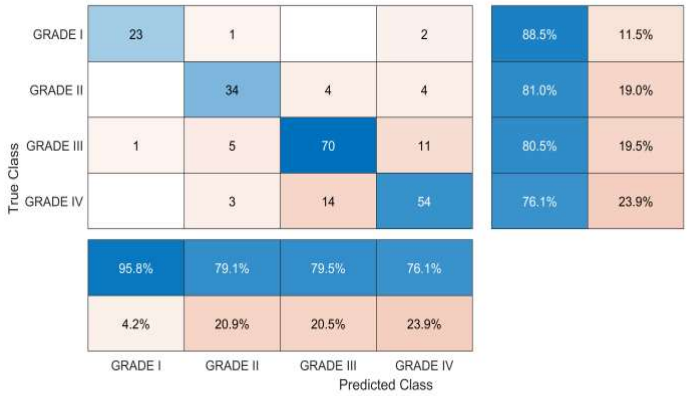


Fig. 4 ResNet50 confusion matrix

In the features extracted with ResNet50, while the accuracy rate of the model reached 98.23% in Grade 1 calculation, 92.48% prediction rate was achieved in Grade 2 estimation. ResNet50 appears to achieve excellent sensitivity in automatic grading of brain histopathological images of four phases. It is critical to get successful results for each phase. The misclassification rate for each grade has remained very low. The incorrect classification number was 45 for a total of 226 samples. In addition, the classification level of the patient samples in Grade2 and Grade3 phases is very low. Total accuracy rate is 90.45% for all grades. Table 4 shows other performance statistics.

Table 4 ResNet50 performance statistics

Grades	**Accuracy(%)**	**Precision**	**Recall**	**F1 Score**
Grade I	98.23	0.96	0.88	0.92
Grade II	92.48	0.79	0.81	0.80
Grade III	84.51	0.80	0.80	0.80
Grade IV	84.96	0.76	0.76	0.76
Avg	***90.45***	***0.83***	***0.81***	***0.82***

Precision and Recall values were obtained from the confusion matrix. The value F1 represents the harmonic average of the precision and recall values. Fig. 5 shows ResNet101 confusion matrix.

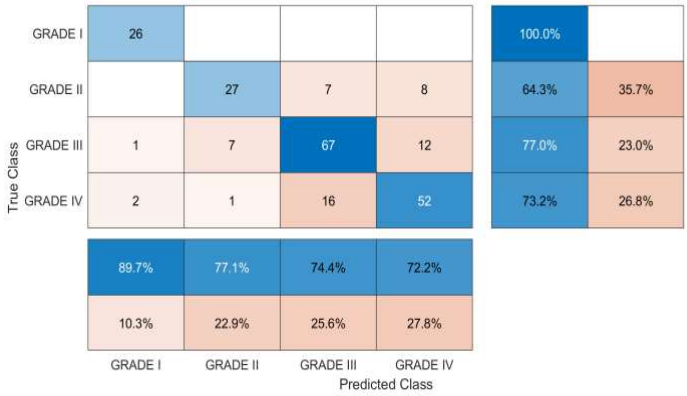


Fig. 5 ResNet101 confusion matrix

In the features obtained with ResNet101, the model achieved a 98.67% accuracy rate in the Grade I estimate and it achieved an 89.82% accuracy rate in the Grade II prediction. As with Resnet50, success results have been obtained in the tests performed with Resnet101. 54 incorrect classifications were made for a total of 226 samples. Table 5 shows other performance statistics.

Table 5 ResNet101 performance statistics

Grades	**Accuracy(%)**	**Precision**	**Recall**	**F1 Score**
Grade I	98.67	0.90	1.00	0.95
Grade II	89.82	0.77	0.64	0.70
Grade III	80.97	0.74	0.77	0.76
Grade IV	82.74	0.72	0.73	0.73
Avg	***88.05***	***0.78***	***0.79***	***0.79***

Considering all the results obtained, ResNet50 model was found to be more successful with an average of 90.45% in terms of classification of all stages. ResNet101 was more successful than ResNet50 in determining Grade I, where the differentiation is the least. The reason for this is thought to be due to the deeper architecture of ResNet101. However, since the differentiation in Grade III and Grade IV is more than other stages, it has lagged behind ResNet50.

In the tests performed in this study, SVM, which is a widely used classification method, was used together with CNN. By making tests with different classifiers, it may be possible to compare them and produce more successful results. However, the model will also have to be changed accordingly. For this reason, studies to be carried out using a new model with different classifiers will be evaluated later. For this reason, SVM was preferred in this study because it is only a widely used method.

IV. CONCLUSIONS

In the article, the use of residual networks with the SVM algorithm has been examined. Classification successes of ResNet50 and ResNet101 models with SVM algorithm were compared. The proposed model has many advantages. The problem is the staging of histopathological images of the brain. It is a difficult problem in terms of multiple classifications. The study was automatically performed on the original images, with minimal pre-processing. It also reached a high accuracy rate in a short time, with a minimum GPU requirement. The number of images can be increased in future studies. Hybrid methods can be used with different networks.

ACKNOWLEDGMENT

There is no conflict of interest among the authors.

REFERENCES

[1] L. M. DeAngelis, “Brain tumors,” *New England journal of medicine,* vol. 344, no. 2, pp. 114-123, 2001.

[2] W. Penfield, “The classification of gliomas and neuroglia cell types,” *Archives of Neurology & Psychiatry,* vol. 26, no. 4, pp. 745-753, 1931.

[3] K. Herholz, K.-J. Langen, C. Schiepers, and J. M. Mountz, "Brain tumors.", *Seminars in nuclear medicine*, vol. 42, no. 6, pp. 356-370, 2012.

[4] N. Coudray, P. S. Ocampo, T. Sakellaropoulos, N. Narula, M. Snuderl, D. Fenyö, A. L. Moreira, N. Razavian, and A. Tsirigos, “Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning,” *Nature medicine,* vol. 24, no. 10, pp. 1559-1567, 2018.

[5] S. Khan, N. Islam, Z. Jan, I. U. Din, and J. J. C. Rodrigues, “A novel deep learning based framework for the detection and classification of breast cancer using transfer learning,” *Pattern Recognition Letters,* vol. 125, pp. 1-6, 2019.

[6] J. J. Nirschl, A. Janowczyk, E. G. Peyster, R. Frank, K. B. Margulies, M. D. Feldman, and A. Madabhushi, "Deep learning tissue segmentation in cardiac histopathology images," *Deep Learning for Medical Image Analysis*, pp. 179-195: Elsevier, 2017.

[7] S. Kostopoulos, P. Ravazoula, P. Asvestas, I. Kalatzis, G. Xenogiannopoulos, D. Cavouras, and D. Glotsos, “Development of a reference image collection library for histopathology image processing, analysis and decision support systems research,” *Journal of digital imaging,* vol. 30, no. 3, pp. 287-295, 2017.

[8] D. Glotsos, I. Kalatzis, P. Spyridonos, S. Kostopoulos, A. Daskalakis, E. Athanasiadis, P. Ravazoula, G. Nikiforidis, and D. Cavouras, “Improving accuracy in astrocytomas grading by integrating a robust least squares mapping driven support vector machine classifier into a two level grade classification scheme,” *computer methods and programs in biomedicine,* vol. 90, no. 3, pp. 251-261, 2008.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks.", *Neural information processing systems*, vol. 25, no. 2, pp. 1097-1105, 2012.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition.", IEEE conf. computer vision and pattern recognition(CVPR), pp. 770-778, 2016.

[11] I. L. S. V. R. Challenge, “Available online: http://www. image-net. org/challenges,” *LSVRC/(accessed on 26 February 2019)*, 2014.