NATURENGS
E-ISSN 2717-8013
https://dergipark.org.tr/en/pub/naturengs

Automated Grading of Glioma Using Deep Neural Networks

Muhammed Yildirim [a,*], Serpil Aslan [b], Emine Cengil [c], Sercan Yalçın [d]

[a] Malatya Turgut Ozal University, Department of Computer Engineering, Malatya Türkiye - 44210
[b] Malatya Turgut Ozal University, Department of Software Engineering, Malatya Türkiye - 44210
[c] Bitlis Eren University, Department of Computer Engineering, Bitlis Türkiye - 13100
[d] Adıyaman University, Department of Computer Engineering, Adıyaman Türkiye - 02040

* corresponding author

ARTICLE INFO

Received 28.11.2023
Accepted 18.12.2023

Doi: 10.46572/naturengs.1397010

ABSTRACT

Gliomas are one of the most common tumors in the brain. Gliomas can be classified as Low-Grade Glioma (LGG) and Glioblastoma Multiforme (GBM). Clinical and molecular/mutation factors come to the fore in the grading of gliomas. Molecular tests used to grade glioma are expensive and time consuming. A new deep learning-based model has been developed for glioma grading to reduce the workload of experts and to be able to use these computer-aided systems in non-specialist locations. In the proposed model, long short-term memory (LSTM) and Convolutional neural network (CNN) were used together. The developed model was also compared with six different classifiers accepted in the literature. The developed model achieved the highest performance among the models used in the study.

Keywords: Classifiers, CNN, Deep Learning, Glioma, LSTM

1. Introduction

Gliomas are a type of tumor that most often occurs in the brain or spinal cord. Gliomas arise from glial cells, which are cells that form the support tissue of the nervous system. Gliomas usually occur as a result of abnormal and uncontrolled proliferation of cells. This abnormal growth can put pressure on normal brain tissue and disrupt nervous system functions. Gliomas are divided into several types. Gliomas can be low-grade or high-grade. High-grade gliomas tend to grow faster and are more difficult to treat [1, 2].

Glioma symptoms may differ depending on the region of the brain where it is located. Common symptoms include headache, nausea, vomiting, seizures, memory and concentration problems, loss of coordination, and behavioral changes. These symptoms may differ from person to person and may be associated with other health problems, so medical support is required to make a diagnosis [3].

Treatment options for gliomas usually include surgery, radiotherapy, and chemotherapy. Treatment options are determined depending on the type of tumor, its size, location, and the patient's general health condition. Surgical intervention is performed to remove as much of the tumor as possible. Radiotherapy and chemotherapy are other treatment methods used to control the growth of the tumor or shrink it [4].

Gliomas usually cannot be cured or completely eliminated, but treatments are used to relieve symptoms and control tumor growth. After treatment, it is important to follow up with patients and have regular check-ups [5]. There are studies in the literature on glioma grading. While most of these studies use MR images, there are some studies using molecular features.

Tasci et al. used 2 different data sets in their study for glioma grading. Researchers classified the features they obtained using the Lasso feature selection method into 5 different supervised classifiers. In this study, the researchers obtained accuracy values of 87.60% and 79.66%, respectively [6].

Cengil et al. In their study, they performed the grading and localization of glioma and meningioma tumors. While the researchers used the Efficientnet architecture for feature extraction, they used the PANet network to create the feature pyramid. Finally, object detection was performed using YOLO [7].

Yang et al. In their study, they used MRI images for glioma grading. In the study, researchers preferred to use Googlenet and Alexnet architectures, which are CNN architectures. In the study, they obtained accuracy values of 86.7%, 90.9%, and 93.9%. In this study conducted using transfer learning methods, researchers stated that these architectures can be used in glioma grading [8].

* Corresponding author. e-mail address: muhammed.yildirim@ozal.edu.tr
ORCID : 0000-0003-1866-4721

Xiao et al. In this study for glioma grading, they used the BraTS data set consisting of 285 subjects. In the study, three different feature groups were obtained with the VGG method. The obtained features were classified in different classifiers. In their study, the researchers obtained an AUC value of 94.4% [9].

Molecular tests used to grade glioma are expensive and time-consuming [10]. To avoid this disadvantage and alleviate the workload of experts, deep learning networks were used for glioma grading in this study. Thanks to this computer-aided system, the developed model can be used for preliminary diagnosis in non-expert places. CNN and LSTM layers were used together for glioma grading. In this way, a more effective model was brought to the fore.

In the rest of the article, the methods used in the study were examined, then the results were presented and the article was completed with the conclusion section.

2. Background

In this section, the model developed for glioma grading, classifiers, and the glioma dataset used in the study are examined.

2.1. Proposed Model for Glioma Grading

In this paper, an LSTM and CNN-based model was developed for Glioma grading. Convolution, Max pooling, Dropout, LSTM, Flatten, and Dense layers were used in the developed model. The model developed using CNN and LSTM architecture is presented in Figure 1.

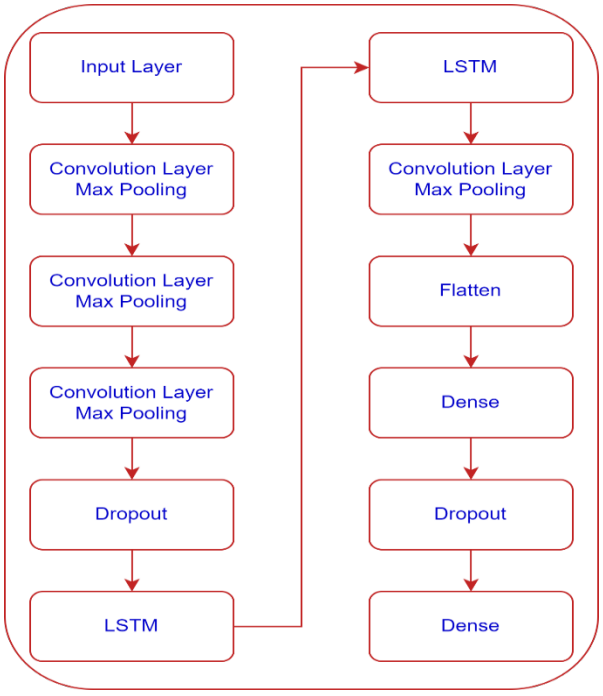

Figure 1. Proposed model for glioma grading

The model developed for glioma grading is summarized in Figure 2.

Layer (type)	Output Shape	Param #
conv1d_36 (Conv1D)	(None, 23, 16)	48
max_pooling1d_36 (MaxPooling1D)	(None, 12, 16)	0
conv1d_37 (Conv1D)	(None, 12, 32)	1056
max_pooling1d_37 (MaxPooling1D)	(None, 6, 32)	0
conv1d_38 (Conv1D)	(None, 6, 64)	4160
max_pooling1d_38 (MaxPooling1D)	(None, 3, 64)	0
dropout_24 (Dropout)	(None, 3, 64)	0
lstm_27 (LSTM)	(None, 3, 32)	12416
lstm_28 (LSTM)	(None, 64)	24832
flatten_12 (Flatten)	(None, 64)	0
dense_24 (Dense)	(None, 32)	2080
dropout_25 (Dropout)	(None, 32)	0
dense_25 (Dense)	(None, 2)	66

Total params: 44658 (174.45 KB)
Trainable params: 44658 (174.45 KB)
Non-trainable params: 0 (0.00 Byte)

Figure 2. Summary of the proposed model

When grading glioma in the proposed model, 70% of the data in the data set was used for training and 30% was used for testing. In the recommended model, batch size 64 and epoch value 210 are selected.

2.2. Classifiers

In the study, a model was developed for glioma grading. To test the performance of the developed model, results were also obtained for 6 different classifiers. The classifiers used for glioma grading are explained respectively.

AdaBoost is an ensemble algorithm consisting of weak learners. The AdaBoost algorithm sequentially trains weak classifiers using the weights of the examples in the dataset, usually black box algorithms, and combines them to create a strong classifier. Each weak classifier works harder on data samples focusing on previous errors, thus getting better over time. In addition to being a successful classification algorithm, AdaBoost is the basis for many learning algorithms that provide good results in various application areas [11, 12].

Random Forest is an algorithm used in classification and regression problems in machine learning. Random Forest creates an ensemble feature by combining multiple decision trees. While each tree may have limited ability to make predictions, the combination of many trees provides more accurate and stable predictions. The Random Forest algorithm uses the principle of randomness when creating the decision tree to increase the diversity of the structure and prediction of each tree. This randomness occurs first by randomly sampling data samples and also by randomly selecting features [13].

Naive Bayes algorithm is based on the Bayes theorem. Bayes' theorem is used to calculate the probability of one event occurring, given that another event occurs. The Naive Bayes algorithm applies Bayes' theorem to classification problems. It creates a model using a pre-given labeled dataset and uses this model to classify the test data. Naive Bayes classifier models the relationship between different classes in the classification problem. For a test sample, it calculates the conditional probabilities of the classes and predicts the class with the highest probability. Naive Bayes assumes that the probability values found by trial and error are independent. Since independence is assumed between features, it does not need a model structure to predict relationships between features in the dataset [14].

XGBoost is a machine learning algorithm used for classification and regression problems. It is based on the gradient boosting method and makes predictions by combining multiple trees. XGBoost shows high performance, especially on tabular data. Unlike models based on Boosting methods, XGBoost uses a special regularization term as well as the error function in the Gradient Boosting method when building trees. This reduces overfitting and provides better generalization. XGBoost is a simple and effective machine-learning algorithm and can be applied to many different problems [15].

LightGBM is a classification and regression algorithm known for being fast and scalable. It was developed by Microsoft and is faster and higher performing than other commonly used gradient boosting methods. LightGBM is especially advantageous when working with large data sets. This algorithm is known for its low memory usage and fast training time. It can run quickly on multi-core CPUs using parallel processing capabilities. In classification problems, LightGBM generally has an accurate classification rate and high prediction performance. Additionally, it is easy to tune hyperparameters and provides good scalability, making LightGBM a popular choice [16].

KNN is basically a sample-based classification algorithm that classifies a new data point based on the labels of neighboring points around it. As its working principle, the KNN algorithm classifies a new data point that we want to classify by determining its k nearest neighbors among previously labeled data points. The class labels of these neighbors are often used and the new data point is assigned to that class. An important parameter of KNN is the k value, this value determines the number of neighbors. As the value of k increases, the model focuses more on the surroundings, and at lower values, the model can become more specific. However, choosing the k value correctly can determine the effectiveness of KNN [17, 18].

2.3. Dataset

The dataset used for glioma grading in the study was downloaded from the UCI Machine Learning repository. In the relevant dataset, 20 genes and 3 clinical features were considered for glioma grading. Diagnosis of Lower-Grade Glioma (LGG) and Glioblastoma Multiforme (GBM) was performed using 23 features. The relevant dataset was funded by The Cancer Genome Atlas (TCGA) Project [6, 19].

3. Experimental Results

The application results of the model developed for glioma grading were obtained in the Python environment. The performance of the model developed for glioma grading was compared with classifiers accepted in the literature. Accuracy (ACC), Sensitivity (SEN), Specificity (SPC), Negative Predictive Value (NPV), False Positive Rate (FPR), False Discovery Rate (FDR), False Negative Rate (FNR), Matthews Correlation Coefficient (MCC) and F1-score metrics were used to compare the performances of the developed model and classifiers [20].

3.1. Results of the Proposed Model

The confusion matrix obtained in the model developed for glioma grading is presented in Figure 3.

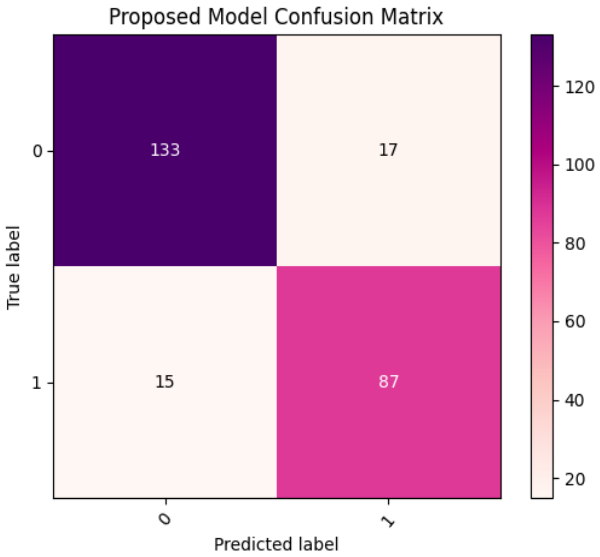

Figure 3. Proposed model Confusion Matrix (0-LGG, 1-GBM)

When Figure 3 is examined, the model proposed for glioma grading correctly predicted that 133 of the test data of 150 patients belonging to the LGG class were LGG, while it predicted the LGG data of 17 patients as GBM. While the proposed model correctly predicted 87 of the GBM data of 102 patients, it incorrectly predicted the data of 15 patients. The accuracy curve of the proposed model is presented in Figure 4, and the loss curve is presented in Figure 5.

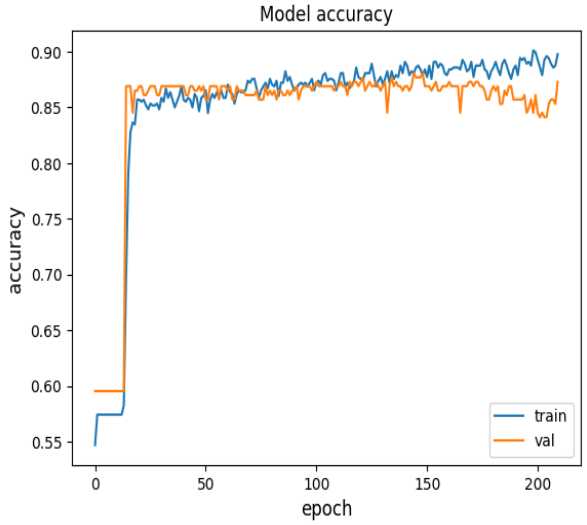

Figure 4. Accuracy curve of proposed model

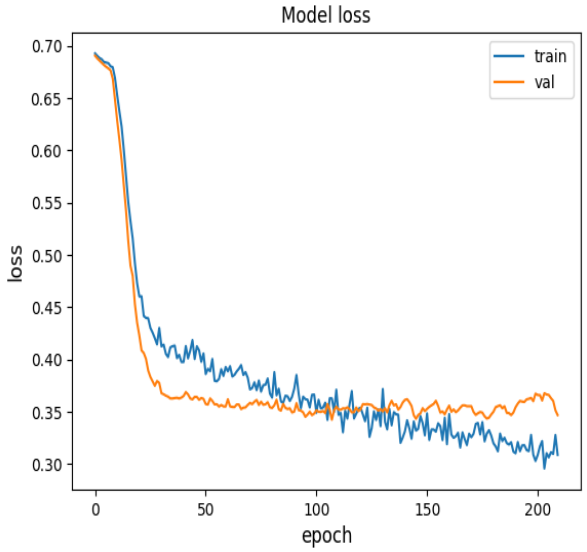

Figure 5. Loss Curve of proposed model

Performance measurement metrics obtained in the model proposed for glioma grading are presented in Table 1.

When Table 1 is examined, it is seen that the model proposed for glioma grading produces successful results. The accuracy value of the model proposed for glioma grading is 87.30%.

Table 1. Performance metrics of proposed model (%)

Model	Performance
ACC	87.30
SPC	83.65
SEN	89.86
NPV	85.29
FPR	16.35
FDR	11.33
FNR	10.14
MCC	73.74
F1-Score	89.26

3.2. Results of the Classifiers

In order to evaluate the performance of the model developed for glioma grading, glioma grading was also done using classical machine learning methods. The first model used for comparison is Adaboost. The confusion matrix and learning curve obtained in the Adaboost classifier are shown in Figure 6 and the performance metrics of Adaboost are given in Table 2.

Table 2. Performance metrics of Adaboost (%)

Model	Performance
ACC	86.90
SPC	78.05
SEN	95.35
NPV	94.12
FPR	21.95
FDR	18
FNR	4.65
MCC	74.75
F1-Score	88.17

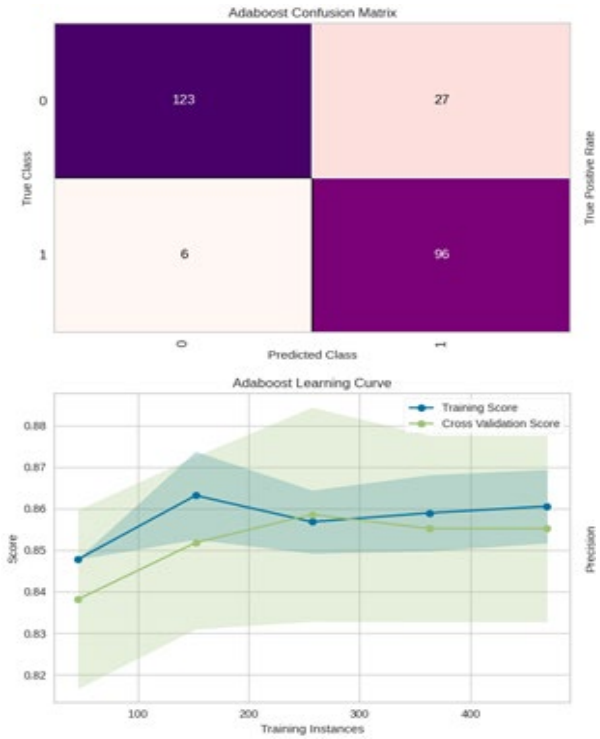

Figure 6. Confusion matrix and Learning Curve of Adaboost

When Figure 6 is examined, it is seen that the Adaboost classifier classified 123 of 150 LGG data correctly and 27 incorrectly. In the GBM class, it is seen that it correctly predicted 96 of 102 patient data as GBM and misclassified the data of 6 patients as LGG. The accuracy value achieved by the Adaboost classifier in glioma grading was 86.90%. The second classifier used in glioma grading is Random Forest. The confusion matrix and learning curve obtained when glioma grading is performed with the Random Forest classifier are shown in Figure 7.

When Figure 7 is examined, it is seen that the Random Forest classifier classified 129 of 150 LGG data correctly and 21 incorrectly. In the GBM class, it is seen that it correctly predicted 87 of 102 patient data as GBM and misclassified the data of 15 patients as LGG. The accuracy value obtained by the Random Forest classifier in glioma grading was 85.71%. Performance metrics of Random Forest are given in Table 3.

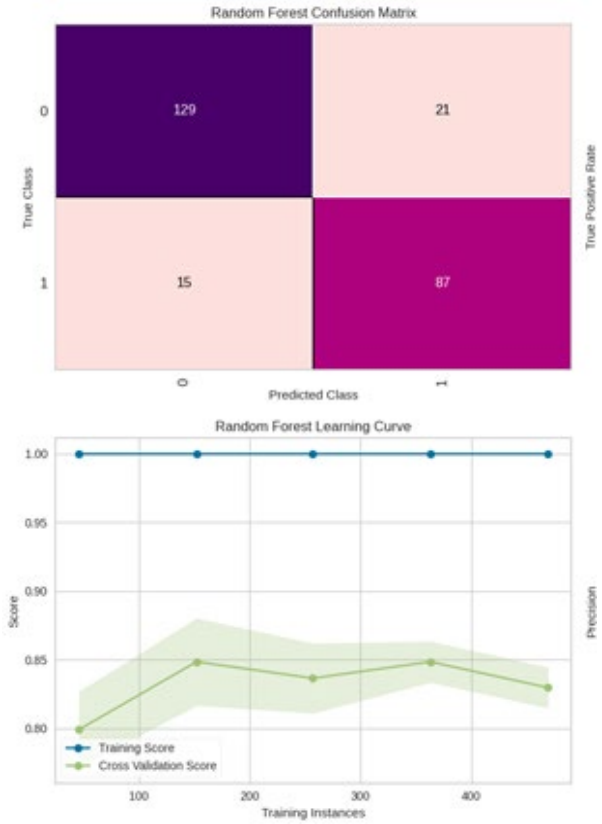

Figure 7. Confusion matrix and learning curve of Random Forest

Table 3. Performance metrics of Random Forest (%)

Model	Performance
ACC	85.71
SPC	80.56
SEN	89.58
NPV	85.29
FPR	19.44
FDR	14.00
FNR	10.42
MCC	70.71
F1-Score	87.76

The third classifier used in glioma grading is Naïve Bayes. The confusion matrix and Learning Curve obtained when glioma grading is performed with the Naïve Bayes classifier is shown in Figure 8.

When Figure 8 is examined, it is seen that the Naïve Bayes classifier classified 115 of 150 LGG data correctly and 35 as incorrect. In the GBM class, it is seen that it correctly predicted 96 of 102 patient data as GBM and misclassified the data of 6 patients as LGG. The accuracy value obtained by the Naïve Bayes classifier in glioma grading was 83.73%. Performance metrics of Naïve Bayes are given in Table 4.

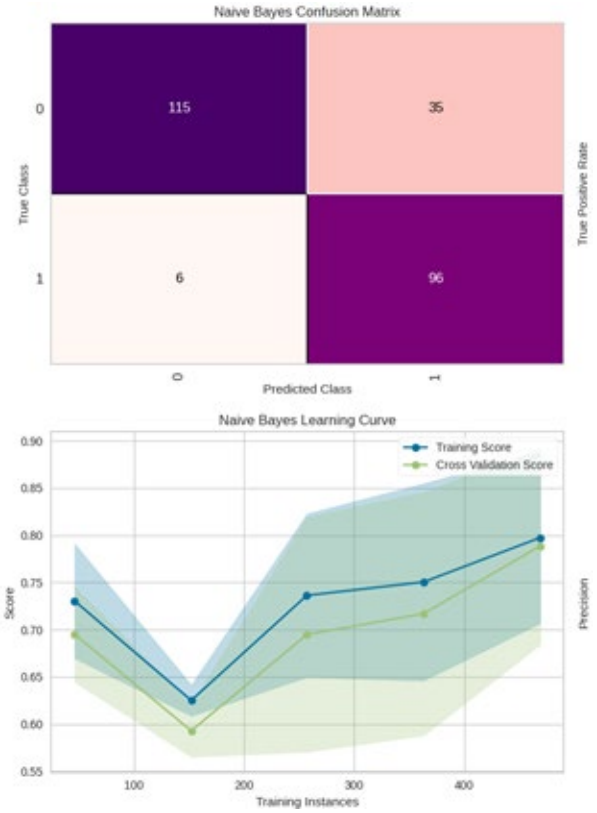

Figure 8. Confusion matrix and learning curve of Naïve Bayes

Table 4. Performance metrics of Naïve Bayes (%)

Model	Performance
ACC	83.73
SPC	73.28
SEN	95.04
NPV	94.12
FPR	26.72
FDR	23.33
FNR	4.96
MCC	69.54
F1-Score	84.87

The fourth classifier used in glioma grading is XGBoost. The confusion matrix and Learning Curve obtained when glioma grading is performed with the XGBoost classifier is shown in Figure 9.

When Figure 9 is examined, it is seen that the XGBoost classifier predicted 127 of 150 LGG patient data correctly and 23 incorrectly. In the GBM class, it is seen that 82 of 102 patient data were correctly predicted as GBM and 20 patients' data were incorrectly predicted as LGG. The accuracy value obtained by the XGBoost classifier in glioma grading was 82.94%. Performance metrics of XGBoost are given in Table 5.

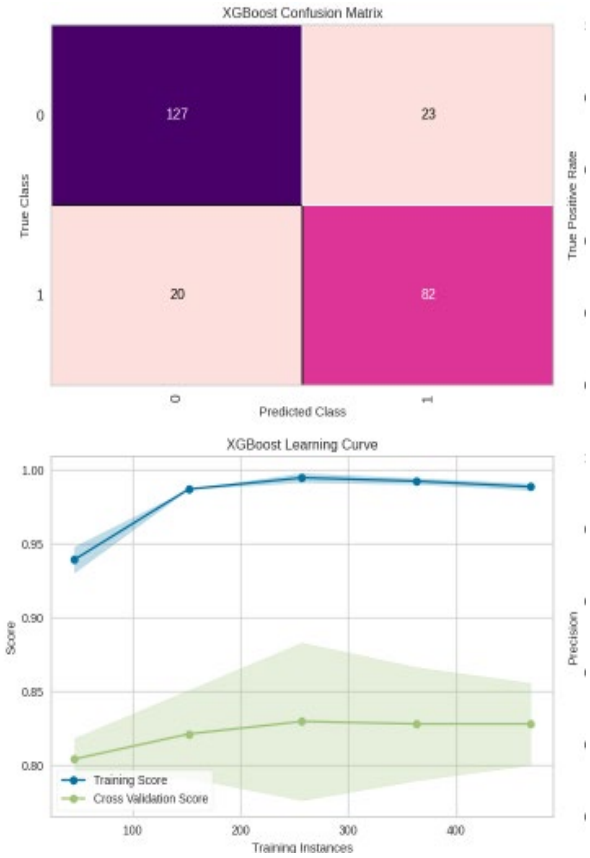

Figure 9. Confusion matrix and learning curve of XGBoost

Table 5. Performance metrics of XGBoost (%)

Model	Performance
ACC	82.94
SPC	7810
SEN	86.39
NPV	80.39
FPR	21.90
FDR	15.33
FNR	13.61
MCC	64.77
F1-Score	85.52

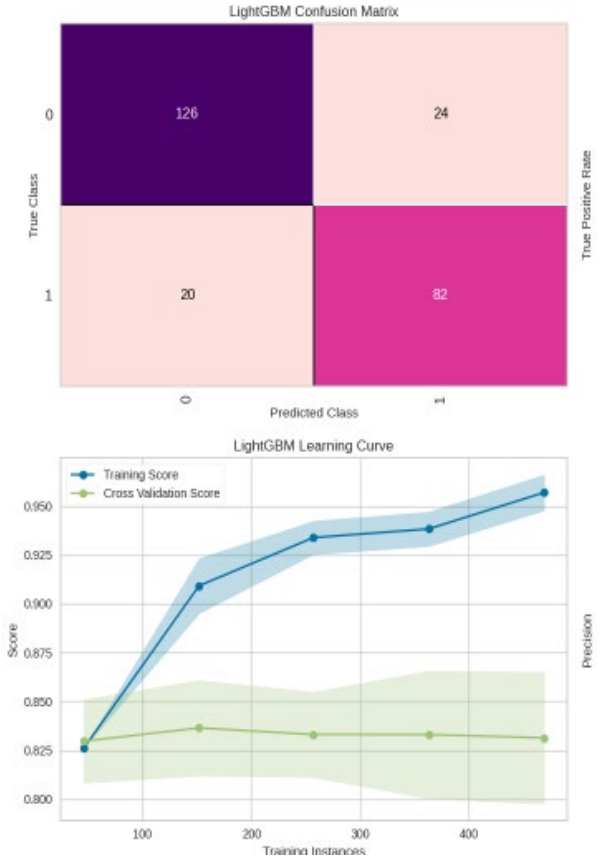

Figure 10. Confusion matrix and learning curve of LightGBM

Table 6. Performance metrics of LightGBM (%)

Model	Performance
ACC	82.54
SPC	77.36
SEN	86.30
NPV	80.39
FPR	22.64
FDR	16.00
FNR	13.70
MCC	64.02
F1-Score	85.14

When Figure 10 is examined, it is seen that the LightGBM classifier predicted 126 of 150 LGG patient data correctly and 24 incorrectly. In the GBM class, it is seen that 82 of 102 patient data were correctly predicted as GBM and 20 patients' data were incorrectly predicted as LGG. The accuracy value obtained by the LightGBM classifier in glioma grading was 82.54%. The performance metrics of LightGBM are given in Table 6.

Another classifier used in glioma grading is LightGBM. The confusion matrix and Learning Curve obtained when glioma grading is performed with the LightGBM classifier are shown in Figure 10.

Another classifier used in glioma grading is KNN. The confusion matrix and Learning Curve obtained when glioma grading is performed with the KNN classifier are shown in Figure 11.

When Figure 11 is examined, it is seen that the KNN classifier predicted 122 of 150 LGG patient data correctly and 28 incorrectly. In the GBM class, it is seen that 82 out of 102 patient data were correctly predicted as GBM and 20 patients' data were incorrectly predicted as LGG. While the KNN classifier correctly predicted 204 of 252 test data, it incorrectly predicted 48 test data. The accuracy value obtained by the KNN classifier in glioma grading was 80.95%. The performance metrics of LightGBM are given in Table 7.

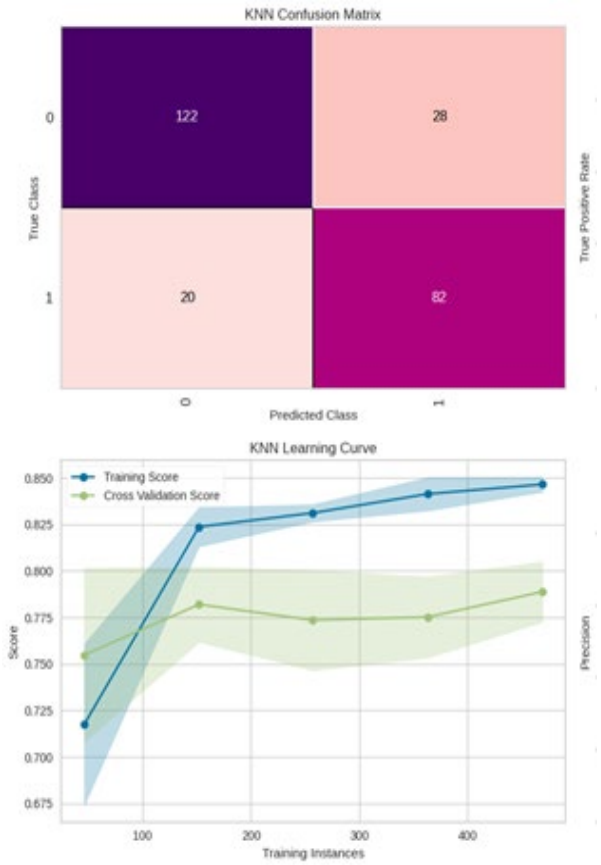

Figure 11. Confusion matrix and learning curve of KNN

Table 7. Performance metrics of KNN (%)

Model	Performance
ACC	80.95
SPC	74.55
SEN	85.92
NPV	80.39
FPR	25.45
FDR	18.67
FNR	14.08
MCC	61.09
F1-Score	83.56

3.3. Comparison of all Models

In this study for glioma grading, a model consisting of CNN and LSTM structures was developed. The developed model was classified with 6 different classifiers. The accuracy values obtained from the models used in the study are presented in Figure 12.

The highest accuracy value in glioma grading was achieved in the proposed model at 87.30%. This was followed by AdaBoost at 86.9%, Random Forest at 85.71%, Naive Bayes at 83.73%, XGBoost at 82.93%, LightGBM at 82.53%, and KNN classifiers at 80.95%, respectively. Using CNN and LSTM networks together has a great impact in achieving a higher accuracy value in the proposed model. CNN and LSTM networks are frequently used in the literature [21-23].

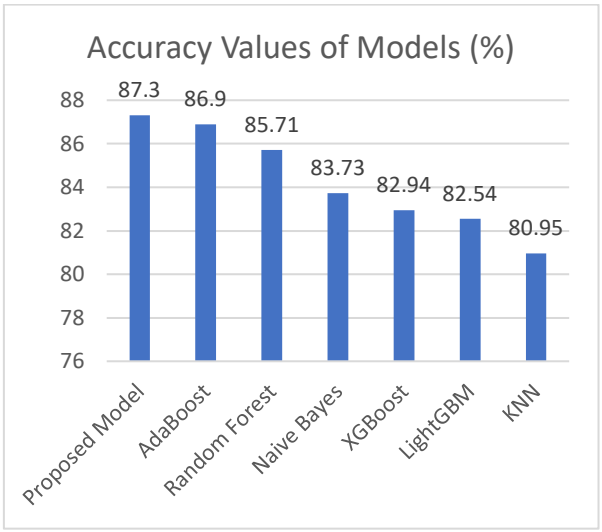

Figure 12. Accuracy values of models

4. Conclusions

Gliomas are tumors that form inside the brain or spinal cord. Grading these tumors is important to determine how fast the tumor is growing and how aggressive it is. Glioma grading plays an important role in determining the treatment approach for the tumor. While low-grade gliomas generally have a better prognosis, higher-grade gliomas can be more aggressive and difficult to treat. Automated glioma grading can be rapid and effective. This will allow the treatment process to start earlier. An accuracy value of 87.30% was achieved in the CNN and LSTM-based model we developed for automatic glioma grading. We believe that this value can be used in the Glioma grading of the proposed model.

Acknowledgments

The authors thank the owners of the database for sharing their data.

Declaration of Competing Interest

The authors declare that there is no conflict of interest in the study.

Author Contribution

The authors contributed equally to the article.

References

[1] **Hakyemez,** B., Erdogan, C., Ercan, I., Ergin, N., Uysal, S., & Atahan, S. (2005). High-grade and low-grade gliomas: differentiation by using perfusion MR imaging. *Clinical radiology*, *60*(4), 493-502.

[2] **Nandihal,** P., Shetty, V., Guha, T., & Pareek, P. K. (2022, October). Glioma Detection using Improved Artificial Neural Network in MRI Images. In *2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon)* (pp. 1-9). IEEE.

[3] **Gore**, S., Chougule, T., Jagtap, J., Saini, J., & Ingalhalikar, M. (2021). A review of radiomics and deep predictive modeling in glioma characterization. *Academic Radiology*, *28*(11), 1599-1621.

[4] **Saini**, A., Kumar, M., Bhatt, S., Saini, V., & Malik, A. (2020). Cancer causes and treatments. *International*

Journal of Pharmaceutical Sciences and Research, *11*(7), 3121-3134.

[5] **Boele,** F. W., Butler, S., Nicklin, E., Bulbeck, H., Pointon, L., Short, S. C., & Murray, L. (2023). Communication in the context of glioblastoma treatment: A qualitative study of what matters most to patients, caregivers and health care professionals. *Palliative Medicine*, *37*(6), 834-843.

[6] **Tasci,** E., et al., Hierarchical Voting-Based Feature Selection and Ensemble Learning Model Scheme for Glioma Grading with Clinical and Molecular Characteristics. International Journal of Molecular Sciences, 2022. **23**(22): p. 14155.

[7] **Cengil**, E., Eroğlu, Y., Çınar, A., & Yıldırım, M. Detection and Localization of Glioma and Meningioma Tumors in Brain MR Images using Deep Learning. *Sakarya University Journal of Science*, *27*(3), 550-563.

[8] **Yang,** Y., Yan, L. F., Zhang, X., Han, Y., Nan, H. Y., Hu, Y. C., ... & Wang, W. (2018). Glioma grading on conventional MR images: a deep learning study with transfer learning. *Frontiers in neuroscience*, *12*, 804.

[9] **Xiao**, T., Hua, W., Li, C., & Wang, S. (2019, August). Glioma grading prediction by exploring radiomics and deep learning features. In *Proceedings of the Third International Symposium on Image Computing and Digital Medicine* (pp. 208-213).

[10] **Vafaeikia,** P., Wagner, M. W., Hawkins, C., Tabori, U., Ertl-Wagner, B. B., & Khalvati, F. (2023). MRI-Based End-To-End Pediatric Low-Grade Glioma Segmentation and Classification. *Canadian Association of Radiologists Journal*, 08465371231184780.

[11] **Hastie,** T., Rosset, S., Zhu, J., & Zou, H. (2009). Multi-class adaboost. *Statistics and its Interface*, *2*(3), 349-360.

[12] **Turchetti Maia**, T., Pádua Braga, A., & de Carvalho, A. F. (2008). Hybrid classification algorithms based on boosting and support vector machines. *Kybernetes*, *37*(9/10), 1469-1491.

[13] **Qi**, Y. (2012). Random forest for bioinformatics. Ensemble machine learning: methods and applications.

[14] **Rish**, I. (2001, August). An empirical study of the naive Bayes classifier. In *IJCAI 2001 workshop on empirical methods in artificial intelligence* (Vol. 3, No. 22, pp. 41-46).

[15] **Dong**, J., Chen, Y., Yao, B., Zhang, X., & Zeng, N. (2022). A neural network boosting regression model based on XGBoost. *Applied Soft Computing*, *125*, 109067.

[16] **Ke,** G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T. Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, *30*.

[17] **Zhang,** M. L., & Zhou, Z. H. (2007). ML-KNN: A lazy learning approach to multi-label learning. *Pattern recognition*, *40*(7), 2038-2048.

[18] **Yildirim,** M. (2023). Image Visualization and Classification Using Hydatid Cyst Images with an Explainable Hybrid Model. *Applied Sciences*, *13*(17), 9926.

[19] *Url: https://archive.ics.uci.edu/dataset/759/glioma+grading+clinical+and+mutation+features+dataset*.

[20] **Cengil,** E., Çınar, A., & Yıldırım, M. (2022). A hybrid approach for efficient multi-classification of white blood cells based on transfer learning techniques and traditional machine learning methods. *Concurrency and Computation: Practice and Experience*, *34*(6), e6756.

[21] **Özbay,** E., & Özbay, F. A. (2023). Interpretable features fusion with precision MRI images deep hashing for brain tumor detection. *Computer Methods and Programs in Biomedicine*, *231*, 107387.

[22] **Özbay**, E., Çinar, A., & Özbay, F. A. (2021). 3D Human Activity Classification with 3D Zernike Moment Based Convolutional, LSTM-Deep Neural Networks. Traitement du Signal, 38(2), 269-280.

[23] **Yücel,** N., Yildirim, M., & Aslan, S. (2023). Performances of Pre-Trained Models in Classification of Body Cavity Fluid Cytology Images.