ADVANCED INTELLIGENT SYSTEMS
Open Access
www.advintellsyst.com

A Novel Mixed Convolution Transformer Model for the Fast and Accurate Diagnosis of Glioma Subtypes

*S. M. Nuruzzaman Nobel, S. M. Masfequier Rahman Swapno, Md Babul Islam, AKM Azad, Salem A. Alyami, Md Alamin, Pietro Liò, and Mohammad Ali Moni**

Glioblastoma is the most common adult brain tumor, significantly impacts disability and mortality. Early and accurate diagnosis of glioma subtypes is essential, but manual categorization is challenging due to their complexity, prompting the need for automated solutions. We developed an innovative mixed convolution-transformer model to classify glioma subtypes, including astrocytoma, glioblastoma, oligodendroglioma, and normal brain tissue, using whole slide images. The novelty of this model lies in its remarkable efficiency and precise results. Multiple advanced and complex layers are incorporated during its development to enhance its performance, ensuring that it delivers fast and accurate classification results for glioma. This proposed model obtains an overall training accuracy of 97.41%, peaking at 98.12% for validation and 97.35% for testing. Next, our model architecture is independently evaluated by comparing its training performances on the CIFAR-10 and CIFAR-100 datasets with the vision transformer and compact convolutional transformer models. Results across various datasets demonstrate that the model consistently outperforms existing models. This performance underscores the effectiveness of our proposed approach in classifying glioma subtypes accurately and efficiently, highlighting its potential impact on healthcare and disability. This system enhances the classification of glioma subtypes and facilitates swift identification, ensuring appropriate and timely treatment.

1. Introduction

Brain tumors arising from glial cells are malignant growths in the central nervous system (CNS).[1] Before 2016, the World Health Organization (WHO) classified CNS gliomas into grades I–IV according to the histological characteristics of a diverse group of tumors. A particular subset of brain malignancies generated from glial cells is referred to as gliomas.[2] The central[3] and peripheral nervous systems' nerve microenvironment is partly maintained by glial cells, including oligodendrocytes, astrocytes, ependymal cells, and microglia. Owing to the complex biological makeup of glial cells, tumors originating from this subset of nervous system cells known as gliomas can be further classified into several functional subgroups, each of which may have originated from a distinct functional subgroup.[4] Gliomas are the most prevalent primary CNS tumors in the brain, accounting for 80% of brain tumors that are malignant.[5,6] The WHO has classified tumors into grades 2–4 based on their aggressiveness, with higher grades indicating more malignant tumors. Commonly, grade 2 gliomas are commonly known as low-grade

S. M. N. Nobel, S. M. M. R. Swapno
Department of Computer Science and Engineering
Bangladesh University of Business and Technology
Dhaka 1216, Bangladesh

M. B. Islam
Department of Computer, Modeling, Electronic, and System Engineering
University of Calabria
87036 Rende, Italy

A. Azad, S. A. Alyami
Department of Mathematics and Statistics
Faculty of Science
Imam Mohammad Ibn Saud Islamic University (IMSIU)
Riyadh 13318, Saudi Arabia

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/aisy.202400566.

© 2024 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/aisy.202400566

A. Azad, S. A. Alyami
Research Department
King Salman Center for Disability Research
Riyadh 11614, Saudi Arabia

M. Alamin
Department of Mathematics & Physics
School of Engineering & Physical Sciences
North South University
Bashundhara, Dhaka 1229, Bangladesh

P. Liò
Department of Computer Science and Technology
The University of Cambridge
Cambridge CB2 1TN, UK

M. A. Moni
Artificial Intelligence & Digital Health Technology
Artificial Intelligence and Cyber Futures Institute, & Rural Health Research Institute
Charles Sturt University
346 Leeds Parade, Orange NSW 2800, Australia
E-mail: m.moni@uq.edu.au

gliomas (LGG), whereas grade 3–4 gliomas are referred to as high-grade gliomas (HGG).[7] The stereotactic[8] brain biopsy is the established approach used to diagnose brain cancers. However, it is intrusive and carries the potential for inconsistencies between different readers and inaccuracies in the sampling process.

Early diagnosis of glioma[9] is more essential, as glioma subtypes represent some of the most dangerous forms of brain tumors. Identifying the molecular patterns of these subtypes is essential for accurate diagnosis, prognosis, and treatment planning. While initial human classification is unnecessary, it is often inaccurate and time-consuming. As a result, automated classification systems have proven to be more successful in diagnosing glioma subtypes, offering greater accuracy, efficiency, and faster.

Several authors have focused on classifying glioma subtypes, employing machine learning and deep learning[10–12] algorithms to achieve this task. Their approaches highlight various glioma subtypes and utilize different datasets for classification. The paper Martin et al.[13] focuses on analyze low grade Gliomas, in particular most adult grade 2 and 3 gliomas had mutations in Isocitrate Dehydrogenase (IDH) 1 and 2, which was indicative of therapy of distinct glioma subtypes. The prognostic difference between grade 2 and grade 3 was lessened when 1p/19q codeletion was used for categorization. As a result, low-grade glioma formerly used to describe gliomas of grade 2 was substituted with the term lower-grade glioma, which now includes grade 2 and grade 3 tumors. Recent advancements and future directions for the surgical, radiotherapeutic, chemotherapeutic, and long-term management of adult lower-grade gliomas were examined to improve knowledge and treatment approaches. Fatemeh et al.[14] examined three adult-type diffuse glioma subtypes microvasculature without enhancement using vascular size imaging (VSI) magnetic resonance imaging (MRI), Oligo, Astro, and GBM comprised 38% of nonenhancing gliomas. VSI mapping assesses each individual's cerebral blood volume (CBV), microvasculature (μCBV), and vascular size. The histopathology and microvasculature of 21 patient tissue samples were examined. Astro showed lower CBV and μCBV than Oligo

(IDH-mut&1p/19q-codeleted) ($p = 0.004$ and 0.001, respectively) and larger vessels than GBM (IDH-wt) ($p = 0.01$). The histology study revealed that GBM (IDH-wt) exhibited more irregular vessel morphologies than the other two subtypes ($p < 0.05$). VSI helped writers understand the microvasculature of the three adult-type glioma subtypes without augmentation.

Benny et al.[15] assessed WHO-classified adult metabolic markers for the classification of glioma subtypes. The method cross-platform global metabolomic profiling with clinical, genetic, and pathological characterization of 224 glioma tumor —oligodendroglioma ($n = 31$), astrocytoma ($n = 31$), and glioblastoma ($n = 162$). All six glioma subtypes had different metabolic properties, and IDH-mutated 2-hydroxy glutaric acid-expressing subtypes were readily distinguishable from the types. In addition to the altered IDH pathway, IDH mutants had high glycerophosphates, inositols, monosaccharides, and sugar alcohols and low levels of sphingosine and lysoglycero. Grade 4 gliomas had significant glycine and 2-aminoadipic acid, while low-grade astrocytomas and oligodendrogliomas had N-acetyl aspartic acid. The authors Wei et al.[16] supervised deep learning for glioma subtype categorization. This hybrid convolutional neural network (CNN) based glioma subtype classification approach utilizes WSI and multiparametric magnetic resonance imaging. The authors created WSI and mpMRI methodologies. The WSI-based 2D CNN classified glioma subtypes. They avoided label restriction using poorly supervised sample patch extraction for glioma subtype classification. Utilizing mpMRIs, they developed a 3D CNN. Brain tumors were segregated by mpMRI. Finally, using a confidence measure, they combined WSI- and mpMRI-based data to improve prediction.

Finding Bounding box areas versus annotated tumor, Muhaddisa et al.[17] trained a deep learning classifier to identify glioma subtypes using tumor ROIs defined by elliptical bounding boxes versus manually annotated data. US and TCGA multimodality MRI datasets were used for experiments, with the US dataset only encompassing diffuse low-grade gliomas. The US dataset predicted 1p/19q codeletion status at 69.86% and the TCGA dataset predicted IDH mutation/wild type at 79.50%. The average deterioration using annotated GT tumor data for training was 3.0% (2.92% for 1p/19q codeletion status and 3.23% for IDH genotype). Sihan et al.[18] focused TCGA and GEO data to classify glioma subtypes (GSE4412). They searched GSEA for tumor microenvironment-related genes. Through gene set variation analysis (GSVA) enrichment, cancer subtypes were classified. GSVA enrichment analysis found three glioma subtypes. Additional survival prognosis and biological function research identified 13 tumor microenvironment gene sets and 14 core genes that impacted patient survival. Targeted medicines and illness detection might target these genes. The main problem is the inadequate capacity to accurately classify glioma subtypes.

The motivation for this research is the possibility of transforming present diagnostic processes by incorporating sophisticated deep-learning algorithms into medical systems. Implementing automated classification methods can significantly increase the efficiency of glioma diagnosis, reduce the work of healthcare employees, and ultimately improve patient care and outcomes. Glioblastoma is the most common primary brain tumor and is a significant cause of death in individuals with cancer. Fast and precise identification of several forms of glioma is crucial for successful therapy and enhanced patient results. The current diagnostic techniques need help due to the complex nature of glioma subtyping, often depending on human categorization, which is time-consuming and susceptible to mistakes. As a result, there is a strong need for automated classification systems that can provide improved accuracy, efficiency, and time-saving advantages. The advancement and utilization of deep learning models provide a hopeful approach to fulfilling this requirement.

To enhance healthcare through improved automatic classification of glioma subtypes, we focus on specific benchmark methods to ensure successful classification. This study involved the identification of many subtypes of glioma, such as astrocytes, glioblastomas, oligodendrogliomas, and normal brain tissue. These categories are very high risk and require fast categorization to guarantee the patient's well-being. **Figure 1** depicts examples of various glioma variants we classify in this study. By applying advanced deep learning and deep learning techniques, we aim to achieve precise glioma subtype classification, important for patient diagnosis and treatment planning. Our approach implements a highly advanced and customized layer-based novel model, mixed convolution transformer (MCT), designed for efficient, fast, and accurate classification of glioma subtypes.

© 2024 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH

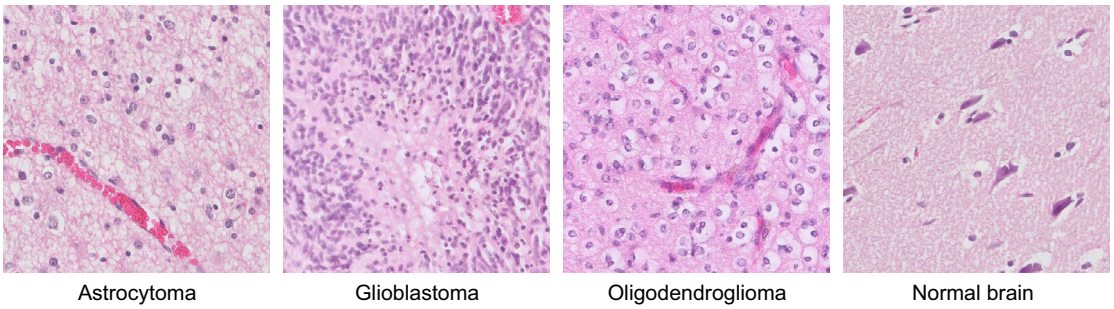


Figure 1. The figure provides an example of a representative sample of four different groups of images from the dataset.

This innovative model leverages cutting-edge technology to provide robust performance, significantly improving the accuracy and speed of medical image analysis in glioma subtype identification.

2. Methodology

This study focuses on categorizing the glioma subtype across various stages, presenting a distinctive MCT model. To assess the robustness of the MCT model, we conducted experiments on two distinct datasets, namely CIFAR-10 and CIFAR-100. **Figure 2** illustrates the comprehensive categorization criteria using our proposed MCT model alongside the evaluation criteria for other models on different datasets. This figure showcases the various datasets utilized within our system, detailing the preprocessing tasks undertaken and the different models implemented to achieve our results. This framework thoroughly compares the MCT model's performance against alternative models across

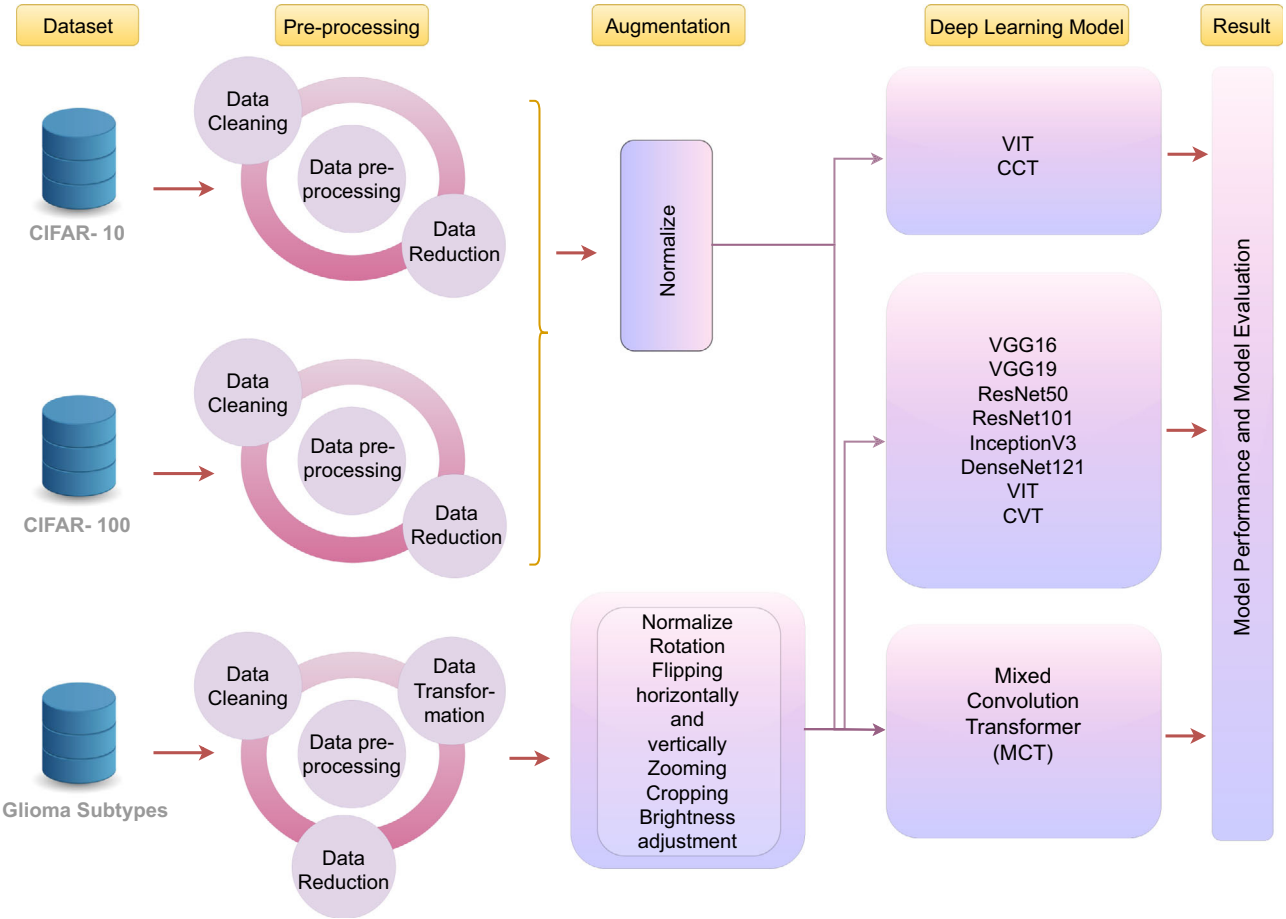


Figure 2. The schematic overview of the Glioma subtype classification workflow. The data handling process is shown in the diagram, which starts with data collection and continues with preprocessing and augmentation. Specifically designed models for the classification of glioma subtypes, such as MCT and transfer learning, are then applied to the data.

© 2024 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH

varied datasets, enlightening on its effectiveness and versatility in glioma subtype categorization.

2.1. Data Collection

In this study, we have collected an image dataset (sized of 19.2 terabytes) from the Zenodo repository[19] that represents the core characteristics of 28 different instances of adult-type diffuse glioma. The cases were carefully collected at the National Center of Pathology, housed within the Luxembourg National Health Laboratory (Laboratoire national de santé), between 2017 and 2021. The resource mainly contains WSI, which is highly judged for its detailed depiction of internal subtleties. These images were taken with a Philips IntelliSite Ultra Fast digital slide scanner. They provide a high level of detail with a 20x/0.75 NA Plan Apo objective lens, resulting in an average slide resolution of 0.25 µm pixel^{-1}. Referred to as the 20x magnification data, this extensive compilation contains 42 718 images, each showcasing the distinct characteristics of various glioma subtypes. These images have great meaning and may be divided into five categories: astrocytoma, glioblastoma, oligodendroglioma, normal brain, and necrosis. Each category expresses its own nature, representing the core of the glioma spectrum. The subtypes of this dataset and the quantity of data used for our model's train and test are displayed in **Table 1**.

2.2. Data Preprocessing

In this study, we used advanced image augmentation techniques to solve dataset overfitting. These guide us toward a balanced and robust dataset. For preprocessing, we used several techniques that help effective training. First, the data cleaning step is employed to remove errors, inconsistencies, and irrelevant data, ensuring the accuracy and quality of the dataset. To balance our data, we applied a data reduction technique, ensuring that the class with an excess of data was reduced to align with the other classes. This approach allowed us to achieve a balanced dataset across all classes. Next, we utilized an effective preprocessing technique known as augmentation. Augmentation is used in data preprocessing to artificially expand the size and diversity of a dataset, improving the model's generalization through the input data. This helps to prevent overfitting and enhances the robustness of the model. During the augmentation step, we utilized various techniques such as normalization, rotation, flipping (both horizontally and vertically), zooming, cropping, and brightness adjustment. Before augmentation, the distribution of our training data was as follows: necrosis class accounted for 1.0%, astrocytoma class accounted for 9.7%, glioblastoma accounted for 4.2%, oligodendroglioma accounted for 8.8%, and normal accounted for 75.2%. Before augmentation, the test data consisted of the following proportions: 2.2% necrosis class, 11.1% astrocytoma class, 5.8% glioblastoma, 10.3% oligodendroglioma, and 70.6% normal. Due to the poor performance of our model on this dataset, we have decided to employ the augmentation strategy. Following augmentation, our dataset has been expanded. After augmentation, the training data distribution for each class is as follows: astrocytoma, 24.1%; glioblastoma, 19.2%; oligodendroglioma, 25.3%; and normal, 31.4%. We excluded the necrosis class from the analysis due to its limited data and its behavior as an outlier. After augmentation, the test data for the astrocytoma class is 25.0%. The test data for glioblastoma is also 25.0%, while the test data for oligodendroglioma and normal cases is 25.0%. This data is highly beneficial for model use. **Figure 3** displays the graphical depiction of data distribution before and after augmentation. Each of our creative methods is carefully intended to balance and diversify our repository. These methods include geometric and geographical adjustments to arrange the dataset. Flipping images—horizontally or vertically—provide a larger view of our data. We've also welcomed random cropping, rotation augmentation, and modest but powerful image translation. Each move adds depth and complexity to our dataset, like a brushstroke. Strategic zooming enhances our dataset's depth, revealing delicate information that might otherwise be lost. Our technique relies on attentively augmenting training data to broaden our dataset. This ambitious expansion protects our models from overfitting and allows them to generalize and understand the domain's many patterns and complexity. The training method data and the after-augmentation outcome are displayed in **Table 2**.

Table 1. Subtype-wise training and testing image distributions within the dataset.

Class	Train data	Test data
Astrocytoma	3755	465
Glioblastoma	1633	241
Oligodendroglioma	3384	431
Normal brain	29 383	2947
Necrosis	389	90

We attentively split our dataset into an 80–20 split, allocating 80% for training and 20% for validation to create a balanced and dependable model. To balance the test data, we apply under-sampling, ensuring that each class has an equal representation. Specifically, we select 241 instances from each class. In this carefully created environment, we found an imbalance in the class Necrosis, which might cause overfitting. A deliberate choice was made to address this imbalance and traverse this difficulty. Thus, we remove Necrosis to focus on astrocytoma, glioblastoma, oligodendroglioma, and normal brain. This systematic protection conscious to join and present our dataset more effectively, reducing overfitting and improving model performance. To protect our test data, we used a thorough method. Precision under-sampling improved the quality and dependability of our test set. To strengthen our testing routine, a careful selection method was used to select the most representative and high-quality images from this modified test set of 964 images. **Table 3** displays the four classes—train, validation, and test—for which we have finally completed system development.

2.3. Mixed Convolution Transformer (MCT) Model Implementation

This article presented a novel deep learning model, namely the MCT. It is an innovative design that interacts with the Conv2D and Transformer portions. The Transformer is an encoder in this

© 2024 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH

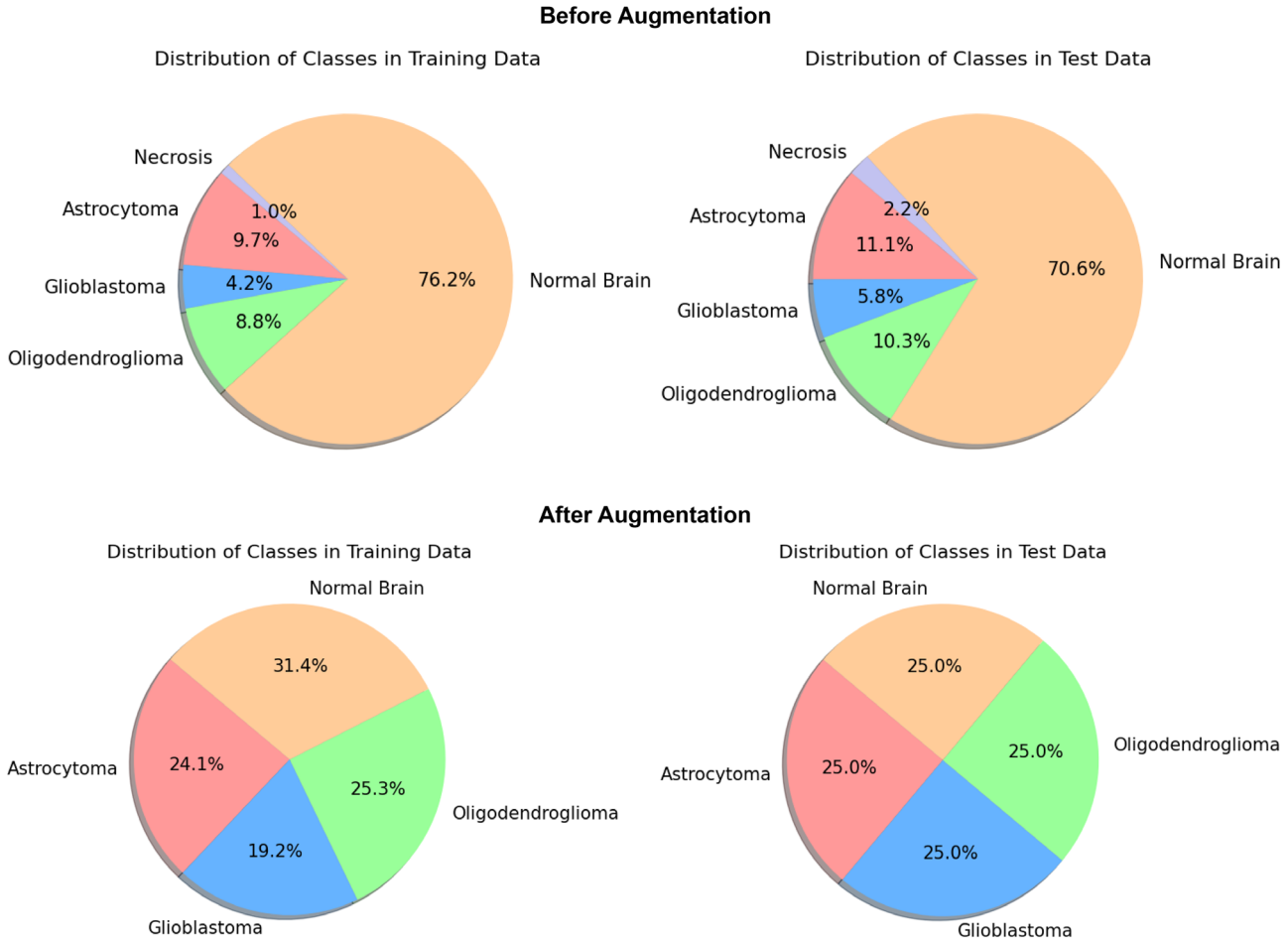


Figure 3. The effectiveness of augmentation strategies in expanding dataset proportions across multiple categories is demonstrated by a pie chart that compares the percentage distribution of data before and after augmentation.

Table 2. Comprehensive table showing the distribution of original training images, enhanced training images, and the total number of training examples in the dataset per class.

Class	Train	Augmented data	Total train
Astrocytoma	3755	18 775	23 530
Glioblastoma	1633	16 330	17 963
Oligodendroglioma	3384	3384	20 304
Normal brain	29 383	0	29 383
Necrosis	389	0	0

Table 3. Tabular representation providing a thorough overview of the dataset, showing how images are distributed throughout training, testing, and validation sets.

Class	Train	Validation	Test
Astrocytoma	18 024	4506	241
Glioblastoma	14 371	3592	241
Oligodendroglioma	18 951	4737	241
Normal brain	23 507	5876	241

innovative model, which uses Conv2D for data training. Several parameters in the Conv2D and Transformer sections are essential to implementation and optimization. In the end, the classification of glioma subtypes is the function of this model. Our design, which captures the complex interaction between Conv2D and Transformer components, is depicted in **Figure 4** and describes the whole glioma subtype classification process.

2.3.1. Conv2D Layer Execution Process

Convolutional neural network models were used to train the wall quality dataset to identify the critical elements from the images. The Gaussian function used for feature extraction is shown in Equation (1). The Gaussian function's variance is indicated by the parameter *r*.

© 2024 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH

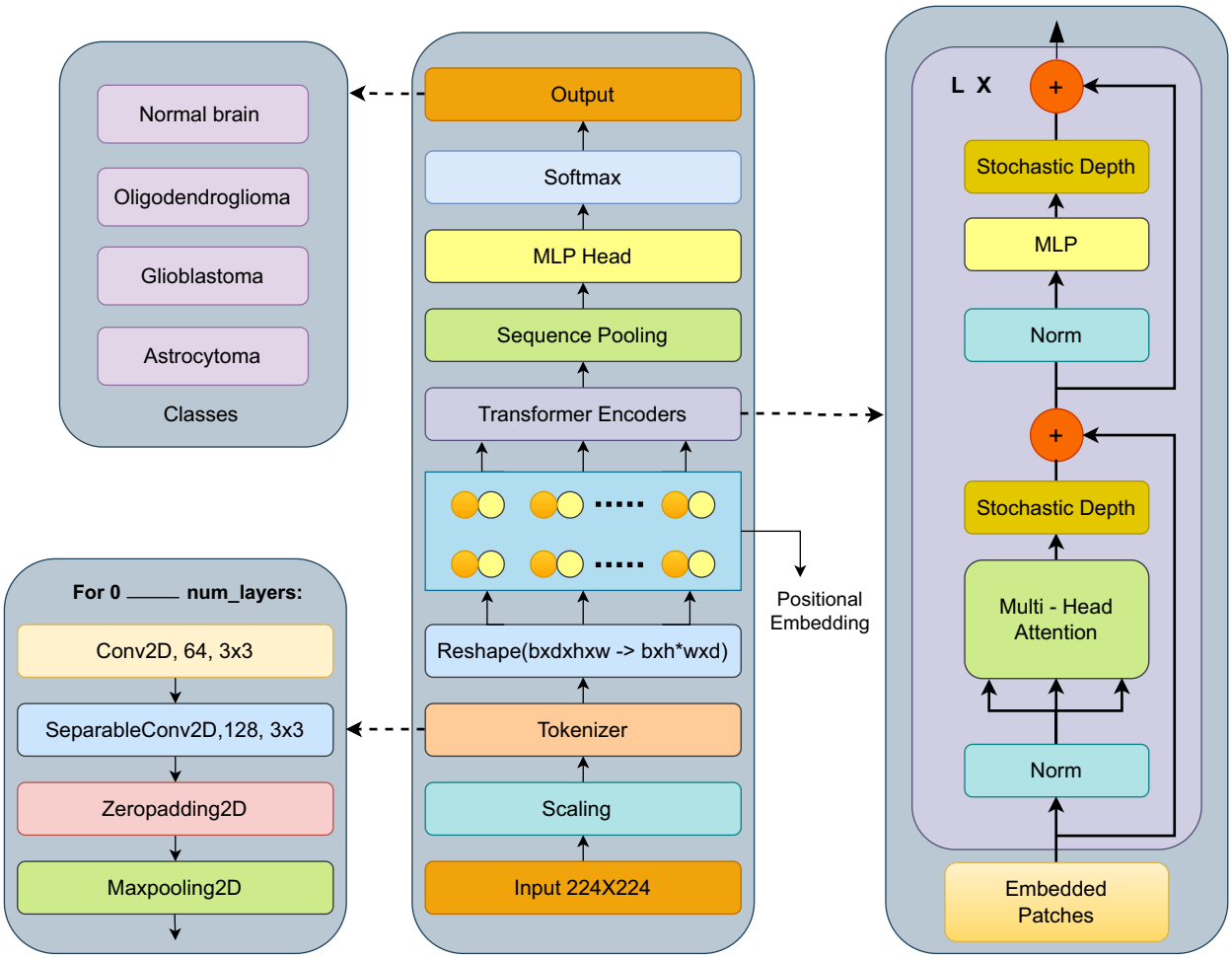


Figure 4. MCT, our novel proposed architecture, is conceptually visualized.

$$\text{Feature}\ (i,j,r) = \frac{1}{\sqrt{2\pi r}}\exp\left(\frac{i^2+j^2}{2r^2}\right) \tag{1}$$

Equation (2) illustrates the Gaussian orientation function utilized for image filtering. Here the features are declared as F.

$$O\,(i,j,r,\theta) = F_{ii}\cos^2(\theta) + F_{ii}\cos(\theta) + F_{jj}\sin^2(\theta) \tag{2}$$

Consequently, shown in Equation (3)–(5), Feature_{ii}, Feature_{ij}, Feature_{jj} stands for the Gaussian function's second derivatives.

$$\text{Feature}\ (i,j,r) = \frac{(i^2-r^2)\exp\frac{\left(-\left(\widetilde{i^2+j^2}\right)\right)}{\sqrt{2\pi^2}}}{\sqrt{2\pi r}} \tag{3}$$

$$\text{Feature}\ (i,j,r) = \frac{(j^2-r^2)\exp\frac{\left(-\left(\widetilde{i^2+j^2}\right)\right)}{\sqrt{2\pi^2}}}{\sqrt{2\pi r}} \tag{4}$$

$$\text{Feature}\ (i,j,r) = \frac{ij\exp\frac{\left(-\left(\widetilde{i^2+j^2}\right)\right)}{\sqrt{2\pi^2}}}{\sqrt{2\pi r}} \tag{5}$$

The final output of the wall quality dataset was produced by applying four-layer Conv2D layers to the input image, represented by Equation (6)–(8). These layers included convolution filtering, sigmoid filtering, linear transformation, and linear sigmoid.

$$Z = \text{Feature}_{ij} \times \text{Filter} \tag{6}$$

$$\text{Con}\,[R,C] = \text{Feature}_{ij} \times \text{Filter}\,[R,C] \tag{7}$$

$$\text{Con}\,[R,C] = \sum j,k\,\text{Kernel}\,[j,k]\,\text{Feature}\,[R-j][C-k] \tag{8}$$

where the rows and columns of the input image matrix are denoted by R, C. Apply the sigmoid function as demonstrated in (9) and (10) to the preceding problem.

$$\text{ASig} = \text{Sigmoid}\,(Z) \tag{9}$$

© 2024 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH

$$\text{Sigmoid}\,(Z) = \frac{1}{1+e^{-z}} \tag{10}$$

In order to process the CNN's third layer, the linear transformation was performed to the layer above, as shown in Equation (11).

$$\text{ZLinear} = \text{Weight}_{\text{T}} \times \text{ASig} + \text{Bias} \tag{11}$$

2.3.2. Transformer Layer Execution Process

The transformer portion functions as an encoder, employing various techniques before the multilayer perceptron (MLP) locates the classification process component. The model can attend to input from many representation subspaces at different locations at the same time thanks to multiheaded attention.

$$\text{MultiHead}\,(Q, K, V) = \text{Concat}(\text{head}_1, \ldots., \text{head}_h) \times \omega_0 \tag{12}$$

$$\text{head}_i = \text{Attention}\left(QW_i^{\text{Q}}, KW_i^{\text{K}}, VW_i^{\text{V}}\right) \tag{13}$$

Another complex execution of the transformer is MLP; the MLP is a feed-forward neural network with dense and dropout layers. These MLP blocks get identical stacked layer blocks and structures. For example, consider X as the token characteristics with n sequence length and d dimension. A mathematical definition of each block is as follows.

$$Z = \sigma(XU), \tilde{Z} = s(Z), Y = \tilde{Z}V \tag{14}$$

$$z_0 = \left[x_{\text{class}}; x_{\text{p}}^1 E; x_{\text{p}}^2 E, .., x_{\text{p}}^n E\right] + E_{\text{pos}}, E\varepsilon R^{(N+1)*D} \tag{15}$$

$$z_l^1 = \text{MSA}(LNz_{l-1}) + z_{l-1}, l = 1, \ldots.., L \tag{16}$$

$$z_l = \text{MLP}\left(LNz^l l\right), z_1^l, l = 1, \ldots..L \tag{17}$$

$$Z = LN\left(z_l^0\right) \tag{18}$$

σ is the activation function, U and V represent the channel's linear projection dimensions, and s(.) is the identity mapping. According to Equation (14), the layer captures spatial interaction, but individual tokens are calculated without interaction. In Equation (15) and (16), the layers' class token, learnable embedding placement, and patch embedding are detailed before stacking in Equation (18). The final output of the encoder is shown in Equation (19).

$$S(Z) = Z \odot f_{\text{w,b}}(Z) \tag{19}$$

where the dot product or element-wise multiplication is represented by ⊙. By combining the Separable Conv2D with Conv2D, we created a unique layer. The layer's functionality is completed when these elements come together. The separable convolution's output O has the following mathematical representation.

$$O = I \times K_{\text{r}} \times K_{\text{c}} \tag{20}$$

Our MCT model (Figure 4) illustrates a neural network architecture designed for glioma classification, featuring a hybrid approach combining convolutional layers and transformer encoders. The input is an image resized to 224×224 pixels, which is first processed through a series of convolutional layers, including standard Conv2D and SeparableConv2D, followed by max pooling and zero padding for feature extraction. These features are then tokenized, reshaped, and passed through multiple transformer encoders equipped with multihead self-attention mechanisms to capture global relationships within the image. Positional embeddings are added to retain spatial information. The output from the transformer encoders undergoes sequence pooling and is processed by an MLP head, followed by a softmax function to generate class probabilities. Additionally, the architecture incorporates residual connections and stochastic depth, providing regularization and improving the training process by allowing specific layers to be bypassed or randomly skipped, enhancing the model's ability to generalize well.

The implementation above represents the execution of our MCT model. The functioning of our constructed MCT model is explained in a step-by-step manner. Therefore, our advanced model effectively functions to categorize different subtypes of glioma.

2.4. Hyperparameter Tuning of MCT Model

We have implemented a list of the hyperparameter tuning processes for our MCT model. Hyperparameter tuning refers to selecting the optimal set of hyperparameters for a deep learning model. Hyperparameters are configuration settings used to control the learning process and the structure of the model, which are set before the training process begins. Unlike the model parameters (e.g., weights in a neural network), they are not learnt from the data but can significantly impact the model's performance. Hyperparameter tuning is needed because it helps optimize a machine learning model's performance by finding the best configuration of hyperparameters. Hyperparameters, set before the training process, significantly affect how well a model learns from data and performs on unseen data. Proper tuning can lead to better accuracy, faster convergence, and improved generalization, ensuring the model fits the training data well and performs effectively on new, unseen data. With tuning, models may avoid overfitting, where they learn the training data too well but fail to generalize, or underfitting, where they need to learn the data better. Thus, hyperparameter tuning is crucial for achieving the best possible model performance. We have analyzed many factors, such as weight decay, kernel initiator, pool, padding, optimizer, learning rate, batch size, and attention. To perform hyperparameter tuning, start by defining the search space by selecting the hyperparameters and their possible values. Next, choose a strategy or algorithm. Split the dataset into training and validation sets and then train models using different combinations of hyperparameters. Each model's performance is evaluated using a chosen metric, and finally, the hyperparameters that yield the best validation performance are selected. Here, we identify a specific range of parameter values, constituting the search space for the most optimal results in classifying glioma subtypes. Then, we select a value from this range to implement our system. **Table 4** displays the outcomes of the hyperparameter tuning process for our MCT model.

© 2024 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH

Table 4. Several hyperparameters that were experimented with and highlight the combinations that yielded the optimal performance for our model.

Parameter	Search space	Selected value
Weight decay	[0.0001–0.001]	0.02
Pool	[Max, Average]	Max
Kernel initiator	[he, glorot, normal]	normal
Padding	[Valid, same]	same
Optimizer	[SGD, Adam, RMSprop, Nadam]	Adam
Learning rate	[0.01, 0.001, 0.0001, 0.02]	0.0001
Batch size	[8,16,32]	16
Attention	[Self, Multihead]	Multihead

By conducting methodical experiments, hyperparameter tuning optimizes the model's performance, improving its ability to accurately forecast outcomes and its resilience.

3. Result Analysis and Findings

This section presents the comprehensive results generated by our classification system using the MCT model. We evaluate the accuracy and loss of our system and evaluate the precision, recall, and F1 score to gauge the effectiveness of our model. Our precise evaluation process has led to the creation of a detailed diagram demonstrating the highly efficient subclassification of gliomas.

3.1. Performance Measuring Metrics Evaluation and Benefits

We evaluate our glioma subtype classification working model using a set of important metrics, namely, True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), to determine its accuracy, precision recall, and F1 score. When taken as a whole, these numbers show how well the model performs in classifying glioma subtypes classification. The MCT model has demonstrated outstanding performance, with a training accuracy of 97.12% and an even more remarkable validation accuracy of 98.12%. Training and validation accuracy have different but complementary roles in evaluating the effectiveness of deep learning models. Training accuracy judges the degree to which a model accurately represents the training data all over the learning process. On the other hand, validation accuracy evaluates the model's ability to generalize to new and unseen data. The training accuracy indicates the model's proficiency in recognizing the patterns in the training set. Additionally, the accuracy of validation assesses its capability to make precise predictions on data not used for training. It is essential to monitor both metrics to comprehend a model's performance and guarantee its efficacy in practical situations. The exceptional statistics demonstrate the model's strength and expertise in acquiring and applying patterns from the dataset. **Figure 5**A displays the

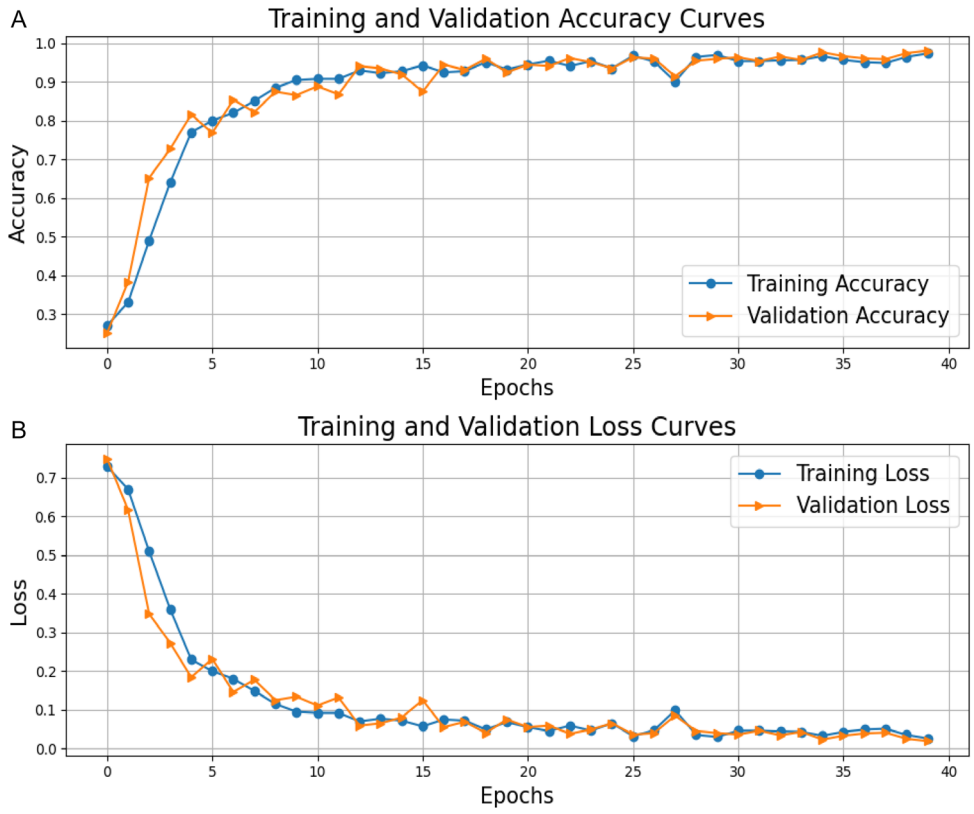


Figure 5. A) Model's training and validation accuracy and B) Model's training and validation loss, to demonstrate its efficacy.

© 2024 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH

graphical depiction of training accuracy and validation accuracy as a function of the number of epochs.

We conducted a comprehensive study that goes beyond measuring accuracy, including examining loss measures in both the training and validation phases. Notably, our approach has surprisingly few losses, which indicate the model's effectiveness and accuracy in learning from the data. The advantage of finding training loss and accuracy is their capacity to offer immediate feedback on the learning progress of a deep learning model based on the data. By tracking the training loss, we can assess the convergence of the model during training, aiding in the adjustment of hyperparameters and the detection of potential problems such as overfitting and underfitting. Training accuracy measures a model's performance on the training dataset, indicating its capacity to learn from the given data. These indicators are essential diagnostic tools for optimizing model performance and directing the training process toward desired outcomes. Figure 5B provides a clear visual image of the train and validation loss. This image highlights the model's ability to reduce loss in training and validation, further confirming its efficiency and reliability in capturing complex data patterns.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{21}$$

Precision, recall, and F1 score are essential measures used to assess the effectiveness of deep learning models. Precision is a metric that evaluates the correctness of optimistic predictions by verifying that the recognized positives are genuinely positive. Recall measures the model's capacity to correctly identify all relevant instances, explicitly capturing the percentage of actual positives that are accurately predicted. The F1 score achieves a trade-off between precision and recall by calculating their harmonic mean. These metrics are crucial for comprehending the effectiveness of a model, particularly in tasks where accurately recognizing positive instances is of absolute importance in medical diagnosis sector. They assist in developing models to reach an optimal trade-off between false positives and false negatives, enhancing overall efficacy and dependability. **Table 5** illustrates the implemented and suggested models' precision, recall, and F1 scores. The formula is expressed below.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{22}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{23}$$

Table 5. Model performance metrics on the test set: a comprehensive overview of the evaluation measures and performance indicators.

Class	Precision	Recall	F1 Score
Astrocytoma	0.84	0.99	0.91
Glioblastoma	0.94	0.81	0.87
Oligodendroglioma	0.80	0.84	0.82
Normal brain	1.00	0.91	0.95

$$F1\,\text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{24}$$

The receiver operating characteristic (ROC) curve thoroughly assesses a binary classifier's performance by considering all potential classification levels. It demonstrates the balance between sensitivity and specificity, assisting in choosing the best threshold for a particular use case. The area under the curve (AUC)–ROC measures the classifier's capacity to differentiate across classes, regardless of the threshold. It is an invaluable tool for evaluating and choosing models based on their overall performance. The class of astrocytomas have an astounding AUC of 0.99, glioblastomas have an AUC of 0.98, oligodendrogliomas have an AUC of 0.96, and a normal brain has an AUC of 0.99. **Figure 6** displays the ROC curve for four classes.

A confusion matrix is an essential machine learning method for assessing a classification model's efficacy. It details the model's classifications and how well we match the real class labels. Usually, the matrix is set up as a table with rows and columns denoting the real classes and the categorized ones. In our investigation, we created a visual representation of the confusion matrix that displays four unique classes: astrocytoma, glioblastoma, oligodendroglioma, and normal brain. The matrix is visually depicted in **Figure 7**, effectively illustrating the categorization ranges for each class using yellow and green boxes. This image accurately represents the model's classification performance in these precise categories, offering a concise and enlightening perspective on its precision and accuracy in this multiclass classification assignment.

3.2. Testing of MCT Architecture with CIFAR-10 and CIFAR-100 Dataset

While assessing the performance of our MCT architecture on the CIFAR-10 and CIFAR-100 datasets, we observed interesting findings compared to the vision transformers (VIT) and compact convolutional transformers (CCT) models at various stages of training, as performed by Hassani et al.[20]

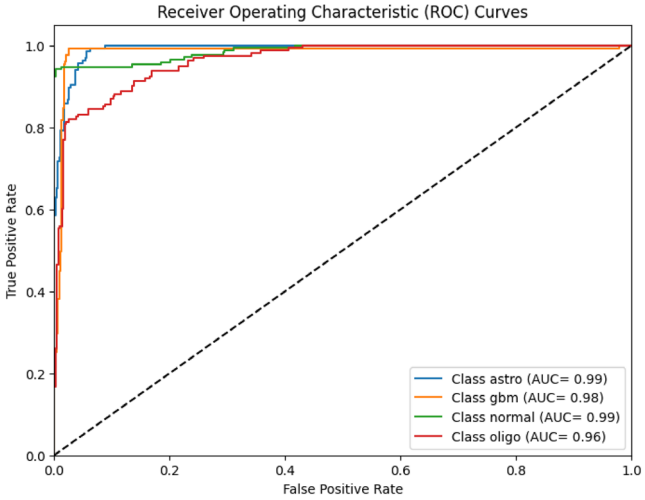


Figure 6. A figure showing ROC curves produced in a test set by the MCT model.

© 2024 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH

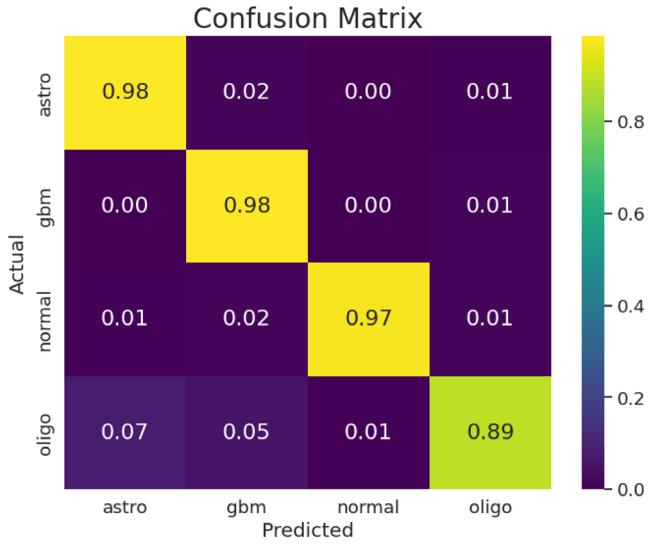


Figure 7. Showing the confusion matrix produced by our proposed model, MCT, which provides a thorough analysis of the classification results.

In CIFAR-10, the VIT models at 100th epoch the training accuracy of 94.36% and a validation accuracy of 80.83%, at 200th epoch the training accuracy of 95.47% and a validation accuracy of 80.99%, and at 300th epoch the training accuracy of 96.01% and a validation accuracy of 82.35%, consisting of 4.7 million parameters. By the CCT models 100th epoch the training accuracy of 86.96% and a validation accuracy of 79.40%, at 200th epoch the training accuracy of 88.68% and a validation accuracy of 83,08%, at 100th epoch the training accuracy of 89.25% and a validation accuracy of 83.10% utilizing 0.4 million parameters. Remarkably, our proposed MCT model, after 100th epochs, attained a training accuracy of 92.44% and a validation accuracy of 83.54%, at 200th epoch the training accuracy of 94.09% and a validation accuracy of 84.88%, at 300th epoch the training accuracy of 97.02% and a validation accuracy of 85.18%, employing 0.4 million parameters. Our model demonstrated exceptional accuracy while using the same settings, highlighting its ability to handle complex data formats effectively. To summarize, our MCT model not only showed vital accuracy in the fewer epochs but also showed significant improvement, outperforming current models in later stages. This indicates its potential for better learning and representation skills. **Table 6** displays the CIFAR-10 result.

During our assessment of the CIFAR-100 dataset, we performed comparative studies of the VIT, CCT, and MCT models. We ensured that all models kept the kernel sizes and convolutional parameters similar. When analyzing the CIFAR-100 dataset, the VIT models 100th epoch achieved a training accuracy of 41.78%, at 300th epoch the training accuracy of 52.42%, using 3×3 kernel. In the CCT model, after 100th epochs, the training accuracy showed a slight improvement to 69.97%, at the 300th epoch the training accuracy of 80.10%, using a 3×3 kernel and a convolution setting of 2. Remarkably, our suggested MCT model achieved superior performance compared to both predecessors, with 100th training accuracy of 82.10%, at the 300th training accuracy of 87.36%. This was accomplished using the identical convolution parameter of 2 and a 3×3 kernel. Our MCT model demonstrated superior performance in terms of accuracy compared to other models with identical convolutional and kernel configurations. The findings emphasize the exceptional performance and flexibility of our MCT model, especially when compared directly with the same convolutional settings. This highlights its potential to improve the learning and representation of complex features in the CIFAR-100 dataset. **Table 7** displays the CIFAR-100 result.

Table 6. On the CIFAR-10 dataset, the following table compares the number of epochs, top-1 training accuracy, and top-1 validation accuracy (Bold are indicating the results generated by proposed model).

Epoch	Model	Train accuracy	Validation accuracy	Parameter
100	VIT	94.36%	80.83%	4.7 M
200	VIT	95.47%	80.99%	4.7 M
300	VIT	96.01%	82.35%	4.7 M
100	CCT	86.96%	79.40%	0.4 M
200	CCT	88.68%	83.08%	0.4 M
300	CCT	89.25%	83.10%	0.4 M
100	**MCT**	**92.44%**	**83.54%**	**0.4** M
200	**MCT**	**94.09%**	**84.88%**	**0.4** M
300	**MCT**	**97.02%**	**85.18%**	**0.4** M

3.3. Improvement and Effectiveness of Our System

Our study has focused on categorizing glioma subtypes using a Zenodo dataset containing ≈42 718 WSI across five classes: astrocytoma, glioblastoma, oligodendroglioma, normal brain, and necrosis. Using a novel MCT model, which integrates complex conv2D layers and transformers, we achieved remarkable results with an overall accuracy of 97.41%. Specifically, our model achieved a validation accuracy of 98.12% and a testing accuracy of 97.35%, surpassing previous methods. Performance evaluations against VIT and CCT models on CIFAR-10 and CIFAR-100 datasets consistently demonstrated superior accuracy throughout various epochs. **Figure 8** displays the dataset used for training, validation, and testing in our system for classifying glioma subtypes. Additionally, we rigorously assessed precision, recall, and F1 scores for each class, providing insights into the

Table 7. On the CIFAR-100 dataset, the following table compares the number of epochs, top-1 training accuracy (Bold are indicating the results generated by proposed model).

Epoch	Model	Train accuracy	Conv	Kernel
100	VIT	41.78%	–	3×3
300	VIT	52.42%	–	3×3
100	CCT	69.97%	2	3×3
300	CCT	80.10%	2	3×3
100	**MCT**	**82.10%**	**2**	3×3
300	**MCT**	**87.36%**	**2**	3×3

© 2024 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH

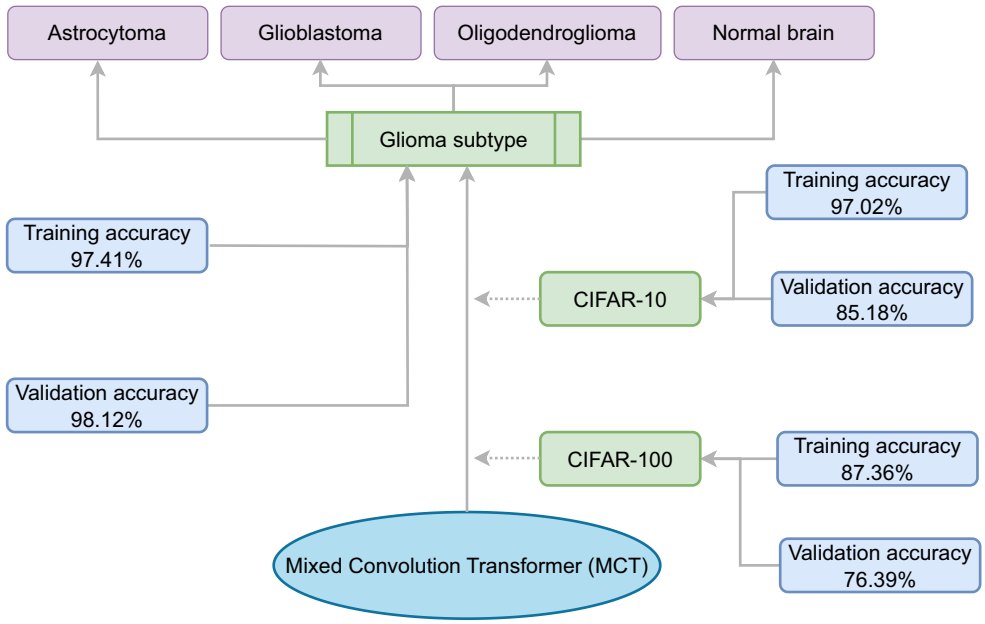


Figure 8. The key discoveries of our system include the classification type, training and validation accuracy, and the results obtained from diverse datasets for our model.

model's strengths and weaknesses. Furthermore, ROC curves illustrated the discriminatory capability of our classification system. The physical expressions of our system's capabilities are presented in **Figure 9**, demonstrating the remarkable alignment between actual and expected outcomes. Each of the four glioma subtypes is defined by its anticipated structure, demonstrating our process's strength and automated nature. In addition to categorization, our approach excels in speed and efficiency, guaranteeing rapid and precise identification of different forms of glioma. Our original glioma subtype classification method represents a significant leap forward in accurately categorizing these complex tumors. Unlike existing systems, which often suffer from limited data and less than precise classification, our study introduces a novel approach that revolutionizes the landscape of glioma classification. By leveraging a comprehensive dataset and employing cutting-edge techniques, our method enhances classification accuracy and offers a deeper understanding of glioma subtypes. This advancement holds profound implications for the field, paving the way for more tailored and effective treatment strategies. In essence, our research marks a pivotal moment in glioma classification, indicating a new era of precision medicine where diagnoses are more refined, diagnosis is more informed, and patient care is significantly improved.

3.4. Finding Robustness of Implementation Models and Comparison Showcase

The scope of our study includes a range of deep learning and deep learning models, which consist of established architectures such as VGG16, VGG19, ResNet50, ResNet101, ResNet201, DenseNet121, InceptionV3, as well as emerging paradigms such as VIT, convolutional vision transformer (CVT), and our original contribution, the MCT. **Table 8** shows these model's performance thoroughly. Our proposed MCT model has shown to be exceptionally successful, demonstrating unmatched accuracy in classifying gliomas. The MCT model has exceptional performance in reliably distinguishing glioma types, with a training accuracy of 97.41%, validation accuracy of 98.12%, and test accuracy of 97.35%, surpassing all other models. This accomplishment highlights the effectiveness and future possibilities of our innovative MCT architecture in applications related to medical image processing. Our model utilizes multichannel information processing to achieve high accuracy and exhibit robustness across various datasets and validation conditions. Our implemented models collectively demonstrate the progression and variety within the field of deep learning for medical imaging. Among these models, our MCT model stands out for its superior accuracy in classifying glioblastoma.

To improve glioma classification, we thoroughly examined recent studies in the field. We compared essential factors such as age, gender, input data, methodology, target conditions, and performance metrics such as AUC and accuracy. After conducting a comprehensive examination, it became evident that the existing models needed to be revised regarding the amount and quality of data they used, leading to suboptimal results. However, our pioneering method, MCT, represents a significant advancement. Utilizing a large dataset of 42 718 images, our model surpasses the constraints of previous models. Using broad and diverse datasets in glioma classification research establishes a new standard, giving our model exceptional depth and robustness. **Table 9** compares the latest research in glioma classification and the outcomes of our suggested model. In spite of that, the ultimate proof of MCT's supremacy can be seen in its performance measurements. Our model achieves exceptional AUC and accuracy rates, surpassing current approaches and establishing a new standard for accuracy in glioma classification.

© 2024 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH

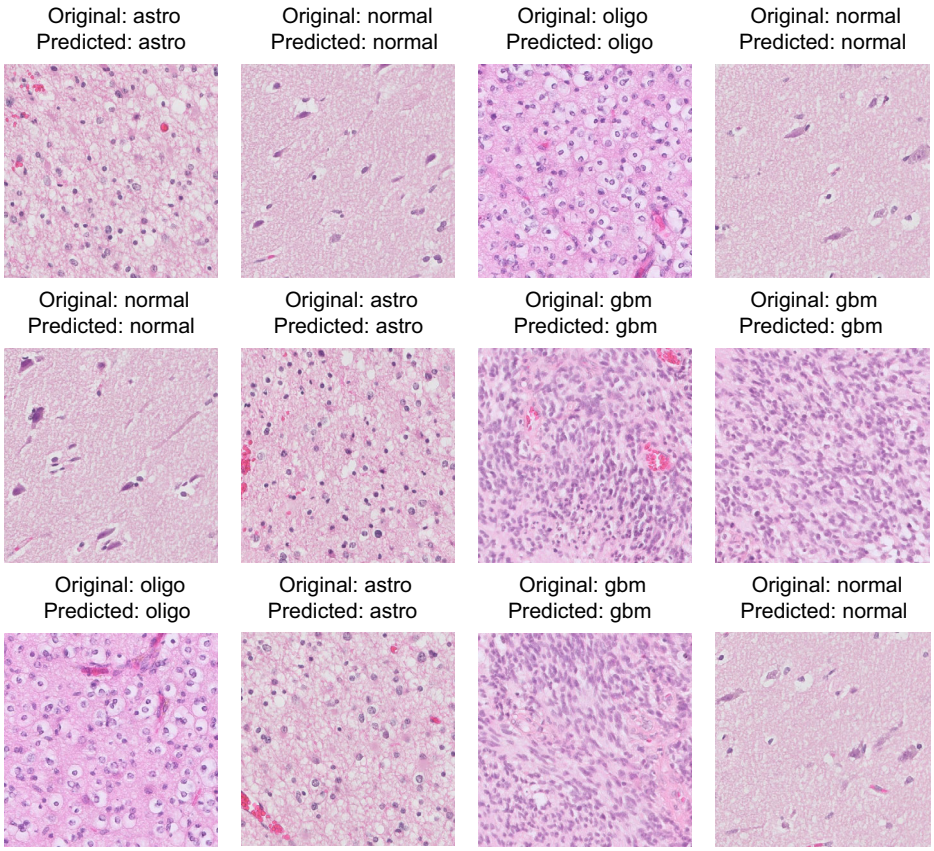


Figure 9. Predicted image produced by our MCT model, demonstrating the precise predictions produced by our proposed model.

Table 8. A thorough overview of model performance on training and validation sets, our proposed model using the glioma subtype dataset and other models. It displays precision and validation accuracy metrics (Bold are indicating the results generated by proposed model).

Model	Train accuracy	Validation accuracy	Test accuracy
VGG16	90.96%	93.43%	92.56%
VGG19	87.61%	80.02%	79.35%
ResNet50	95.45%	92.39%	92.43%
ResNet101	96.22%	92.87%	91.78%
ReseNet201	94.45%	96.61%	95.28%
DenseNet121	87.15%	86.92%	85.61%
InceptionV3	77.18%	75.81%	75.38%
VIT	96.39%	96.48%	95.76%
CVT	84.35%	72.26%	70.28%
MCT(Proposed)	**97.41%**	**98.12%**	**97.35%**

This achievement highlights the effectiveness and dependability of MCT as an advanced tool in detecting and categorizing gliomas. We have initiated a new period in the categorization of gliomas. Using MCT, we have exceeded the constraints of current models and achieved a new level of precision and effectiveness in this crucial field of medical research.

4. Discussion

Our research presents a highly effective glioma subtype classification system that stands out for its accuracy, efficiency, and robustness. We have focused on four subtypes: astrocytoma, glioblastoma, oligodendroglioma, and normal brain. The dataset consisted of high-resolution WSI images. Our system demonstrates exceptional precision in the field, achieving a remarkable training accuracy of 98.41%, an effective validation accuracy of 98.12%, and a testing accuracy of 97.35%. The outstanding outcomes highlight the effectiveness of our methodology in precisely categorizing the wide range of glioma subtypes. This significant advancement in glioma classification is pivotal for clinical and research applications, offering improvements that could directly benefit patient care and therapeutic decision-making. The innovative nature of our system lies in its ability to handle complex data and provide precise results, making it a valuable tool in the healthcare sector. The potential applications of our system extend beyond mere classification; it promises to aid in the diagnosis, treatment planning, and prognostic evaluation of glioma patients. By refining the classification process, our research contributes to the broader goal of advancing medical technology and improving outcomes for patients with gliomas.

By incorporating sophisticated classification criteria, our system surpasses new methodologies in the field. In this study, we demonstrate that our classification approach enhances the

© 2024 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH

Table 9. Comparative performance analysis of our proposed approach with alternative models that used MCT (Bold are indicating the results generated by proposed model).

Author and year	Age	Gender	Method	Target condition	AUC (±SD)	Test accuracy
Ahammed[21]	NR	NR	Deep CNN	Glioma grade	NR	94.64
Bangalore[22]	NR	NR	Dense-UNet	IDH genotype	0.98 ± 0.146	97.14
Batchala[23]	NR	50–52	Multivariate model	1p/19q integrity	NR	81.1
Cao[24]	NR	74–67	Elastic-Net SVM	LGG/HGG	0.915 ± 0.356	NR
Carver[25]	NR	NR	Elastic-Net RGL Model	IDH genotype	NR	74
Fan[26]	46.8	NR	Elastic-Net, RGL, SVM	glioblastoma, oligodendroglioma	0.923	93.8
Gates[27]	NR	NR	Random Forests	Glioma grade	NR	96
Jiang[28]	45.4 ± 13.1	43–44	Fusion Radiomics, SVM	MGMT	0.898 ± 0.323	88.6
Jiang[29]	NR	68–59	SVM, RF, LR	IDH genotype	0.747 ± 0.228	NR
Lee[30]	NR	47–41	KNN, SVM	IDH genotype	NR	83.4
Li[31]	60.0	37–32	SVM	PTEN genotype	0.787	82.5
Lo[32]	NR	28–11	LR	IDH genotype	NR	90
Matsui[33]	42	131–86	DLR	LGG	NR	58.5
Mzoughi[34]	NR	NR	LR, SVM	Glioma grade	NR	96.4
Park[35]	NR	NR	RF	IDH genotype	0.900 ± 0.298	NR
Shboul[36]	NR	NR	LR, MRF	IDH genotype	0.84 ± 0.156	NR
Sun[37]	NR	NR	LR	P53 statusNR	0.709	81.3
Tian[38]	NR	53–35	LR, RN	TERT promoter mutation status	0.889 ± 0.335	84.2
Zhang[39]	NR	61–47	CNN, SVM	LGG/HGG	0.93	94
Zhao[40]	45.0 ± 14.4	19–17	RF	Glioma grade II/III	0.861 ± 0.240	78.1
Proposed Model (2024)	**NR**	**NR**	**MCT**	**astrocytoma, glioblastoma, oligodendroglioma, normal brain**	**0.99, 0.98, 0.99, 96**	**97.35**

precision of glioma subtype identification and we present Table 9, a comprehensive comparison between existing research findings and our own results. It is noticeable from the data that our classification results significantly outperform those reported in previous studies. This highlights the efficacy and robustness of our approach, demonstrating its superiority in achieving higher accuracy and reliability in classification tasks.

To construct this system, we have created a novel MCT model designed explicitly for categorizing glioma subtypes. The model introduces several innovations that enhance its efficiency and performance. First, the MCT model strategy efficiently extracts features at different spatial resolutions, effectively capturing local and global data. Second, stochastic depth regularization improves training depth robustness by randomly dropping layers during training, preventing overfitting and enhancing generalization. Additionally, dynamic data augmentation enriches the training dataset by applying various transformations, enhancing the model's ability to generalize to unseen data. The model utilizes convolutional layers to extract features while preserving spatial information, diverging from traditional tokenization methods. This is complemented by optional positional embeddings that enable the model to understand spatial relationships within the image. Stochastic depth regularization is incorporated to prevent overfitting and enhance generalization, while a comprehensive data augmentation pipeline ensures robustness to variations in input data. The model captures long-range dependencies and hierarchies in image features using transformer blocks, with a sequence pooling mechanism facilitating information aggregation across patches. Additionally, the model uses multihead self-attention transformer blocks to capture complex dependencies across tokens. Sequence pooling is also utilized to aggregate features effectively. The model's attention mechanism allows it to dynamically adjust focus on relevant regions, contributing to its effectiveness in various image-related tasks. While several authors have employed CNN-transformer architectures, our development of the MCT model introduces a novel approach by incorporating two distinct types of convolutional layers and transformers. We used the mixing of Conv2D and SeparableConv2D layers in a neural network model to balance computational efficiency with rich feature extraction. Conv2D layers effectively capture complex spatial patterns by applying multiple filters to the entire input space, making them robust for learning complex features. However, they are computationally expensive. On the other hand, SeparableConv2D layers break down the convolution into depth-wise and point-wise operations, significantly reducing the number of parameters and computations while still capturing essential features. By combining these layers, a model can leverage the detailed feature extraction of Conv2D where needed while using SeparableConv2D to reduce computational costs in other parts of the network, leading to a

© 2024 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH

more efficient and scalable model. This unique combination of CNN and transformer components provided significant advantages, contributing to the novelty and effectiveness of our research. These innovations culminate in a robust, efficient, and adept model handling diverse image classification challenges.

Combining convolutional layers and transformer blocks enhances the model's ability to capture local spatial features and long-range dependencies. This hybrid approach enables more effective feature extraction and improves the model's overall performance in complex image classification tasks. Note, while under-sampling would only allow us to build our model using real datasets, here in our study, we have used an over-sampling strategy rather than under-sampling. The rationales behind that were, first, some of the "normal classes" of our dataset are much higher than the other three remaining classes (astrocytoma, glioblastoma, and oligodendroglioma), which may cause severe over-fitting issues for a multiclass classification problem if trained as it was. Second, although the under-sampling approach may alleviate the class imbalance issue, it would provide us with limited data points for training and testing the model, which would lack generalizability. Third, although the over-sampling strategies create new data but like existing ones, the robustness of the advanced techniques facilitates greater generalizability of the model by reducing the over-fitting or under-fitting issues, which is an extremely important aspect of a multiclass classification task. This system implementation computational criteria was an Intel (R) Core (TM) i7 CPU, 16 GB RAM, and 12 GB GPU was used for the entire training procedure on a Windows 10 computer. TensorFlow 2.2.1 and Python 3.12.3 implemented all offensive automatic glioma subtypes classification models. To edit this system code, we used Jupyter Notebook. Additionally, the memory consumption size of our model was 3.43 MB and the training time for our model was 57 min and 45 s. However, our system also has certain limitations. The MCT model was built explicitly for classification tasks. Its performance may be reduced in the fields of NLP, segmentation, and others.

Our system's enhanced capabilities will set a new standard for glioma classification, fostering advancements in clinical practice and research in neuro-oncology. This accomplishment is a notable advancement in the history of medical science, providing a revolutionary instrument for doctors and researchers. Our research contributes to advancing diagnostic capacities and ultimately leads to improved patient outcomes by examining the historical development in the medical industry.

5. Conclusion and Future Work

This research has successfully tackled the critical issue of glioma subtype classification by introducing an advanced deep learning model, MCT. The classification of glioma subtypes significantly impacts both research and clinical settings. Accurate subtype classification enables a deeper understanding of these diverse brain tumor molecular and genetic foundations, facilitating the identification of distinct prognostic and therapeutic targets. Clinically, it enhances diagnostic precision, allowing for more tailored and effective treatment plans that can improve patient outcomes. It also aids in predicting disease progression and response to therapies, thereby contributing to personalized medicine approaches. Additionally, refined classification systems support the development of more targeted clinical trials, advancing the discovery of novel treatments and ultimately improving survival rates and quality of life for glioma patients.

In the future, we will enhance our model to improve efficiency, by reducing the training time while delivering more accurate results. To enhance the performance metrics more advanced augmentation techniques, such as GANs, may be employed. Additionally, we will incorporate more image data from various hospitals to ensure better performance in glioma diagnosis and the healthcare sector.

Acknowledgements

The authors extend their appreciation to the King Salman Center for Disability Research for funding this work through Research Group number KSRG-2024-113.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

S. M. Nuruzzaman Nobel: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Methodology (equal); Software (equal); Validation (equal); Writing—original draft (equal). **S. M. Masfequier Rahman Swapno**: Conceptualization (lead); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Validation (equal); Writing—original draft (lead). **Md Babul Islam**: Investigation (supporting); Writing—review & editing (supporting). **AKM Azad**: Supervision (lead); Validation (lead); Writing—review & editing (lead). **Salem A. Alyami**: Funding acquisition (equal). **Md Alamin**: Visualization (equal); Writing—review & editing (equal). **Pietro Liò**: Validation (equal); Visualization (equal). **Mohammad Ali Moni**: Funding acquisition (lead); Supervision (lead); Visualization (lead); Writing—review & editing (lead).

Data Availability Statement

The data that support the findings of this study are openly available in DeepHisto: Dataset for glioma subtype classification from Whole Slide Images, reference number 7941080.

Keywords

brain tumor, computer vision, deep learning, disability Researches, glioma diagnosis, vision transformers, whole slide images

Received: July 9, 2024
Revised: August 30, 2024
Published online:

[1] J. Ferlay, H. Shin, F. Bray, D. Forman, C. Mathers, D. M. Parkin, *Int. J. Cancer* **2010**, *127*, 2893.
[2] Q. T. Ostrom, H. Gittleman, L. Stetson, S. M. Virk, J. S. Barnholtz-Sloan, *Cancer Treat. Res.* **2015**, *163*, 1.

© 2024 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH

[3] Y.-D. Cai, S. Zhang, Y.-H. Zhang, X. Pan, K. Feng, L. Chen, T. Huang, X. Kong, *J. Clin. Med.* **2018**, *7*, 350.
[4] M. Shi, D. Liu, Z. Yang, N. Guo, *Cancer Metastasis Rev.* **2013**, *32*, 603.
[5] C. Walker, A. Baborie, D. Crooks, S. Wilkins, M. D. Jenkinson, *Br. J. Radiol.* **2011**, *84*, S90.
[6] M. L. Goodenberger, R. B. Jenkins, *Cancer Genet.* **2012**, *205*, 613.
[7] D. N. Louis, A. Perry, G. Reifenberger, A. Von Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues, D. W. Ellison, *Acta Neuropathol.* **2016**, *131*, 803.
[8] G. S. Tandel, M. Biswas, O. G. Kakde, A. Tiwari, H. S. Suri, M. Turk, J. Laird, C. Asare, A. A. Ankrah, N. N. Khanna, B. K. Madhusudhan, L. Saba, J. S. Suri, *Cancers* **2019**, *11*, 111.
[9] Q. T. Ostrom, N. Patil, G. Cioffi, K. Waite, C. Kruchko, J. S. Barnholtz-Sloan, *Neuro-oncology* **2020**, *22*, iv1.
[10] S. M. M. R. Swapno, G. Chhabra, K. Kaushik, S. M. Nuruzzaman Nobel, M. Babul Islam, M. Shahiduzzaman, in *2023 Annual Int. Conf. on Emerging Research Areas: Int. Conf. on Intelligent Systems (AICERA/ICIS)*, IEEE **2023**, pp. 1–6.
[11] S. M. N. Nobel, S. M. M. R. Swapno, M. R. Islam, M. Safran, S. Alfarhood, M. F. Mridha, *Sci. Rep.* **2024**, *14*, 14435.
[12] M. B. R. Bappi, S. M. M. R. Swapno, M. M. F. Rabbi, in *2024 6th Int. Conf. on Electrical Engineering and Information & Communication Technology (ICEEICT)*, IEEE **2024**, pp. 445–450.
[13] M. C. Tom, D. P. Cahill, J. C. Buckner, J. Dietrich, M. W. Parsons, J. S. Yu, *Am. Soc. Clin. Oncol. Educ. Book* **2019**, *39*, 133.
[14] F. Arzanforoosh, S. R. van der Voort, F. Incekara, A. Vincent, M. van den Bent, J. M. Kros, M. Smits, E. A. H. Warnert, J. M. Kros, M. Smits, E. A. H. Warnert, *Cancers* **2023**, *15*, 2135.
[15] B. Björkblom, C. Wibom, M. Eriksson, A. T. Bergenheim, R. L. Sjöberg, P. Jonsson, T. Brännström, H. Antti, M. Sandström, B. Melin, *Neuro-oncology* **2022**, *24*, 1454.
[16] W.-W. Hsu, J.-M. Guo, L. Pei, L.-A. Chiang, Y.-F. Li, J.-C. Hsiao, R. Colen, P. Liu, *Sci. Rep.* **2022**, *12*, 6111.
[17] M. B. Ali, I. Y.-H. Gu, A. Lidemar, M. S. Berger, G. Widhalm, A. S. Jakola, *BMC Biomed. Eng.* **2022**, *4*, 4.
[18] S. Chen, *Front. Endocrinol.* **2022**, *13*, 879233.
[19] DeepHisto: Dataset for Glioma Subtype Classification from Whole Slide Images — zenodo.org., https://zenodo.org/records/7941080 (accessed: April 2024).
[20] A. Hassani, S. Walton, N. Shah, A. Abuduweili, J. Li, H. Shi (Preprint), arXiv:2104.05704, v4, Submitted: Apr. **2021**.
[21] A. M. K. V., V. R. Rajendran, P. J. K., *J. Med. Syst.* **2019**, *43*, 113.
[22] C. G. Bangalore Yogananda, B. R. Shah, M. Vejdani-Jahromi, S. S. Nalawade, G. K. Murugesan, F. F. Yu, M. C. Pinho, B. C. Wagner, B. Mickey, T. R. Patel, B. Fei, A. J. Madhuranthakam, J. A. Maldjian, *Neuro-oncology* **2020**, *22*, 402.
[23] P. P. Batchala, T. J. E. Muttikkal, J. H. Donahue, J. T. Patrie, D. Schiff, C. E. Fadul, E. K. Mrachek, M.-B. Lopes, R. Jain, S. H. Patel, *Am. J. Neuroradiol.* **2019**, *40*, 426.
[24] H. Cao, E. Z. Erson-Omay, X. Li, M. Günel, J. Moliterno, R. K. Fulbright, *Eur. Radio.* **2020**, *30*, 3073.
[25] E. Carver, J. Snyder, B. Griffith, N. Wen, *Neuro-Oncology* **2019**, *21*, vi61.
[26] Y. Fan, C. Chen, F. Zhao, Z. Tian, J. Wang, X. Ma, J. Xu, *Front. Oncol.* **2019**, *9*, 1164.
[27] E. D. H. Gates, J. S. Lin, J. S. Weinberg, S. S. Prabhu, J. Hamilton, J. D. Hazle, G. N. Fuller, V. Baladandayuthapani, D. T. Fuentes, D. Schellingerhout, *Am. J. Neuroradiol.* **2020**, *41*, 400.
[28] C. Jiang, Z. Kong, S. Liu, S. Feng, Y. Zhang, R. Zhu, W. Chen, Y. Wang, Y. Lyu, H. You, D. Zhao, R. Wang, Y. Wang, W. Ma, F. Feng, *Eur. J. Radiol.* **2019**, *121*, 108714.
[29] C. Jiang, Z. Kong, Y. Zhang, S. Liu, Z. Liu, W. Chen, P. Liu, D. Liu, Y. Wang, Y. Lyu, D. Zhao, Y. Wang, H. You, F. Feng, W. Ma, *Neuroradiology* **2020**, *62*, 803.
[30] M. H. Lee, J. Kim, S.-T. Kim, H.-M. Shin, H.-J. You, J. W. Choi, H. J. Seol, D.-H. Nam, J.-I. Lee, D.-S. Kong, *World neurosurg.* **2019**, *125*, e688.
[31] Y. Li, Y. Liang, Z. Sun, K. Xu, X. Fan, S. Li, Z. Zhang, T. Jiang, X. Liu, Y. Wang, *Neuroradiology* **2019**, *61*, 1229.
[32] C.-M. Lo, R.-C. Weng, S.-J. Cheng, H.-J. Wang, K. L.-C. Hsieh, *Medicine* **2020**, *99*, e19123.
[33] Y. Matsui, T. Maruyama, M. Nitta, T. Saito, S. Tsuzuki, M. Tamura, K. Kusuda, Y. Fukuya, H. Asano, T. Kawamata, K. Masamune, Y. Muragaki, *J. Neuro-oncol.* **2020**, *146*, 321.
[34] H. Mzoughi, I. Njeh, A. Wali, M. B. Slima, A. Benhamida, C. Mhiri, K. B. Mahfoudhe, *J. Digital Imaging* **2020**, *33*, 903.
[35] C. J. Park, Y. S. Choi, Y. W. Park, S. S. Ahn, S.-G. Kang, J.-H. Chang, S. H. Kim, S.-K. Lee, *Neuroradiology* **2020**, *62*, 319.
[36] Z. A. Shboul, J. Chen, K. M. Iftekharuddin, *Sci. Rep.* **2020**, *10*, 3711.
[37] X. Sun, P. Pang, L. Lou, Q. Feng, Z. Ding, J. Zhou, *J. Int. Med. Res.* **2020**, *48*, 0300060520914466.
[38] H. Tian, H. Wu, G. Wu, G. Xu, *BioMed Res. Int.* **2020**, *2020*, 3872314.
[39] Z. Zhang, J. Xiao, S. Wu, F. Lv, J. Gong, L. Jiang, R. Yu, T. Luo, *J. Digital Imaging* **2020**, *33*, 826.
[40] S.-S. Zhao, X.-L. Feng, Y.-C. Hu, Y. Han, Q. Tian, Y.-Z. Sun, J. Zhang, X.-W. Ge, S.-C. Cheng, X.-L. Li, L. Mao, S.-N. Shen, L.-F. Yan, G.-B. Cui, W. Wang, *BMC Neurol.* **2020**, *20*, 1.

© 2024 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH