# Evaluation of Classification of Brain Tumors Using Convolutional Neural Network Algorithm

Andrew Chain*
*Department of Computer Science and Engineering*
*Texas A&M University*
Texas, USA
andrew.chian@gmail.comline
*Corresponding author

Logan Coons
*Department of Chemistry and Biochemistry*
*University of Nevada*
Las Vegas, USA
coonsl2@unlv.nevada.edu

Adam Ramos
*Department of Biochemistry and Biophysics*
*Texas A&M University*
Texas, USA
adam.ramosmora@tamu.edu

Kya Richardson
*Department of Bioengineering and Biomedical Engineering*
*North Carolina A&T State University*
North Carolina, USA
krichardson1@aggies.ncat.edu

Kamaluddeen Usman Danyaro
*Department of Computer and Information Science*
*Universiti Teknologi PETRONAS*
Seri Iskandar, Malaysia
kamaluddeen.usman@utp.edu.my

Bimal Nepal
*Department of Engineering Technology and Industrial Distribution*
*Texas A&M University*
Texas, USA
nepal@tamu.edu

Mujaheed Abdullahi
*Department of Computer and Information Science*
*Universiti Teknologi PETRONAS*
Seri Iskandar, Malaysia
abdullahi_18001208@utp.edu.my

Hashir Sohrab
*Department of Industrial and Systems Engineering*
*Texas A&M University*
Texas, USA
hashirsohrab@tamu.edu

***Abstract*— Brain cancer is on the rise globally, with a significant increase in adult brain tumor cases in the last two decades. Detecting and treating brain tumors is challenging due to delayed diagnosis, asymptomatic presentation, and size and shape variations. Gliomas, slow-growing brain tumors, are classified by grade and type. These classifications are useful in predicting the tumor's growth rate and likelihood of recurrence. Brain tumors are categorized as benign or malignant. Medical image processing methods can be time-consuming, and accurate grading and typing guidance are scarce. To overcome these challenges, convolutional neural networks (CNN) are a deep learning model that can automatically learn and extract notable features from MRI images and selected machine learning tools to accomplish accurate classification of brain tumors. It is important to recognize brain tumors early on so that treatment can be given early in the progression of the disease. In this study, we propose evaluating of classification of brain tumors using CNN algorithms. The result achieved better performance, where the results of the confusion matrix achieved, an accuracy of 97%, precision of 97.7%, recall of 96.5%, and F1-score of 97.1%. The model has been validated using benchmark datasets from Kaggle which contains 3060 MRI images of brain tumors. The study findings indicate that overall model performance shows potential effectiveness for brain tumor classification. Finally, further readings have been provided.**

***Keywords*— *brain tumor classification, convolutional neural network, MRI images, high-grade glioma, image processing***

## I. Introduction

A brain tumor is one of the deadliest illnesses which occurs when there is an abnormal growth of cells in the brain. The tumor is a mass or lump formed by the uncontrolled division and proliferation of these cells. It can be classified as either benign (non-cancerous) or malignant (cancerous). It is important to achieve detection and diagnosis as early as possible due to the severity and unpredictability of this illness.

The most common method of detection comes via Magnetic Resonance Imaging (MRI) scans. MRI is a non-invasive medical imaging technique that can visualize and evaluate the location, size, and characteristics of brain tumors. Multiple types of images may be provided from an MRI scan. Commonly used types of MRI include T1-weighted, T2-weighted, and contrast-enhanced T1-weighted images. Each kind provides different information about the tumor's structure, composition, and vascularity. T1-weighted images can provide notable detail on the anatomy of the brain tumor, which may help identify the location and size of the tumor. T2-weighted images can differentiate tissue, water content, and swelling. Finally, contrast-enhanced T1-weighted images can differentiate areas of vascularity and blood-brain barrier disruption [1], [2]. Using MRI images to improve brain tumor classification can help determine the most appropriate treatment strategy for individual cases. In addition, it can help develop new treatment strategies in the future as well as serve as a steppingstone for future research. Classification algorithms may help forecast the future of the tumor as they can have different growth rates.

Moreover, technological advances have introduced many suggested methods to segment, analyze, and classify brain tumors. The aim of using convolutional neural networks (CNNs) is to improve diagnosis, treatment planning, and the monitoring of brain tumor patients [3]. CNNs are a deep learning model that can automatically learn and extract notable features from MRI images. CNNs aid medical professionals in the task of early detection and diagnosis. Following this, tumors can be accurately segmented as CNNs can learn to identify boundaries and regions of interest. The size, location, and shape of the brain tumor can be predicted as well [4], [5]. Additionally, accurate segmentation can help determine how much radiation is required for specific tumorous areas, which can help preserve healthy tissue. The goal is to accurately aid in the monitoring of the progression and effectiveness of the treatment.

The key characteristic of CNNs is their ability to automatically learn hierarchical representations of visual data. This has made it the preferred method over time as CNNs generally have better training and testing rates when compared to KNN and SVM as less preprocessing is required for CNNs [6]. They accomplish this through multiple layers that perform specific operations based on the input data. Convolutional layers apply learnable filters, which may be referred to as kernels, to the input image. Each filter extracts local patterns by sliding the filter across the image. From here, dot products can be computed between the filter weights and the corresponding image patches. Multiple filters are used to capture different types of features [3]. CNNs typically involve pooling layers and fully connected layers. Pooling layers can speed up training and compress data into "acceptable" dimensions to fit parameters. The most common pooling operation is max pooling, where some maximum value is used in a local neighborhood. Pooling helps reduce the spatial resolution, making the representation more compact while preserving important features. Fully connected layers are traditional neural network layers where each neuron is connected to every neuron in the previous and subsequent layers [3].

To give neural networks non-linearity, activation functions such as Rectified Linear Unit (ReLU) are generally used to improve the accuracy of a CNN model. Activation functions aid in the algorithm's ability to learn complex relationships and make predictions [3]. In the case of ReLU, negative values are set to zero, and positive values are left unchanged. During training, CNNs learn to optimize their parameters by giving specific weights to features and pixels in MRI images. This hierarchical architecture helps the algorithm optimize pattern recognition and feature selection. This is generally done by minimizing a loss function that quantifies the difference between what the algorithm predicts and what the actual target results are [3].

## II. LITERATURE REVIEW

Fox states that though the exact incidence of metastatic brain tumors is unknown, the incidence is significantly higher than that of primary brain tumors [7]. Metastatic brain tumors are the most frequent type of brain tumor in adults and knowing the epidemiologic factors associated with this tumor can decrease the likelihood of this disease.

Salehiniya and Farmanfarma [8] investigated the epidemiology and risk factors of neural cancer in the world and found the mortality rate of neural system cancers is estimated at 3.4 per 100,000 across the world. Some of the factors that are associated with an increase in the incidence and death of brain tumors in developed countries include urbanization, increased life expectancy, and lifestyle changes. Collins identified that classification and malignancy grading can be extremely difficult or impossible if therapy is performed before histological diagnosis [9]. Because of the difficulty in treating and detecting brain tumors, it is one of the deadliest diseases in the world [10]. Histopathology and molecular information can be beneficial in providing details of any pathological process present which will inform the appropriate treatment and prognosis needed. This study found that methods for recognizing targets of interest can demonstrate the expression by the tumor cells of an antigen typically expressed by a particular cell type which can assist in classification. It is important to note, though, that there are no antibodies that unequivocally identify the different tumor types.

The foundational challenges of brain tumor prediction via algorithms hinder accurate classification, segmentation, and forecasting promptly [11]. Touching or overlapping tumors can make segmentation difficult as discrepancies between unique tumorous and non-tumorous zones can be unclear [12]. Post-operative scar tissue and hemorrhages contribute to this issue [13]. The structure of brain tumors can also change as a patient improves or worsens. Challenges can extend beyond the nature of brain tumors as overfitting can result in machine learning giving accurate predictions for training data but not new data, thus limiting our understanding of the relationship between the variables [14].

Digital preprocessing allows for high-quality imaging machines to extract a variety of parameters. The spatial domain features and hybrid features are parameters that filter the image to extract important information. Iqbal walks through one brain tumor classification process of preprocessing to segmentation to feature extraction and then multiclass classification. Segmentation is another process applied to images to assist in classification. It allows for the brain tumor to be separated from normal brain tissue in the image analysis process. This review found that the classification process with or without segmentation is equally comparable [11].

Technological advances have introduced many suggested methods to analyze and classify brain tumors. Texture analysis of the apparent diffusion coefficient (ADC) map is a process that allows for the tumor volume to be analyzed to provide the Glioma grade. Skogen reports that the evaluation of tumor heterogeneity, attenuation, and size is correlated with the tumor grade which assists in the classification and grading of brain tumors. In agreement with the use of texture analysis, this study identified coarse texture as the best discriminator between low and high Glioma grades [15]. This also introduces the suggested improvement of quantifying tumor heterogeneity through data acquisition technologies. Krabbe specifies that the process of measuring ADC in enhancing and non-enhancing areas of the tumors allows for the examination of primary and secondary intracranial tumors. They found that ADC in contrast-enhancing areas within the cerebral metastases was statistically significantly higher than ADC in contrast-enhancing areas in high-grade Giloma [16]. Additionally, this research proposes an automated scheme that includes region-of-interest definition, feature extraction, feature selection, and classification. The classification accuracy of this automated tool is comparable to or higher than in other studies that do not use spectroscopy or diffusion sensor imaging. A limitation of this study is the need for region of interest tracing which causes the approach to be semiautomatic and subject to intra-observer and inter-observer variability [16].

When it comes to selecting a method that produces accurate results for brain tumor prediction and mitigates challenges, the preprocessing of MRI images is highly recommended. This methodology identifies the geometric features of a raw image given the original contrast, size, false boundaries, and noise content. Laplacian and Gaussian filtering (LoG) with contrastlimited Adaptive Histogram Equalization (CLAHE) are two staples of the preprocessing approach. CLAHE can remove false boundaries and improve image contrast, while LoG helps to smoothen and enhance image clarity. This is achieved through the removal of noise artifacts and the manipulation of weighted pixels [17], [4].

Before the tumors can be segmented, it is important to improve the image quality using a Gray-level Co-occurrence Matrix (GLCM), Spatial Gray-level Dependence (SGLDM), or similar algorithms [14], [4]. The purpose of these algorithms is to reflect the spatial distribution of gray levels in a region by examining the texture of the pixels in the image. The pixels can be viewed as variables as their angles and distances can reveal values for contrast, correlation, energy, and entropy. Morphological Edge Detection can examine pixels in similar ways helping to determine erosion and dilation, which helps outline the tumor [17]. Many of the studies use all the algorithms mentioned above in that order before proceeding with feature selection, classification, and segmentation.

The best algorithms to accomplish these goals are CNN's. Not only are they the most widely used algorithm for predicting brain tumors, but they are also the most accurate. Two specific CNN algorithms stood out in terms of accuracy and training time. AlexNet is a CNN that achieved 99.04% accuracy in a comparison with 4 other commonly used CNN algorithms. However, it was assisted with a SGDM optimizer that, when paired with AlexNet, proved to be the most efficient CNN and optimizer posting a 46 min and 31 sec training time, and the fastest training times with each algorithm [1]. PSO is a CNN algorithm that achieved a study-best 99.60% accuracy, 84.54% recall, and a 90.10 F-score in a comparison amongst 7 other commonly used CNN algorithms [2].

An optimizer such as the Stochastic Gradient of Descent with Momentum (SGDM) will accelerate the descent of a desired path and reduce its oscillations. This can help stop network overfitting as susceptible algorithms may deviate from the desired path due to over-training. In a comparison of 5 algorithms and 3 optimizers, SGDM proved to be the most efficient posting a 46 min and 31 sec training time, and the fastest training times with each algorithm [1]. To further minimize training time and overfitting, the sweet-spot training to testing/validation ratio should be 70% training to 30% testing and validation or 80/20 according to other sources [1]. K-Nearest Neighbor (KNN) is also commonly used and reported 97.65% accuracy [18]. However, KNN localizes features and only remembers them for a specific spot [19]. In other words, they rely on the nearest neighborhood data point to make decisions. KNNs have also proven to be slow and inefficient when handling large and high-dimensional datasets [20]. CNNs are better at generalizing large and complex sets of data to any part of the image as they extract features from the input data [5]. In one other instance, KNN reported an accuracy of only 76% [4]. One other paper favors a CNN model against others based on key features and picks KNN as the least favorable [16].

Optimizers such as ADAM and RMSProp reported longer training times with the 5 tested algorithms compared to SGDM. AlexNet was recommended in my analysis because it reported the highest average accuracy for benign and malignant tumors regardless of the optimizer being used. SVM is a powerful classification model in machine learning that finds an optimal hyperplane that best separates its feature vectors into classes [5]. CNN includes convolutional neural networking and deep learning. Using large sample sets results in an SVM accuracy of 0.88 and a CNN accuracy of 0.98. However, when using smaller sample sets, the accuracy of SVM is 0.86 and CNN is 0.83. Generally, the best method for smaller datasets is SVM as they have a training complexity that grows quadratically with the number of samples. Given the fact that potentially over 1000 MRI images will be analyzed in a study, we can conclude that CNN's will be the more accurate approach [5].

## III. METHODOLOGY

A dataset of 3,000 MRI images was preprocessed to fit a standardized size and color palette. After this step, images were split into a training and experimental data set, with a 80%/20% split of preprocessed MRI images. Utilizing API layers from Keras, we developed a 5-layer convolutional neural network. CNN works to segment the images inputted into the CNN machine, which further classifies the images as malignant or not. Output is then given to a physician who will make the final decision and diagnosis.

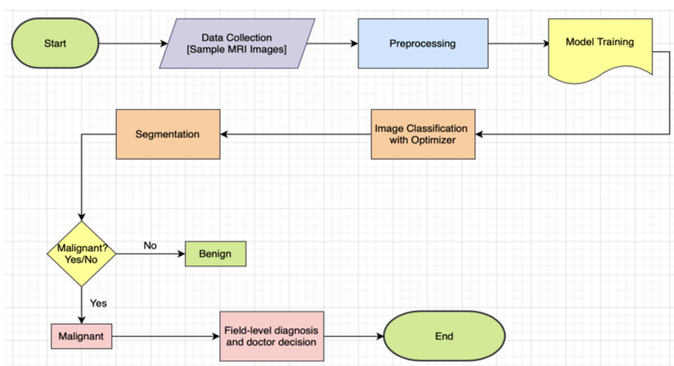


Fig. 1. Methodology Flow Chart

### A. Dataset

To present a fundamental design applying the convolutional neural network in classifying brain tumors, a publicly available Kaggle dataset was obtained from https://www.kaggle.com/datasets/abhranta/brain-tumor-detection-mri. The dataset contains a total of 3060 images and will be split into 2 separate parts, one for training purposes and one for testing purposes. The dataset is split into 1,500 images containing MRI scans containing a brain tumor and 1,500 images containing MRI scans that do not contain a brain tumor. 60 images within the dataset are MRI scans that are unlabeled and are used for making predictions. The dataset is made up entirely of MRI scans and all the MRI scans were used in the research.

### B. Image Preprocessing

To make sure the images can be used effectively during model compilation and training; they first need to be pre-processed. During this stage, images will all be resized and converted to grayscale to make sure they are consistent during the training phase. After pre-processing, each image will have a dimension of 150 x 150 pixels. When the images are resized to these dimensions, the aspect ratio will be retained. After pre-processing the images, the model performance will be enhanced.

## *C. Model Training*

The architecture used for this research is a Convolutional Neural Network (CNN). CNN is often used for classification tasks and will take images as input. It involves multiple layers and pooling operations before making a final decision. The model is trained using a 2D convolutional layer with 64 filters and a 5x5 kernel size. After this, the Leaky ReLu activation function is used to avoid losing values. Following this, a max pooling layer is used with a dropout layer. After the convolutional blocks, the data is flattened into a 1-dimensional vector called the flattened layers. The final output layer will utilize a sigmoid function to output a probability between 0 and 1.


Fig. 2. Dataset

## *D. Metric Evaluation*

The metric evaluation for this study employed accuracy, precision, recall, and F1-score which are calculated as follows:

Accuracy is the proportion of all predictions that are correct. It is calculated as follows:

$$Accuracy = \frac{True\ Positive + True\ Negative}{Total} \quad (1)$$

Precision is the proportion of predicted positives that are actually positive. It is calculated as follows:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (2)$$

Recall is the proportion of all actual positives that are correctly predicted. It is calculated as follows:

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (3)$$

The score is a measure that combines precision and recall into a single metric. It is calculated as follows:

$$F1 - score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (3)$$

# IV. RESULTS AND DISCUSSION

This section provides the result and discussion of the proposed method. The convolutional neural network (CNN) algorithm's training and validation accuracy for brain tumor classification is depicted in Figure 3. The graph demonstrates how, as the algorithm is trained on an increasing number of examples, the training accuracy rises steadily. Although it does not increase as quickly as the training accuracy, the validation accuracy still rises. This suggests that the algorithm is not overfitting to the training set, which is encouraging. Although the 97% accuracy is excellent, it's crucial to remember that this is the validation accuracy. It is likely that the test accuracy will be lower, i.e., the accuracy on a set of examples that the algorithm has never seen before. Nonetheless, a 97% validation accuracy indicates that the algorithm should work well with fresh data. Overall, the training and validation accuracy graph indicates that the CNN algorithm is a potentially useful method for brain tumor classification.

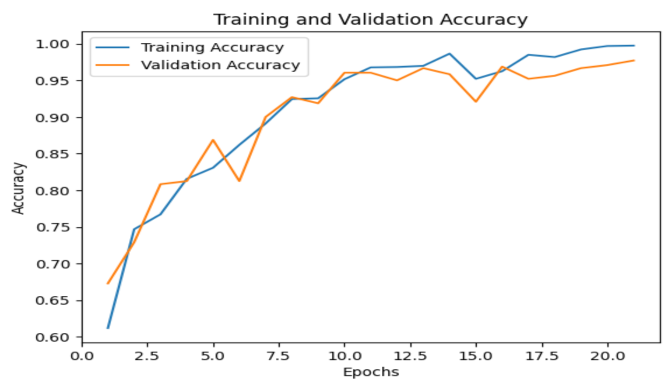


Fig. 3. Training and validation accuracy Plot.

The training and validation loss of an algorithm for classifying brain tumors using convolutional neural networks (CNNs) is depicted in Figure 4. The average algorithmic error on training data is known as the training loss, and the average algorithmic error on validation data is known as the validation loss. The graph demonstrates how, as the algorithm is trained on an increasing number of examples, the training loss gradually drops.

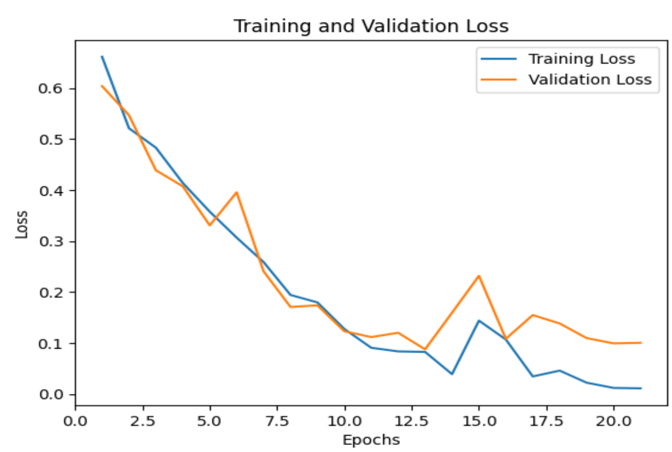


Fig. 4. Training & Validation Loss Plot.

Although it does not decrease as quickly as the training loss, the validation loss still decreases. This suggests that the algorithm is not overfitting to the training set, which is encouraging. A loss of 0.6% is excellent for a difficult task like brain tumor classification. It implies that even with fresh data, the algorithm can produce very accurate predictions. All things considered, the training and validation loss graph indicates that the CNN algorithm is a potentially useful method for brain tumor classification.

Figure 5 indicates the confusion matrix result, where there are 302 instances of true positive, 11 instances of false negative, 7 instances of false positive, and 280 instances of true negative. A confusion matrix can give us much more information about the model's performance including, accuracy, precision, recall, and F1-score. Using the results of the confusion matrix, an accuracy of 97%, a precision of 97.7%, a recall of 96.5%, and an F1-score of 97.1% is achieved.

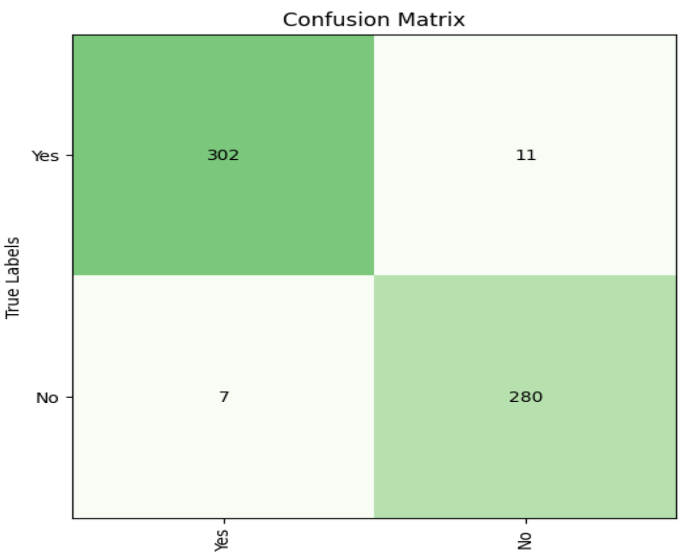


Fig. 5. Confusion Matrix.

## V. CONCLUSION

In this study, CNNs were successfully used to classify a set of MRI images both accurately and precisely. The chosen method was well-suited for this larger, more complex dataset as it showed high training and testing rates as well as sufficient ability to generalize features to any part of the MRI image. Limitations such as overtraining were mitigated by utilizing an 80:20 training-to-testing split. Although an accuracy rate of 97% and a precision rate of 96.22% were reported, improvements can be made to the study. Classification and segmentation are made difficult as every brain tumor has its own unique boundaries, scar tissue, and level of aggression. For these reasons, a larger, more diverse set of MRI images would be more representative of the variety of brain tumors being diagnosed daily. Most studies, including this one, were done on pre-operative patients, but patients in the post-operative state may still benefit from accurate classification. Therefore, another way this study could be improved is by implementing post-operative MRI images into the dataset. This will aid in the development of CNNs and other methods of classification such as KNNs and SVM as they learn to classify in the presence of scar tissue and hemorrhages.


## ACKNOWLEDGEMENT

The authors would like to thank Universiti Teknologi PETRONAS (UTP) in Malaysia, University of Nevada Las Vegas (UNLV), Texas A&M University (TAMU), North Carolina Agricultural Technical State University (NCAT) in the USA. This work was supported by the National Science Foundation's International Experience for Students (IRES) Site grant (Grant Numbers: OISE# 1952490-TAMU, 2208801-NCAT, and 1952497-UNLV). Any opinions, findings, conclusions, or recommendations presented are those of the authors and do not necessarily reflect the views of the National Science Foundation.


## REFERENCES


[1] R. Mehrotra, M. A. Ansari, R. Agrawal, and R. S. Anand, "A Transfer Learning approach for AI-based classification of brain tumors," *Machine Learning with Applications*, vol. 2, p. 100003, 2020, doi: https://doi.org/10.1016/j.mlwa.2020.100003.

[2] M. Saeidifar, M. Yazdi, and A. Zolghadrasli, "Performance Improvement in Brain Tumor Detection in MRI Images Using a Combination of Evolutionary Algorithms and Active Contour Method," *J Digit Imaging*, vol. 34, no. 5, pp. 1209–1224, 2021, doi: 10.1007/s10278-021-00514-6.

[3] H. Mzoughi *et al.*, "Deep Multi-Scale 3D Convolutional Neural Network (CNN) for MRI Gliomas Brain Tumor Classification," *J Digit Imaging*, vol. 33, no. 4, pp. 903–915, 2020, doi: 10.1007/s10278-020-00347-9.

[4] R. Thillaikkarasi and S. Saravanan, "An Enhancement of Deep Learning Algorithm for Brain Tumor Segmentation Using Kernel Based CNN with M-SVM," *J Med Syst*, vol. 43, no. 4, p. 84, 2019, doi: 10.1007/s10916-019-1223-7.

[5] P. Wang, E. Fan, and P. Wang, "Comparative analysis of image classification algorithms based on traditional machine learning and deep learning," *Pattern Recognit Lett*, vol. 141, pp. 61–67, 2021, doi: https://doi.org/10.1016/j.patrec.2020.07.042.

[6] M. M. Taye, "Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions," *Computation*, vol. 11, no. 3, 2023, doi: 10.3390/computation11030052.

[7] B. D. Fox, V. J. Cheung, A. J. Patel, D. Suki, and G. Rao, "Epidemiology of Metastatic Brain Tumors," *Neurosurg Clin N Am*, vol. 22, no. 1, pp. 1–6, 2011, doi: https://doi.org/10.1016/j.nec.2010.08.007.

[8] H. Salehiniya and K. Kalan Farmanfarma, "Brain Cancer in the World: An Epidemiological Review," Mar. 2020, doi: https://doi.org/10.32113/wcrj_20197_1356.

[9] V. P. Collins, "Brain tumours: classification and genes," *Journal of Neurology, Neurosurgery*

& *Psychiatry*, vol. 75, no. suppl 2, p. ii2, Jun. 2004, doi: 10.1136/jnnp.2004.040337.

[10] M. A. Gómez-Guzmán *et al.*, "Classifying Brain Tumors on Magnetic Resonance Imaging by Using Convolutional Neural Networks," *Electronics (Basel)*, vol. 12, no. 4, 2023, doi: 10.3390/electronics12040955.

[11] S. Iqbal, M. U. G. Khan, T. Saba, and A. Rehman, "Computer-assisted brain tumor type discrimination using magnetic resonance imaging features," *Biomedical Engineering Letters*, vol. 8, no. 1. Springer Verlag, pp. 5–28, Feb. 01, 2018. doi: 10.1007/s13534-017-0050-3.

[12] M. A. Khan *et al.*, "Brain tumor detection and classification: A framework of marker-based watershed algorithm and multilevel priority features selection," *Microsc Res Tech*, vol. 82, no. 6, pp. 909–922, Jun. 2019, doi: https://doi.org/10.1002/jemt.23238.

[13] P. J. Sørensen *et al.*, "Evaluation of the HD-GLIO Deep Learning Algorithm for Brain Tumour Segmentation on Postoperative MRI," *Diagnostics*, vol. 13, no. 3, 2023, doi: 10.3390/diagnostics13030363.

[14] V. Jeevanantham and G. MohanBabu, "Detection and diagnosis of brain tumors-framework using extreme machine learning and CANFIS classification algorithms," *Int J Imaging Syst Technol*, vol. 31, no. 2, pp. 540–547, Jun. 2021, doi: https://doi.org/10.1002/ima.22479.

[15] K. Skogen, B. Ganeshan, C. Good, G. Critchley, and K. Miles, "Measurements of heterogeneity in gliomas on computed tomography relationship to tumour grade," *J Neurooncol*, vol. 111, no. 2, pp. 213–219, 2013, doi: 10.1007/s11060-012-1010-5.

[16] K. Krabbe, P. Gideon, P. Wagn, U. Hansen, C. Thomsen, and F. Madsen, "MR diffusion imaging of human intracranial tumours," *Neuroradiology*, vol. 39, no. 7, pp. 483–489, 1997, doi: 10.1007/s002340050450.

[17] C. J. J. Sheela and G. Suganthi, "Morphological edge detection and brain tumor segmentation in Magnetic Resonance (MR) images based on region growing and performance evaluation of modified Fuzzy C-Means (FCM) algorithm," *Multimed Tools Appl*, vol. 79, no. 25, pp. 17483–17496, 2020, doi: 10.1007/s11042-020-08636-9.

[18] L. El Fatimi and H. Boucheneb, "Tumor Growth Prediction and Classification Based on the KNN Algorithm and Discrete-Time Markov Chains (DTMC)," *Neural Comput. Appl.*, vol. 35, no. 13, pp. 9739–9751, Jan. 2023, doi: 10.1007/s00521-023-08212-w.

[19] R. S. Chugh, V. Bhatia, K. Khanna, and V. Bhatia, "A Comparative Analysis of Classifiers for Image Classification," in *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 2020, pp. 248–253. doi: 10.1109/Confluence47617.2020.9058042.

[20] I. Triguero, D. García-Gil, J. Maillo, J. Luengo, S. García, and F. Herrera, "Transforming big data into smart data: An insight on the use of the k-nearest neighbors algorithm to obtain quality data," *WIREs Data Mining and Knowledge Discovery*, vol. 9, no. 2, p. e1289, Mar. 2019, doi: https://doi.org/10.1002/widm.1289.

**Authors' background**

| Name | Prefix | Research Field | Email | Personal website |
|---|---|---|---|---|
| Andrew Chain | Undergraduate Student | | andrew.chian@gmail.com | |
| Logan Coons | Undergraduate Student | | coonsl2@unlv.nevada.edu | |
| Adam Ramos | Undergraduate Student | | adam.ramosmora@tamu.edu | |
| Kya Richardson | Undergraduate Student | | krichardson1@aggies.ncat.edu | |
| Kamaluddeen Usman Danyaro | Lecturer | | kamaluddeen.usman@utp.edu.my | |
| Bimal Nepal | Professor | | nepal@tamu.edu | |
| Mujaheed Abdullahi | PhD Candidate | | abdullahi_18001208@utp.edu.my | |
| Hashir Sohrab | PhD Candidate | | hashirsohrab@tamu.edu | |

**Note:**

[1] **This form helps us to understand your paper better; the form itself will not be published.**

[2] *Prefix*: can be chosen from Master Student, PhD Candidate, Assistant Professor, Lecturer, Senior Lecturer, Associate Professor, Full Professor