

Brain tumor detection

Amelia Lunning, Eddie Jang

December 14, 2025

1 Introduction

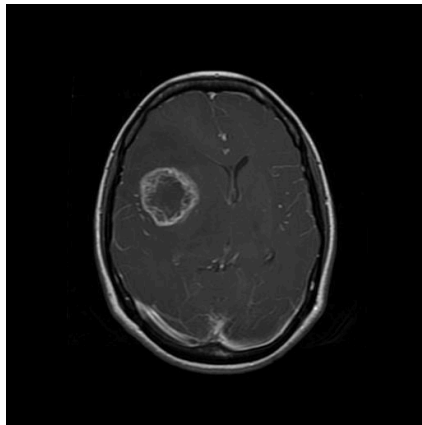
Brain tumors are destructive, and a misdiagnosis leads to incorrect medical care. An accurate diagnosis is critical for the health and survival of the patient. Brain tumors fall into two types: malignant and benign. Malignant tumors are cancerous and grow quickly, while benign tumors are non-cancerous and grow slowly. Some common types of tumors are glioma, meningioma, and pituitary.

Gliomas are tumors that grow within the substance of the brain and are the most common type of brain tumors. High-grade glioma tumors are one of the most aggressive types of brain tumors with a minimal survival of two years (Rehman et al.). Meningiomas are the most common type of benign tumors, they originate from the membranes surrounding the brain and spinal cord. Pituitary tumors are an irregular growth of brain cells that develop in the pituitary gland of the brain, and they're commonly benign.

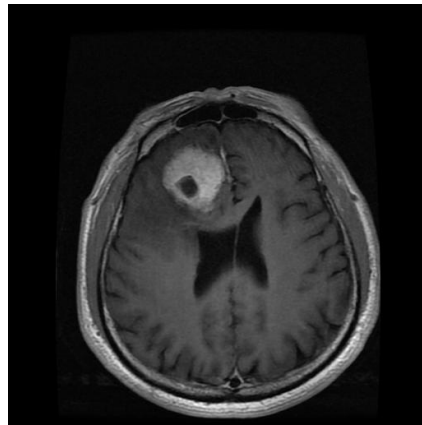
We wanted to answer the following question: given a human MRI brain scan, what type of tumor does it have? We did a study of three models to classify brain tumors from MRI scans. The models used were a multiclass KNN model, a CNN model, and the pretrained CNN model VGG-16. Our goal was to classify the MRI scans into the following classes: glioma, meningioma, pituitary, and no tumor.

2 MRI Scan Dataset

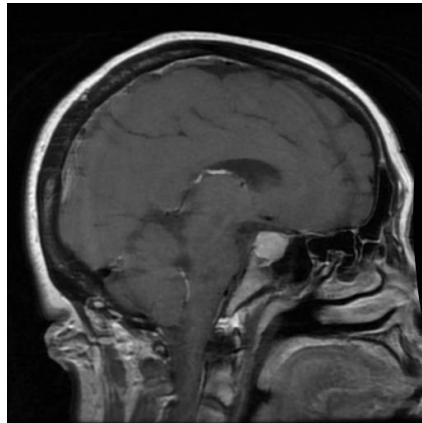
We used the Kaggle Brain Tumor MRI Dataset. It contains 7023 images of human brain MRI images which are classified into 4 classes: glioma, meningioma, no tumor, and pituitary. It is already split into testing/training data and it's roughly balanced, with testing containing 1311 images (300 glioma, 306 meningioma, 405 no tumor, 300 pituitary) and training containing 5712 images (1321 glioma, 1339 meningioma, 1595 no tumor, 1457 pituitary). Each type of tumor can be seen in Figure 1.



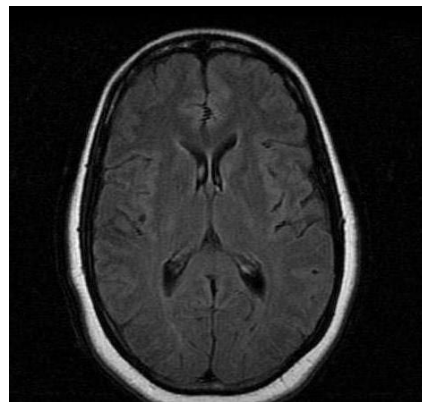
Glioma



Meningioma



Pituitary



No tumor

Figure 1. Four types of brain tumor

As the data is already balanced, no data balancing strategies are needed. Since the data's format is image, we needed to convert them into tensors for our models. We used `ToTensor()` transform. This transform is provided in the `torchvision.transforms` package. Using this transform, we can convert a PIL image into a torch tensor of data type `torch.float32` in range 0 and 1. Then the image was normalized with standard deviation and mean of ImageNet, ensuring all features contribute equally. Each model had its own unique approach of data augmentation and preprocessing.

3 KNN

We implemented a K-Nearest Neighbor algorithm (KNN) that used the training dataset as neighbors. We evaluated the accuracy of the testing set using every K value in the range [1, 100] to determine the best performing number of neighbors, which gave us the recommendation of a value between 3 and 5. We also used `GridSearchCV` for tuning the hyperparameters k and weights. The best performing configuration was $k = 3$ and `weights = 'distance'`.

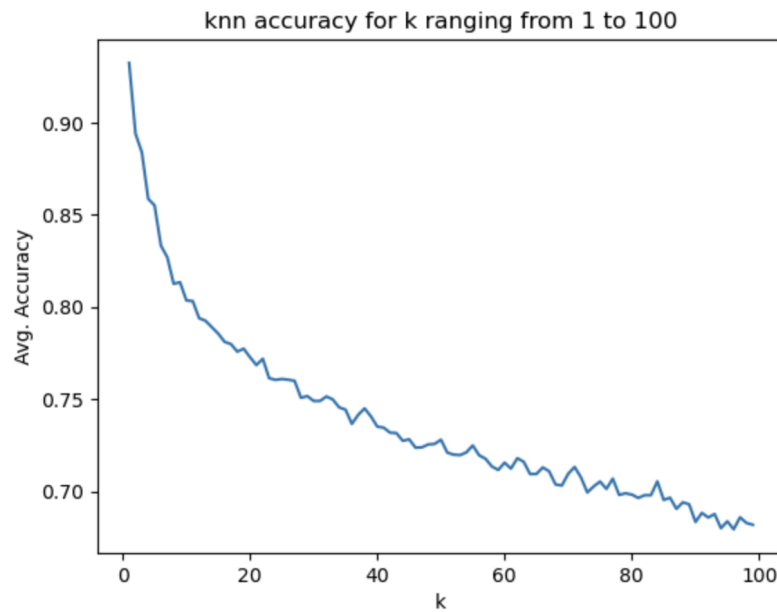


Figure 2. KNN Accuracy for K ranging from 1 to 100

We created a custom `MRICropDataset` which dynamically cropped the images to isolate the brain and reduce unnecessary background noise. The image vectors

were normalized using StandardScaler, which standardizes each feature to zero mean and unit variance.

4 Neural Network

A ten layer feed-forward convolutional neural network with max-pooling, flattening, ReLu, and convolutions was used for further classification analysis. We implemented L2 regularization by adding weight_decay that is greater than zero on the Adam optimizer. We also used various methods to prevent overfitting such as normalization, data augmentation, train-test split, and early stopping. Using Cross Entropy Loss with Stochastic Gradient Descent at batch size of 8 has improved accuracy of the model. It was trained on epoch of 10, and all the hyperparameters were tuned by trial and error.

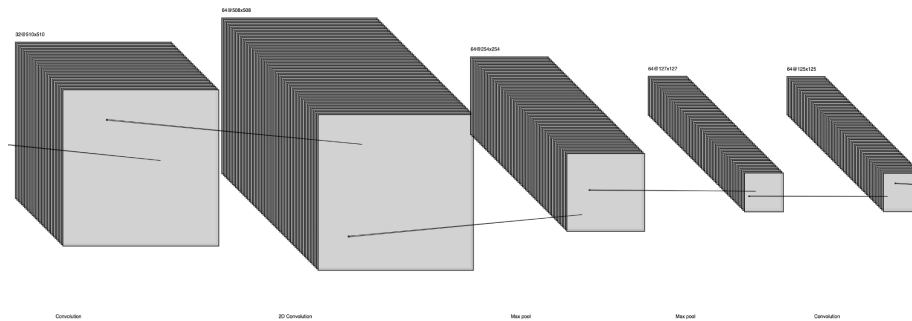


Figure 3. CNN Layers (Image cut due to size limitation)

5 Pretrained Model (VGG-16)

We used a pretrained VGG-16 convolutional neural network from the Torchvision library. VGG-16 was originally trained on ImageNet, which is a large database of labeled images. We chose to use VGG-16 because of previous research done that compared the accuracy of the pretrained models AlexNet, GoogLeNet, and VGG-16. VGG-16 had the highest accuracy of 98.69% (Rehman et al.).

VGG-16 was used with ImageNet weights as a feature extractor. The model was trained for 5 epochs using the Adam optimizer and binary cross-entropy loss. Input images were resized and normalized to match VGG-16 input requirements, which is 224.

6 Results

To evaluate model performance, we report overall accuracy, per-class precision, recall, and F1-score, as well as confusion matrices to analyze class-specific errors. We also examine the three most common misclassifications to better understand model limitations.

6.1 KNN

The KNN model achieved an overall accuracy of 96% on the test set. Table 1 reports the per-class precision, recall, and F1-scores. The KNN model performed strongly across all classes, with the highest glioma recall score of 96%.

	Precision	Recall	F1-Score
Glioma	0.90	0.96	0.93
Meningioma	0.96	0.87	0.91
No Tumor	0.98	0.99	0.99
Pituitary	0.99	1.00	0.99

Table 1. KNN Class Metrics

As shown in Figure 4 and Figure 5, the most common misclassification was predicting glioma when it was meningioma, which occurred in 30 cases. This is likely due to the similar visual characteristics between the two tumor types.

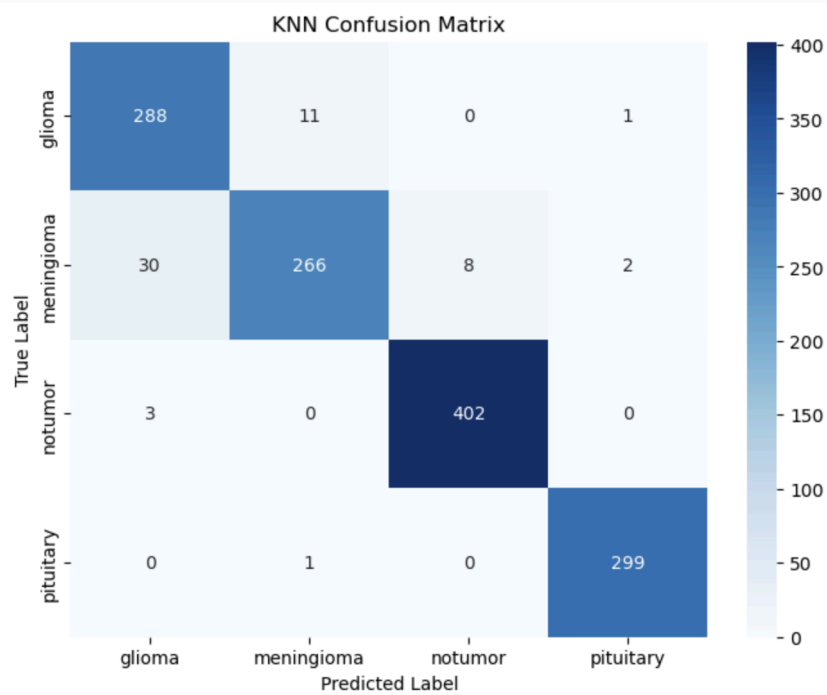


Figure 4. KNN Confusion Matrix

Top 3 Most Common Misclassifications

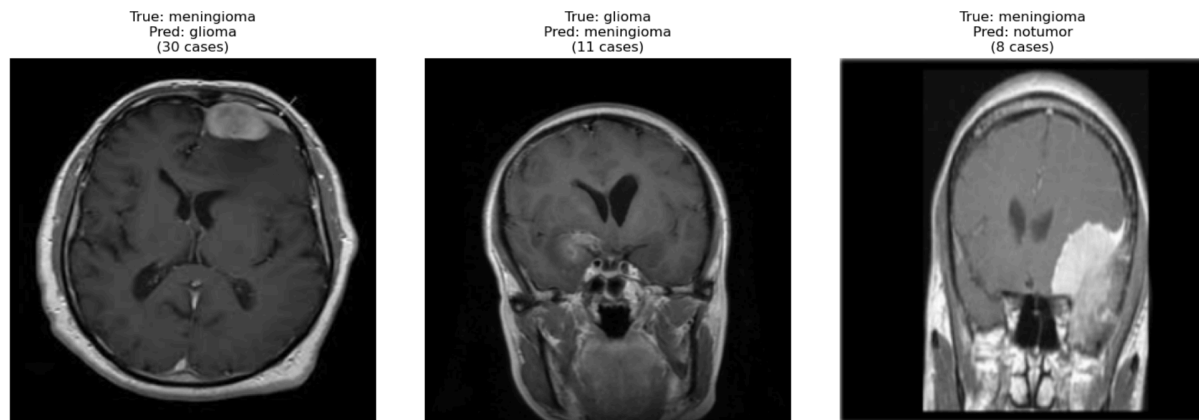


Figure 5. Three Most Common KNN Misclassifications

6.2 Neural Network

Among all the models tested, the CNN model performed the worst, but still high of 95% accuracy on the test set. Table 2 reports the per-class precision, recall, and F1-scores. Similar to KNN, it has a high recall on Glioma but it showed particularly low recall on meningioma. This is due to the fact that many misclassifications happen among glioma and meningioma.

	Precision	Recall	F1-Score
Glioma	0.88	0.94	0.91
Meningioma	0.94	0.77	0.85
No Tumor	0.94	0.98	0.96
Pituitary	0.94	0.98	0.96

Table 2. CNN Class Metrics

Similar to the KNN model, the most common misclassification was predicting glioma when it was meningioma, which occurred in 40 cases as shown in Figure 6 and Figure 7.

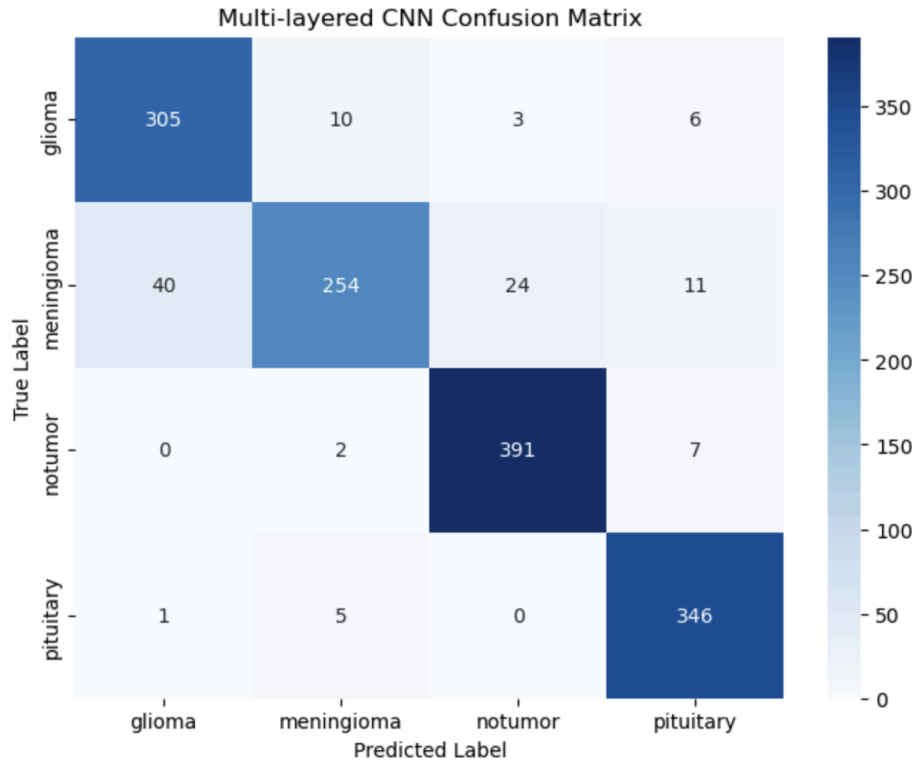


Figure 6. CNN Confusion Matrix

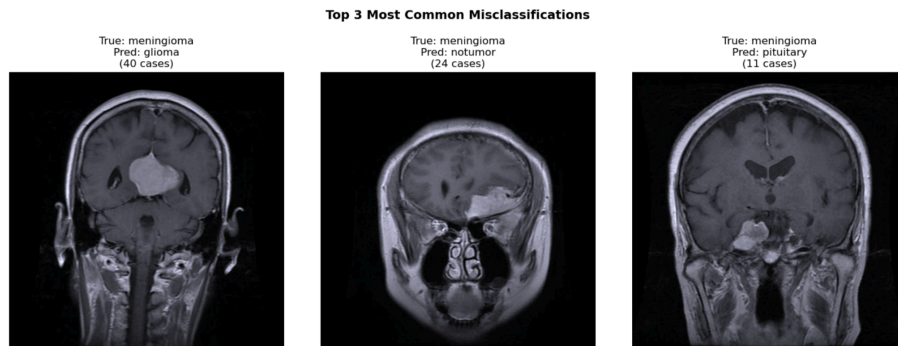


Figure 7. Three Most Common CNN Misclassifications

6.3 Pretrained Model (VGG-16)

The VGG-16 model also had an overall accuracy of 96% on the test set. Table 3 shows the per-class metrics for the model. VGG-16 achieved perfect scores for the no tumor classification.

	Precision	Recall	F1-Score
Glioma	0.95	0.90	0.93
Meningioma	0.90	0.93	0.92
No Tumor	1.00	1.00	1.00
Pituitary	0.97	0.98	0.98

Table 3. VGG-16 Class Metrics

As illustrated in the confusion matrix in Figure 8 and the misclassification diagram in Figure 9, the most common misclassification was classifying meningioma as glioma, which occurred in 24 cases.

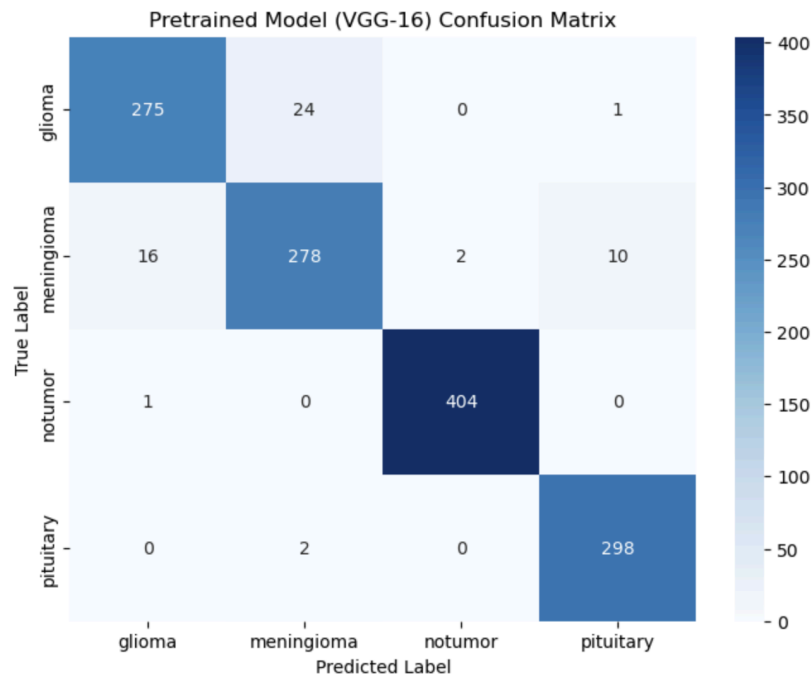


Figure 8. VGG-16 Confusion Matrix



Figure 9. Three Most Common VGG-16 Misclassifications

7 Discussion

All three models achieved strong performance, with KNN and VGG-16 scoring the highest accuracies of 96%. The most frequent misclassification among all the KNN and CNN models was predicting glioma when it was meningioma. The CNN model did this the most, with 40 total cases.

We realized that having high accuracy does not always mean the best model to use for classification; the context of misclassification is more important, particularly in case of medical conditions. With glioma being the most aggressive brain tumor among the ones we classified, there is greater importance in correctly identifying these. In this case, recall is the most important because misclassification of glioma tumors can drastically reduce survival rate. The pretrained KNN model has the highest recall for glioma, achieving 96%.

It is recommended to avoid using the pretrained model because it misclassifies glioma as meningioma. When dealing with medical data classification, we want to minimize false negatives. Although there is no true and positive since the model was a multi-class classification model, we could divide the tumors into two types: malignant and benign where malignant tumors could be treated as positives (for being cancerous) and benign tumors could be treated as negatives. Since glioma

tumors fall into positive categories, misclassifying glioma as meningiomas could cause critical outcomes.

Between the KNN model and the CNN model, using the KNN model is recommended because not only it has higher accuracy but also it excels in identifying glioma, the most aggressive type of tumor. In addition, since all precision, recall, and F1-Score values of the KNN model are higher than that of the CNN model, it is recommended to use the KNN model.

Surprisingly, the CNN model performed the worst, even worse than KNN. This might be due to overcomplication of the layers (e.g. having too many layers). For future reference, not only tuning hyperparameters but also tuning the number and structure of the convolutional layer could improve the performance.

Another limitation of the model is that the dataset used is relatively small, with only 7023 total images, which makes the model prone to overfitting. To avoid overfitting, early stopping and train-test validation along with different methods such as regularization were adopted. Potential use cases for the models in the real world would be a diagnosis assistant model for neuroradiologists; however, before these models could be integrated into the real world system, more data should be tested, and enhancing the results like accuracy and recall for glioma would be necessary to minimize false negatives.

8 Acknowledgments

Rehman, Arshia, et al. "A Deep Learning-Based Framework for Automatic Brain Tumors Classification Using Transfer Learning." *Circuits, Systems, and Signal Processing*, vol. 39, no. 2, 10 Sept. 2019, pp. 757–775, <https://doi.org/10.1007/s00034-019-01246-3>.

Msoud Nickparvar. (2021). Brain Tumor MRI Dataset [Data set]. Kaggle. <https://doi.org/10.34740/KAGGLE/DSV/2645886>