Brain Tumor Classification: CS 464 Group 7 Project Final Report

Mehmet Can Eroğlu 21801979 / Murat Sevinç 21702603 / Oğuz Salar 21803752 / Matin Moradi 22104108 / Arda Barış Örtlek 21903472

1. Introduction and Background Information

Brain cancers are the diseases which should be detected as early as possible for benefits of patients. Moreover, whether the tumor is malignant or benignant and which tumor type is are other important information after detection of a tumor.

In our work, we consider 4 situations: **no tumor**, **meningioma** tumors (which are usually benignant), **glioma** tumors (which are more prevalent and malignant) and **pituitary** tumors (which are less prevalent than previous ones and whose located in pituitary gland) [2]. Our dataset (which was provided from [1]) includes 2000 no tumor photos, 1645 meningioma photos, 1621 glioma photos and 1757 pituitary photos.

We have been aiming to train deep learning models that classify brain MRI images. We have planned to use Convolutional Neural Network (CNN) models and a Vision Transformer (ViT) model for this classification duty.

2. Classification with Convolutional Neural Networks

After dataset was prepared, firstly we tried to code (single) CNN models thanks to Pytorch Library via Google Colab and we examined performances of some pre-trained models [3] in this project. Pre-trained CNN models provide us special networks, which are designed carefully with classification purpose, and pre-trained weights, which ease our training process because of their previous classification trainings.

We have written our code utilizing and modifying similar Pytorch image classification projects in the Internet: [4] and [5]. Our project takes rar file dataset and splits this into training/validation/test folders with given ratios (70% / 20% / 10%). Then a pre-trained model is uploaded and data loaders are built with some image augmentations/transforms. Training function definition with selected loss functions and optimizers, plotting the results of training, testing the trained model with test dataset and making prediction for internet images are also included in our Colab Project.

We examined 3 pretrained CNN models: Resnet152, Densenet 161 and Efficient Net B4. We connected these with a simple classifier network consisting of 1024 and 4 (number of classes) neurons. As data augmentation/transform methods, we applied random rotation (between 0 and 20 degrees), random flipping, random changing the sharpness and cropping. Then we have made some trainings with some optimizers and loss function options to have intuition about their performances.

2.1 First Training: Resnet152 / Adam Optimizer / NLL Loss

After 50 epochs with Adam Optimizer and Negative Likelihood Loss, validation accuracy became around 95%. Max validation accuracy rate (94.6) was reached at 47th epoch. The loss and accuracy values can be seen below:

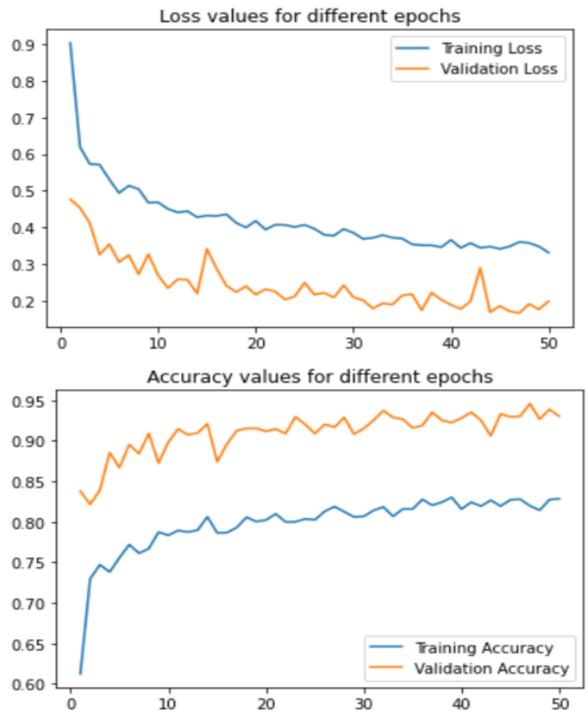

Fig. 1: Results of First Training

The accuracy rate of the trained model for test dataset is 93.57%. Then this model is used for some pictures from the internet whose labels are respectively: **pituitary**, **meningioma** and **no tumor**. Prediction scores can be seen from Figure 3. Prediction 1s for each picture are predictions.

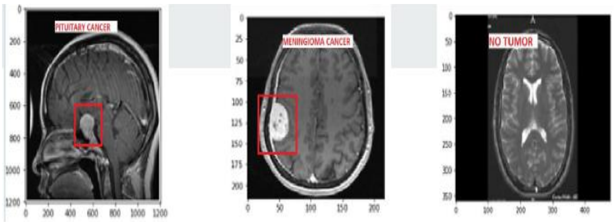

Fig. 2: Pituitary, Meningioma and No Tumor Photos

```
Prediction 1 : pituitary, Score: 51.82591676712036%
Prediction 2 : meningioma, Score: 43.813639879226685%
Prediction 3 : glioma, Score: 2.435963600873947%
Prediction 4 : notumor, Score: 1.9244838505983353%

Prediction 1 : meningioma, Score: 87.5436782836914%
Prediction 2 : notumor, Score: 12.12019994854927%
Prediction 3 : pituitary, Score: 0.23530283942818642%
Prediction 4 : glioma, Score: 0.10083294473588467%

Prediction 1 : notumor, Score: 99.93513822555542%
Prediction 2 : meningioma, Score: 0.05890784668673616%
Prediction 3 : glioma, Score: 0.005747644536313601%
Prediction 4 : pituitary, Score: 0.00021990547338646138%
```

Fig. 3: Prediction Scores for First Model

2.2 Second Training: Resnet152 / AdamW Optimizer / NLL Loss

After 50 epochs with AdamW Optimizer and Negative Likelihood Loss, validation accuracy became around 95%. Max validation accuracy rate (94.59) was reached at 49th epoch. The loss and accuracy values can be seen at Figure 4. We can say that there is no big difference between performances of AdamW and Adam Optimizers.

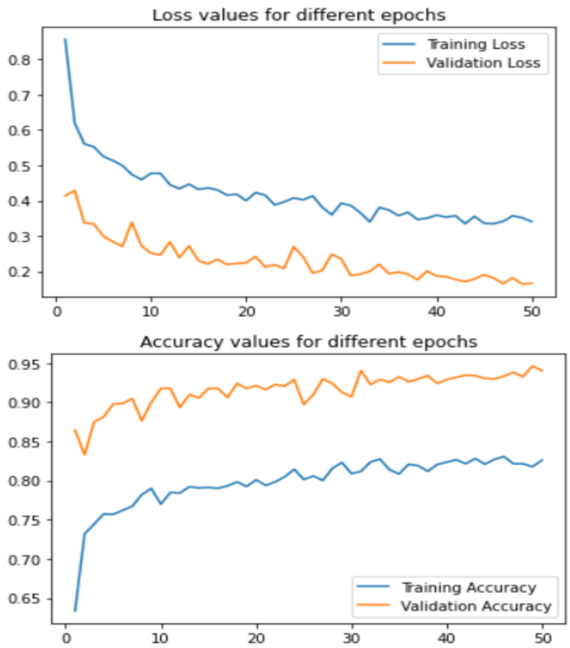

Fig. 4: Results of Second Training

The accuracy rate of the 50 epoch trained model for test dataset is 93.71%. Prediction scores can be seen from Figure 5.

2.3 Third Training: Resnet152 / AdamW Optimizer / CrossEntropy Loss

After 50 epochs with AdamW Optimizer and CrossEntropy Loss, validation accuracy became around 94%. Max validation accuracy rate (94.1) was reached at 50th epoch. The loss and accuracy values can be seen at Figure 5. We can say that there is no big difference between performances of CrossEntropy Loss and Negative Likelihood Loss Functions.

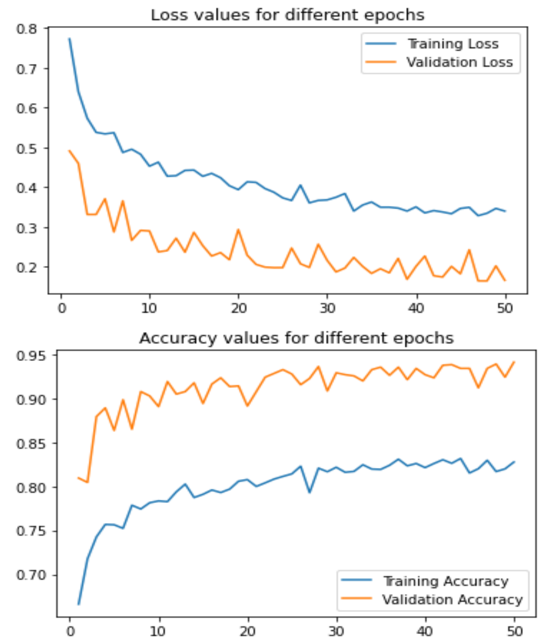

Fig. 5: Results of Third Training

The accuracy rate of the 50 epoch trained model for test dataset is 93.57%.

2.4 Fourth Training: DenseNet161 / AdamW Optimizer / NLL Loss

After 50 epochs with AdamW Optimizer and NLL Loss, validation accuracy became around 96%. Max validation accuracy rate (96.23) was reached at 49th epoch. The loss and accuracy values can be seen at Figure 6. We can say that the performance of DenseNet161 model is better than Resnet152.

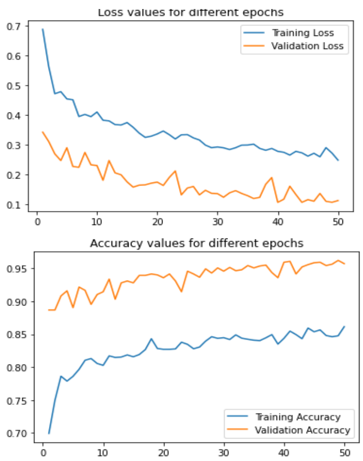

Fig. 6: Results of Fourth Training

The accuracy rate of the 50 epoch trained model for the testing dataset is 95.71%. Prediction scores of that model for the images at Figure 2 can be seen at Figure 7.

```
Prediction 1 : pituitary, Score: 83.00074934959412%
Prediction 2 : meningioma, Score: 16.82167947292328%
Prediction 3 : glioma, Score: 0.146414409391582%
Prediction 4 : notumor, Score: 0.031159963691607118%

Prediction 1 : meningioma, Score: 98.99172782897949%
Prediction 2 : notumor, Score: 0.9170649573206902%
Prediction 3 : glioma, Score: 0.066582887666299994%
Prediction 4 : pituitary, Score: 0.024625263176858425%

Prediction 1 : notumor, Score: 99.99998807907104%
Prediction 2 : glioma, Score: 1.1666584498470911e-05%
Prediction 3 : meningioma, Score: 1.8491654429908522e-06%
Prediction 4 : pituitary, Score: 8.472977874873777e-07%
```

Fig. 7: Prediction Scores for Fourth Model

2.5 Fifth Training: EfficientNetB4 / AdamW Optimizer / NLL Loss

After 50 epochs with AdamW Optimizer and NLL Loss, validation accuracy became around 95%. Max validation accuracy rate (95.16) was reached at 34^{th} epoch. The loss and accuracy values can be seen at Figure 8.

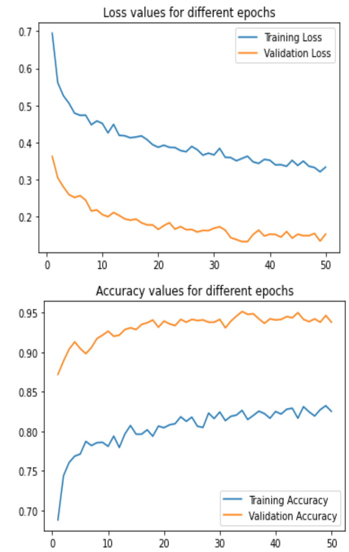

Fig. 8: Results of Fifth Training

The accuracy rate of the 50 epoch trained model for the testing dataset is 95.28%.

Table 1: Comparison of Pre-trained CNN Models with Accuracy Rates as Performance Metrics

Model	Accuracy of Test Dataset
ResNet 152 + Adam Optimizer + NLL Loss	93.57 %
ResNet 152 + AdamW Optimizer + NLL Loss	93.71 %
ResNet 152 + AdamW Optimizer + CrossEntropy Loss	93.57 %
DenseNet161 + AdamW Optimizer + NLL Loss	95.71 %
EfficientNetB4 + AdamW Optimizer + NLL Loss	95.28

2.6 Comment on the Results of CNN Models

After first three trainings (ResNet152 with Adam/AdamW optimizers and CrossEntropy/NLL loss functions), We saw that there was no significant difference between Adam AdamW optimizers; between NLL Loss CrossEntropy Loss functions in terms of providing more accuracy. Therefore, with other CNN models in fourth and fifth trainings, we used AdamW Optimizer and NLL Loss function.

According to the table above, 1^{st} training (ResNet152) gives smallest accuracy, whereas, 4^{th} one (DenseNet161) gives biggest one. This evaluation were compatible with the inference results (prediction scores as confidence levels) in Fig. 3 and Fig 7. At the same time, 5^{th} training accuracy (EfficientNetB4) was also successful and close to the accuracy of the best one.

3. Dataset Analysis in terms of Perspectives of MRI Images

By taking a closer look, it can be realized that the dataset consists of three different views/ perspectives of the brain: **Axial**, **Coronal** and **Sagittal**. The first one is a horizontal plane dividing the body into superior (upper) and inferior (lower) sections. The second one is a longitudinal plane dividing the body into anterior (front) and posterior (back) sections. The last one is a longitudinal plane, dividing the body into right and left parts. The figure below shows the difference of these perspectives for Pituitary Tumor Class.

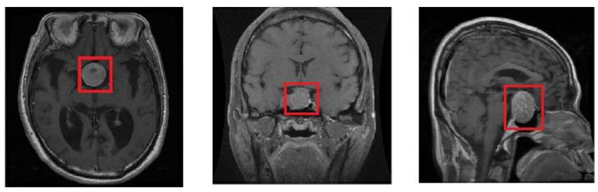

Fig. 9: Axial/ Coronal/ Sagital MRI Images for Pituitary Class

To increase efficiency, we differentiate MRI images based on these views and split the dataset into 3 different datasets. The table below depicts the distribution of these views in our 4 classes. As we can see, there is an imbalance due to no tumor dataset. There is not enough image of no tumor dataset for sagittal and coronal perspectives while there is a redundancy of axial perspective images of no tumor dataset.

Table 2: Comparison of Classes of The Dataset In terms of The Number of Photos from Different Perspectives

	Axial	Sagittal	Coronal	Total
Glioma	594	452	575	1621
Meningioma	567	558	520	1645
No Tumor	1864	97	39	2000
Pituitary	552	610	596	1758
Total	3577	1717	1730	7024

Not to affect this imbalance during training, we did our trainings with axial photos only after reducing some redundant notumor - axial images.

4. Classification with Convolutional Neural Networks with Modified Dataset

Because EfficientNetB4 and DenseNet161 reached good accuracy levels, we used these again with modified dataset. The confusion matrices obtained after 50 epoches can be seen below.

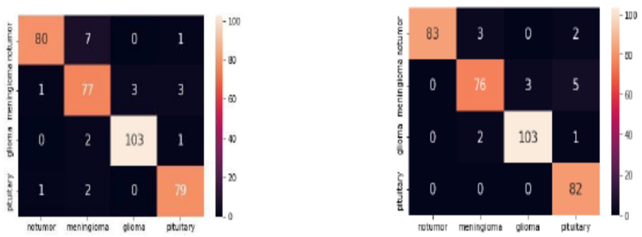

Fig. 10: Confusion Matrices of 50 Epoch Trainings of EfficientNetB4 and DenseNet161 Models

5. Classification with Vision Transformer Method

Another option for this classification duty is Vision Trasformers (ViT). Vision Transformers take images and divide these into patches. Then, these patches with their location (in the original images) information pass through a network. Output gives the classes of the input images. We used the specific ViT model: ***google/vit-base-patch16-224-in21k*** from Hugging Face platform [6]. This model takes 224x224 resolution images; then it creates and use 16x16 resolution patches of these images [6]. Necessary codes for this duty with this ViT model was written with the help of another project for classification of endoscopy images [7].

After 20 epoches of the ViT training of modified dataset, we obtained 98.73% accuracy and the confusion matrix below:

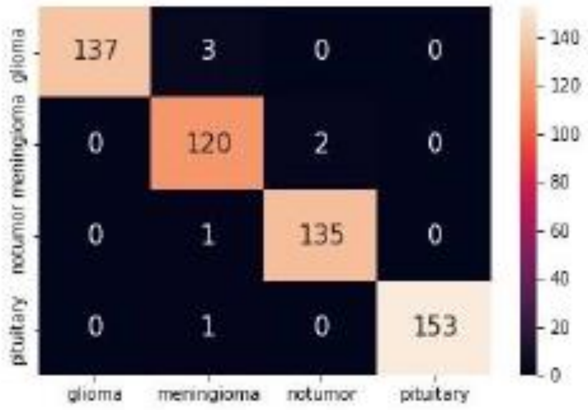

Fig. 11: Confusion Matricex of 20 Epoch Training of the Vision Transformer Model

6. Comparison of Results and Conclusion

Considering and using the confusion matrices in Fig. 10 and 11, we reached the performance metrics below:

Table 3: Comparison of Precision Results of Different Models for Different Classes

	No Tumor	Meningioma	Glioma	Pituitary
ViT	97.83	94.49	97.86	99.35
Efficient NetB4	88.89	81.05	94.50	90.80
DenseN et161	94.32	85.39	94.50	91.11

Table 4: Comparison of Recall Results of Different Models for Different Classes

	No Tumor	Meningioma	Glioma	Pituitary
ViT	99.46	98.73	99.46	99.82
Efficient NetB4	97.22	95	98.33	97.78
DenseN et161	98.61	95.83	98.33	97.78

Table 5: Comparison of Sensitivity Results of Different Models for Different Classes

	No Tumor	Meningioma	Glioma	Pituitary
ViT	98.54	96	100	100
Efficient NetB4	97.56	87.50	97.17	94.05
DenseN et161	100	93.83	97.17	91.11

In terms of accuracy, the Vision Transformer Model gave best accuracy value (98.73%). The tables above (Table 3, 4 and 5) also show the success of the Vision Transformer Model for every classes in terms of precision, recall and sensitivity.

Briefly, it can be said that in this project we examined some popular CNN models with different optimizers/loss functions. Then, we separated the images in the dataset considering their perspectives and we modified the dataset. At the end, we compare the testing performances of two successful CNN models and a ViT model. We realized that the ViT technique is very successful for this image classification duty.

REFERENCES

1. M. Nickparvar, "Brain Tumor MRI Dataset," Kaggle, 24-Sep-2021. [Online]. Available: https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset?select=Training.

2. "Brain tumors," American Association of Neurological Surgeons. [Online]. Available: https://www.aans.org/en/Patients/Neurosurgical-Conditions-and-Treatments/Brain-Tumors.

3. "Models and Pre-Trained Weights," Models and pre-trained weights - Torchvision 0.12 documentation. [Online]. Available: https://pytorch.org/vision/stable/models.html.

4. S. R. R. Rath, "Brain MRI Classification using PyTorch EfficientNetB0," DebuggerCafe, 23-Jan-2022. [Online]. Available: https://debuggercafe.com/brain-mri-classification-using-pytorch-efficientnetb0/.

5. S. James, "Building an Image Classifier with PyTorch," Hubofcode, 14-Mar-2020. [Online]. Available: https://hubofco.de/machinelearning/2020/03/14/Image-classification-with-Pytorch/.

6. "Google/VIT-base-patch16-224-in21k · hugging face," google/vit-base-patch16-224-in21k · Hugging Face. [Online]. Available: https://huggingface.co/google/vit-base-patch16-224-in21k.

7. Y. Labrak, "How to Train a Custom Vision Transformer (VIT) Image Classifier to Help Endoscopists in Less than 5 min," Medium, 06-Sep-2021. [Online]. Available: https://medium.com/@yanis.labrak/how-to-train-a-custom-vision-transformer-vit-image-classifier-to-help-endoscopists-in-under-5-min-2e7e4110a353.