Analyzing the Effects of Non–Generative Augmentation on MRI–Based Classification of Brain Tumors using Convolutional Neural Networks
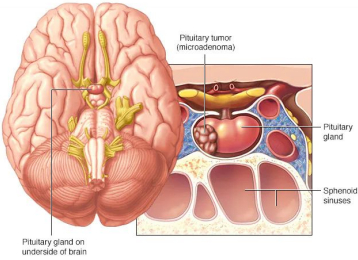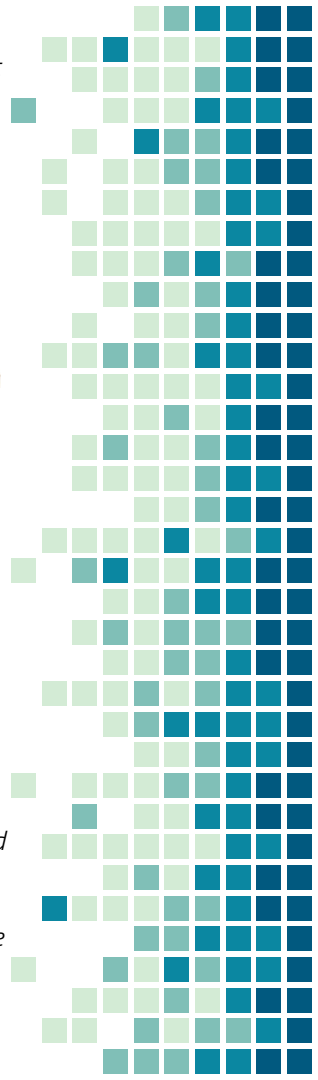
Adam Rolander
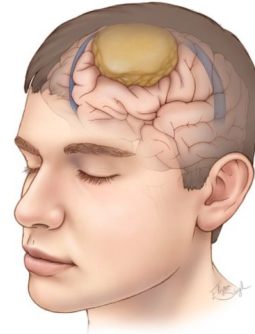
INTRODUCTION I

Brain Tumors

- Three of the four most common brain tumors that originate within the brain are **Gliomas**, **Meningiomas**, and **Pituitary Tumors**.

Meningiomas (right) and Pituitary Tumors (left) are typically benign, but can cause other medical issues depending on their location.

Gliomas (above) are aggressive, malignant tumors, and are the most common brain tumor in adults.

- Due to the devastating impacts that brain tumors can have on one's health, early classification is crucial in reducing mortality and developing a treatment plan.

- Current methods of brain tumor diagnosis typically rely on a surgical biopsy.

- This process is cost-inefficient, time-consuming, and highly invasive, which presents the need for more effective methods of brain tumor classification.

INTRODUCTION II

Convolutional Neural Networks (CNNs)

- Due to the physiological changes that brain tumors cause, they can often be identified with medical imaging techniques, like **Magnetic Resonance Imaging (MRI)**.

      - This identification process can then be automated with Deep Learning, which I sought to do.

- The most commonly used Deep Learning technique for image classification tasks is the **CNN**, an algorithm that employs **feature learning** and **feed forward networks** to classify images. CNN architectures are modeled after the Visual Cortex in the human brain, and they attempt to replicate the process of biological image recognition. A generic CNN Architecture is shown below.

- CNNs have proven to be beneficial because they do not require **image preprocessing** or **feature engineering** (e.g. image segmentation). This results in greater efficiency and applicability to more fields.
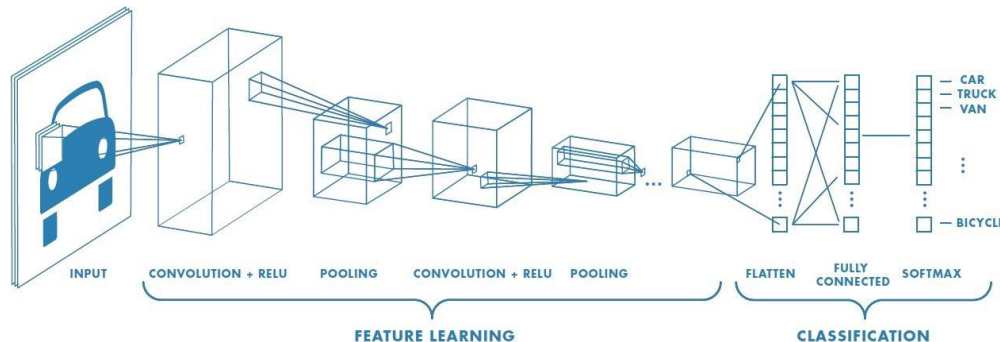
Image Taken From: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

INTRODUCTION III

Non-Generative Data Augmentation & Research Objectives

- CNNs are limited by their need for large training datasets.

- This poses a challenge for medical classification tasks, which typically utilize small datasets due to the rigorous standards for medical data acquisition.

- This issue presents the need for **image data augmentation**, which expands and diversifies a training dataset by altering the original data.

- By expanding a model's training dataset, data augmentation can help prevent the model from **overfitting** during training, which is often correlated with greater image classification capabilities.

Research Questions

1. How successful is my custom CNN architecture at classifying brain tumors?

2. How does the implementation of data augmentation affect my model's success?

Hypothesis

The CNN model utilizing the most data augmentation methods will obtain greater brain tumor classification accuracy than models using the same architecture, preprocessing techniques, and number of training images that do not employ as many augmentation methods.

Objectives: In this study, I wanted to develop a **CNN model** to classify **gliomas, meningiomas, pituitary tumors, and no tumors** from MRIs. I then wanted to test and analyze how different methods of **data augmentation** affected the model's capabilities when used individually and in combination with each other.
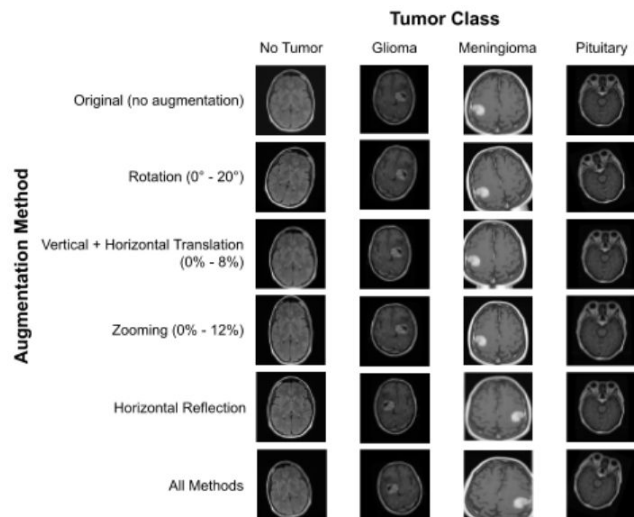
METHODS I

Data Acquisition & Image Preprocessing

- The imaging data used in this study consists of **7023 MRIs** collected from the Kaggle Database and from Nanfang Hospital, Guangzhou, China and General Hospital, Tianjin Medical University, China.

- It consists of **1621 glioma** images, **1645 meningioma** images, **2000 no-tumor** images, and **1757 pituitary tumor** images.

- The MRIs were collected in three planes - coronal, axial, and sagittal.

- Preprocessing: Images were resized to 224 x 224 px, rescaled, and split into training and testing datasets (containing 80% and 20% of the data, respectively).
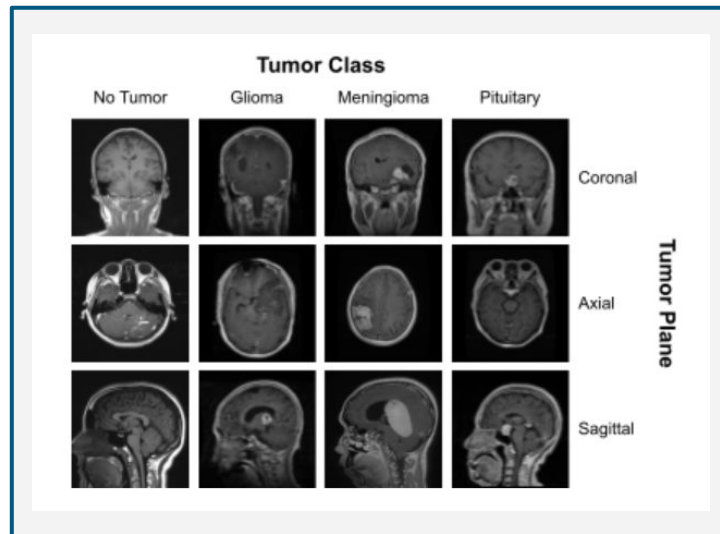
Figure 1. MRIs of No Tumors, Gliomas, Meningiomas, and Pituitary Tumors in different planes. Adapted from [9]. Created by Adam Rolander.

METHODS II

Figure 2. Augmented MRIs of No Tumors, Gliomas, Meningiomas, and Pituitary Tumors in the axial plane. Created by Adam Rolander.

Non-Generative Image Data Augmentation

The following methods of data augmentation were applied to training images for each trial in my study.

(The parameters for each method were chosen based on successful implementations from studies in my Literature Review).

1. **Random Rotation -** Either clockwise or counter-clockwise, by between 0 and 20 degrees.
2. **Random Horizontal Translation -** Either left or right, by between 0% and 8% of the image's total width.
3. **Random Vertical Translation -** Either up or down, by between 0% and 8% of the image's total height.
4. **Random Zoom -** Either in or out, by between 0% and 12%.
5. **Random Horizontal Reflection -** Over the image's y-axis, occurring for 50% of training images (theoretically).

METHODS III

Custom CNN Architecture

- The **Dense Layers Block (DLB)** has a **Flattening Layer** & three **Fully Connected (FC) Layers** (activation function - RELU).
- The first FC layer is followed by a **Dropout Layer** (dropout - 20%)
- The third FC layer (activation function - **Softmax**) is determine a classification for the original image.

- 2D MRIs from the training dataset are provided to the CNN as input in **batch sizes of 32**.

Repeat for 3 Iterations

224

224

Input Layer

Convolutional & RELU

Pooling & Batch Normalization

Feature Learning

Flatten

FC + RELU + Dropout (0.2)

: 128

FC + RELU

: 64

FC + Softmax

Concatenated Dense Layers

Meningioma
Glioma
No Tumor
Pituitary

Classification

Created by Adam Rolander

- The **Feature Learning Block (FLB)** has **Convolutional Layers** (kernel size - (3, 3); strides - (1, 1); padding - valid; activation function - RELU) and **Max Pooling Layers** (pool size - (2, 2); strides - (2, 2); padding - valid).

- After each Convolutional Layer in the FLB, Max Pooling is used to subsample the previous layer. Each image is passed through the FLB three times.

- The **Classification Output** returns a value of 0, 1, 2, or 3 for its predicted classification.

METHODS IV

CNN Implementation & Evaluation

- My CNN Architecture was tested and evaluated over six (6) trials.
- The first used **no** augmentation, the second only used **rotational** augmentation, the third only used **horizontal** and **vertical translations**, the fourth only used **zooming** augmentation, the fifth only used **horizontal reflections** as augmentation, and the final trial used **all** of the discussed methods of augmentation.
- In each trial, the CNN model was trained and tested on the same number of images, regardless of augmentation.

- For each trial, my CNN model was implemented using the **Adam optimizer** (learning rate - 1e-3; *No relation*) and was trained over **20 epochs** with a **batch size of 32**. Training and testing occurred in Google Colaboratory.

$$Accuracy = \frac{Correct\ Predictions}{Total\ Images}$$

$$Precision = \frac{Correct\ Class\ Predictions}{Total\ Class\ Predictions}$$

$$Recall = \frac{Correct\ Class\ Predictions}{Total\ Images\ in\ Class}$$

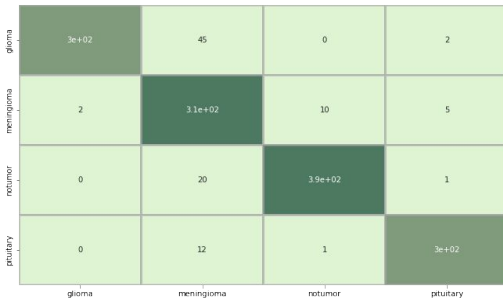$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Figure 3. Formulas for CNN Metrics. Created by Adam Rolander.

- The CNN Models were evaluated according to **Accuracy**, **Precision**, **Recall**, and **F1 Score**
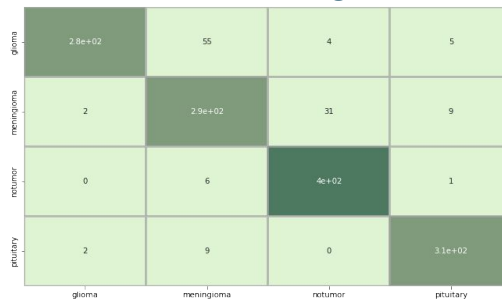
RESULTS I

Confusion Matrices For CNN Trials

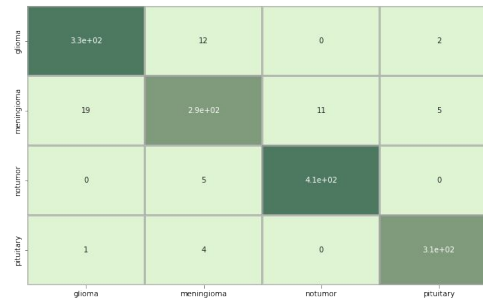### Trial 1 - No Augmentation

Figure 4. Confusion Matrix for No Augmentation Trial. Created by Adam Rolander.

Trial 2 - Rotational Augmentation

Figure 5. Confusion Matrix for Rotational Augmentation Trial. Created by Adam Rolander.

Trial 3 - Translational Augmentation

Figure 6. Confusion Matrix for Translational Augmentation Trial. Created by Adam Rolander.

Trial 4 - Zooming Augmentation
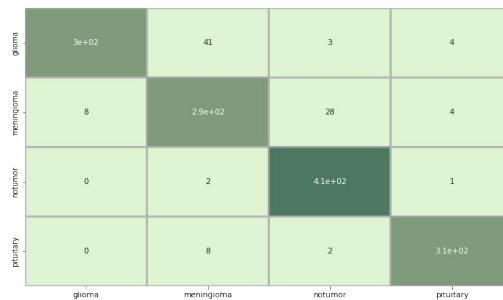
Figure 7. Confusion Matrix for Zooming Augmentation Trial. Created by Adam Rolander.

Trial 5 - Horizontal Reflection Augmentation
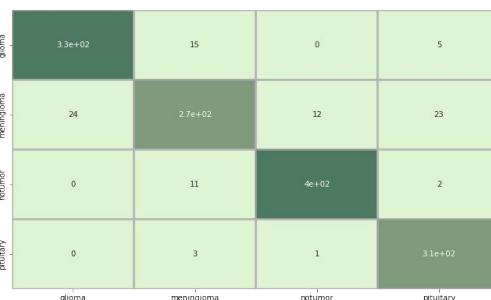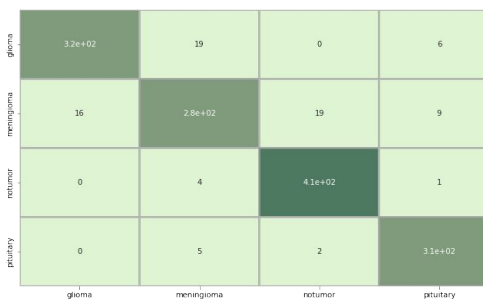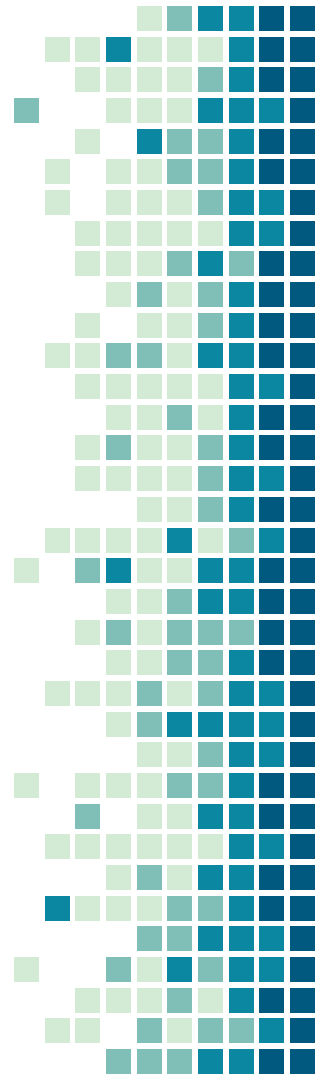

Figure 8. Confusion Matrix for Horizontal Reflection Augmentation Trial. Created by Adam Rolander.

Trial 6 - Combined Augmentation

Figure 9. Confusion Matrix for Combined Augmentation Trial. Created by Adam Rolander.

RESULTS II

Augmentation Effectiveness Summary

All created by Adam Rolander

1 - Translational Augmentation

Tumor Class	Recall (%)	Precision (%)	F1-Score (%)	Accuracy (%)
Glioma	95.97	94.35	95.15	-
Meningioma	89.36	93.33	91.30	-
No Tumor	98.78	97.36	98.06	-
Pituitary	98.42	97.81	98.11	-
Average	95.63	95.71	95.66	95.80

2 - Combined Augmentation

Tumor Class	Recall (%)	Precision (%)	F1-Score (%)	Accuracy (%)
Glioma	92.82	95.28	94.03	-
Meningioma	86.63	91.05	88.79	-
No Tumor	98.78	95.08	96.89	-
Pituitary	97.79	95.09	96.42	-
Average	94.01	94.13	94.03	94.23

3 - Horizontal Reflection Augmentation

Tumor Class	Recall (%)	Precision (%)	F1-Score (%)	Accuracy (%)
Glioma	94.25	93.18	93.71	-
Meningioma	82.07	90.30	85.99	-
No Tumor	96.84	96.84	96.84	-
Pituitary	98.74	91.25	94.85	-
Average	92.98	92.89	92.85	93.17

4 - No Augmentation

Tumor Class	Recall (%)	Precision (%)	F1-Score (%)	Accuracy (%)
Glioma	86.49	99.34	92.47	-
Meningioma	94.83	80.21	86.91	-
No Tumor	94.89	97.20	96.06	-
Pituitary	95.90	97.44	96.66	-
Average	93.03	93.55	93.03	93.02

5 - Zooming Augmentation

Tumor Class	Recall (%)	Precision (%)	F1-Score (%)	Accuracy (%)
Glioma	86.21	97.40	91.46	-
Meningioma	87.84	85.00	86.38	-
No Tumor	99.27	92.52	95.78	-
Pituitary	96.85	97.15	97.00	-
Average	92.54	93.02	92.66	92.81

6 - Rotational Augmentation

Tumor Class	Recall (%)	Precision (%)	F1-Score (%)	Accuracy (%)
Glioma	81.61	98.61	89.31	-
Meningioma	87.23	80.39	83.67	-
No Tumor	98.30	92.03	95.06	-
Pituitary	96.53	95.33	95.93	-
Average	90.92	91.59	90.99	91.17

DISCUSSION & CONCLUSIONS

- The **no-augmentation** control trial achieved a classification **accuracy of 93.02%**.
- When utilizing **rotational augmentation**, the model achieved an overall classification accuracy of **91.17%**, and the model with **zooming augmentation** obtained an accuracy of **92.81%**, suggesting that rotational augmentation and zooming hindered the model's abilities.
- When using **horizontal reflections** as augmentation, my model improved slightly in overall accuracy to **93.17%**, suggesting that horizontal reflections did not have a significant impact.
- The most successful results were achieved with the implementation of **translational augmentation** by itself with an accuracy of **95.80%**, as well as the **combined model** with an accuracy of **94.23%**.

- The results I obtained **partially support my hypothesis** that the CNN model using the most augmentation methods would be most successful at classifying MRIs.
- My hypothesis was correct because the combined model showed a significant improvement in classification capability for each tumor class compared to the no-augmentation model. However, the model using **only horizontal and vertical translations** for augmentation was most **successful**.
- This was unexpected, as data augmentation is most commonly used as a combination of some or all of the methods I tested, but my results show that it was more effective to only use translations.

REFERENCES & FUTURE STEPS

References

1. A neurosurgeon's overview the brain's anatomy
https://www.aans.org/Patients/Neurosurgical-Conditions-and-Treatments/Anatomy-of-the-Brain (accessed Jul 16, 2021).
2. Metastatic Brain Cancer https://www.cancercenter.com/cancer-types/brain-cancer/types/metastatic-brain-cancer (accessed Jul 16, 2021).
3. Benign https://medlineplus.gov/ency/article/002236.htm (accessed Jul 16, 2021).
4. Brain Tumors https://www.aans.org/en/Patients/Neurosurgical-Conditions-and-Treatments/Brain-Tumors (accessed Jul 16, 2021).
5. Cancer https://www.who.int/news-room/fact-sheets/detail/cancer (accessed Jul 16, 2021).
6. Brain Tumor - Diagnosis https://www.cancer.net/cancer-types/brain-tumor/diagnosis (accessed Jul 16, 2021).
7. Cheng, J.; Huang, W.; Cao, S.; Yang, R.; Yang, W.; Yun, Z.; Wang, Z.; Feng, Q. Enhanced Performance of Brain Tumor Classification via Tumor Region Augmentation and Partition. *PLoS One* 2015, *10* (10), e0140381.
8. Díaz-Pernas, F. J.; Martínez-Zarzuela, M.; Antón-Rodríguez, M.; González-Ortega, D. A Deep Learning Approach for Brain Tumor Classification and Segmentation Using a Multiscale Convolutional Neural Network. *Healthcare (Basel)* 2021, *9* (2), 153.
9. Badža, M. M.; Barjaktarović, M. Č. Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network. *Appl. Sci. (Basel)* 2020, *10* (6), 1999.
10. Antoniou, A.; Storkey, A.; Edwards, H. Data Augmentation Generative Adversarial Networks. *arXiv [stat.ML]*, 2017.
11. Brownlee, J. How to configure image data augmentation in keras
https://machinelearningmastery.com/how-to-configure-image-data-augmentation-when-training-deep-learning-neural-networks/ (accessed Jul 16, 2021).
12. Nickparvar, M. Brain Tumor MRI Dataset, 2021 (accessed Jul 16, 2021).
13. Kang, J.; Ullah, Z.; Gwak, J. MRI-Based Brain Tumor Classification Using Ensemble of Deep Features and Machine Learning Classifiers. *Sensors (Basel)* 2021, *21* (6), 2222.
14. Skull-stripping http://brainsuite.org/processing/surfaceextraction/bse/ (accessed Jul 16, 2021).
15. Shmueli, B. Multi-class metrics made simple, part I: Precision and recall
https://towardsdatascience.com/multi-class-metrics-made-simple-part-i-precision-and-recall-9250280bddc2 (accessed Jul 16, 2021).
16. Shmueli, B. Multi-class metrics made simple, part II: The F1-score
https://towardsdatascience.com/multi-class-metrics-made-simple-part-ii-the-f1-score-ebe8b2c2ca1 (accessed Oct 14, 2021).
17. Çinar, A.; Yildirim, M. Detection of Tumors on Brain MRI Images Using the Hybrid Convolutional Neural Network Architecture. *Med. Hypotheses* 2020, *139* (109684), 109684.
18. Saxena, P.; Maheshwari, A.; Tayal, S.; Maheshwari, S. Predictive Modeling of Brain Tumor: A Deep Learning Approach. *arXiv [cs.CV]*, 2019.

Future Steps

Future research should seek to vary and optimize the parameters of the augmentation methods I used. In addition, different augmentation methods should continue to be tested individually on different datasets with different CNN architectures.