DEEP LEARNING CIS4050-N

BRAIN TUMOUR MRI CLASSIFICATION USING DL MODELS

PATEL SNEHALKUMAR RAJESHBHAI(E4202433), CAMPBELL CONNOR(A0031646), WELFORD CHLOE(A0105592), RANGANATHA PALLAVI(S3417639), MILLS-NKETSIAH EKOW (E4041159)

Teesside University

PROJECT BACKGROUND AND AIMS

•Brain tumours pose a global health challenge, with over 300,000 new cases annually, requiring early diagnosis for survival.

•Manual MRI analysis is time-consuming and struggles with variability (e.g., 1.5T vs. 3T scanners), delaying treatment.

•Deep learning enables automated, accurate classification, with transfer learning (e.g., EfficientNet) optimizing performance for clinical use (Vimala, B, 2023).

•**Aim**: Classify brain tumours (glioma, meningioma, pituitary, no tumour) in MRI scans using deep learning models (CNN, ResNet-50, EfficientNetB0, DenseNet-121).

•Focus on achieving high accuracy, computational efficiency, and interpretability via Grad-CAM for clinical deployment.

•Project aims to reduce diagnostic errors and support AI-driven diagnostics in healthcare, particularly in underserved regions.

SOCIAL, ETHICAL AND LEGAL RISKS

- Incorrect interpretation of AI system results by users, leading to legal issues, and further exasperating a disconnect between patient and doctor that such a system could create. To combat, ensure a transparency to all involved about the basic workings of said system, alongside retain clinical thinking & secondary checks when interpreting AI results.
- Over reliance on AI systems leading to reduced critical & clinical thinking skills among users, leading to increased risk on patient health. To combat, stricter regulation & regular refresh of medical licensing (e.g. drug training), alongside maintaining standard practices (e.g. doctor patient consultations, following existing well established routines).
- Tying to legal repercussions, responsibility of those involved, including the event of errors / malpractice, issues in development, & errors caused from implementation phases.
- Risk of bias in AI systems, caused by errors in phases mentioned above, also including: gender & sex (male, female, intersex), persons ethnicity & background (e.g. biological factors & differences in presenting symptoms)
- Inequalities in access to medical services & systems (e.g. 1st vs 3rd world countries, paid vs funded healthcare).
  To combat 3 points above, develop & enforce strong AI principles for use in medicine, alongside balanced training data.
- Data privacy, weak legislation & regulations, from relatively rapid development & implementation, alongside short life-span.

DATASET PRE-PROCESSING AND AUGMENTATION

- Dataset: https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset
- The dataset consists of 7,022 MRI images taken from either the top, side, or back of the patient's head
- Images were pre-classified into 1 of 4 categories: notumor, glioma, meningioma, pituitary
- These classes were distributed well, with only a very small bias toward the 'notumor' category
- Images were already split in to testing and training folders, which had to be combined before processing and augmentation. The data was then split in to training, testing and validation sets with a 70/15/15 split
- All images were scaled to 224x224x3, with smaller images receiving padding to prevent stretching or distortion
- Following this, all images were given an additional 20 pixels of padding to reduce feature loss at the corners of the images (Chaudhary, 2022)
- ImageDataGenerator was used to augment the images, as applying the augmentation manually took a long time. This way also means they were augmented randomly each time, without the need to re-process the images
- Images were only augmented with minor parameters, as too much zooming or flipping would not be realistic. However, minor augmentation is shown to increase the accuracy of models trained on MRIs (Nalepa et al., 2019)

MODEL IMPLEMENTATION

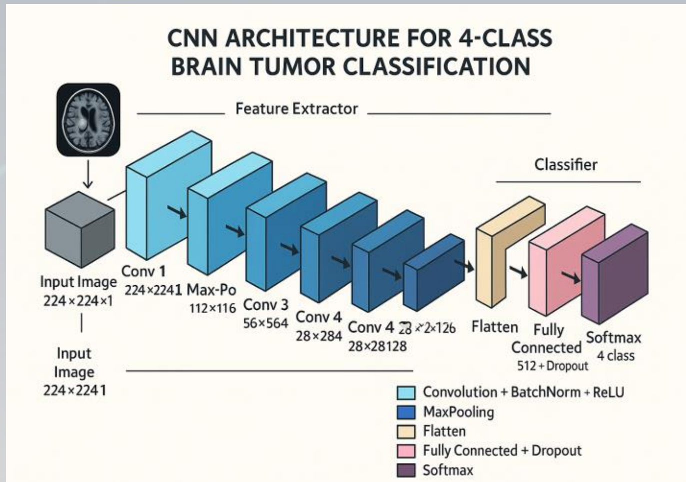
Implemented and evaluated multiple deep learning models for brain tumor classification:

1. CNN Base Model:
- Custom convolutional neural network built from scratch.
- Consisted of convolutional, pooling, dropout, and fully connected layers.
- Tuned with Swish and LeakyReLU activations to improve convergence.

2.Transfer Learning Models:
- ResNet-50, EfficientNetB0, and DenseNet-121 pre-trained Models.
- Top layers were modified for binary classification..
- Fine-tuning included adjusting trainable layers and re-training higher layers while freezing earlier convolutional blocks.

3.Optimization Techniques:
- Adam optimizer with learning rate tuning (1e-4 to 1e-6).
- Batch size increased from 32 to 64 to enhance training efficiency.
- Dropout regularization is applied to prevent overfitting.
- EarlyStopping callback used to halt training.

4.Evaluation and Interpretability:
- Training/validation accuracy and loss curves monitored.
- Grad-CAM visualizations are used to highlight tumor regions, improving model explainability.
- All models were developed and trained using Keras and TensorFlow frameworks in Python.



FINDINGS AND RESULTS

•Fine-tuned transfer learning models outperformed

•the custom CNN for brain tumour MRI classification.

•**EfficientNetB0**: Best balance with 99.05% accuracy, 98.99% precision, 99.00% recall, 99.00% F1-score.

•**DenseNet-121**: Highest validation accuracy (99.05%), But longer training times.

•**ResNet-50**: 98.86% accuracy;

•**Custom CNN**: 98.01% accuracy post fine-tuning.

Model	Accuracy	Precision (Avg)	Recall (Avg)	F1-Score (Avg)
ResNet50	98.86%	98.81%	98.80%	98.80%
EfficientNetB0	99.05%	98.99%	99.00%	99.00%
DenseNet	99.05%	99.00%	99.00%	98.99%
Custom CNN	98.01%	97.96%	97.92%	97.93%

•Fine-tuning strategies improved performance such as switching activation functions (Swish, LeakyReLU), and retraining.

•Increased batch size from 32 to 64, Applied dropout regularization and EarlyStopping to prevent overfitting.

•Evaluation metrics: accuracy, precision, recall, F1-score.

CONCLUSION AND FUTURE WORKS

This project demonstrates the effectiveness of deep learning, particularly transfer learning, in accurately classifying brain tumors from MRI images. Among the models tested, *EfficientNetB0* offered the best trade-off between accuracy and efficiency, while *DenseNet-121* achieved the highest validation accuracy at the cost of increased training time. Fine-tuning techniques, including modified activation functions, increased batch size, dropout regularization, and early stopping, significantly enhanced model generalization and stability. Additionally, Grad-CAM visualizations confirmed the models' ability to focus on relevant tumor regions, an essential factor for clinical trust and interpretability.

Future work must focus on extending the classification to multi-class tumor detection with clinical labels, integrating 3D MRI data for richer spatial context, and deploying the models in a real-time diagnostic tool for clinical settings.

REFERENCES

➢Vimala, B.B., Srinivasan, S., Mathivanan, S.K., Mahalakshmi, P., Jayagopal, P. and Dalu, G.T., 2023. Detection and classification of brain tumor using hybrid deep learning models. https://www.nature.com/articles/s41598-023-50505-6.

➢Chaudhary, V. (2022) *What is Padding in Neural Network?*, *GeeksforGeeks*. Available at: https://www.geeksforgeeks.org/what-is-padding-in-neural-network/ (Accessed: 21 April 2025).

➢Nalepa, J., Marcinkiewicz, M. and Kawulok, M. (2019) 'Data augmentation for brain-tumor segmentation: A Review', *Frontiers in Computational Neuroscience*, 13. doi:10.3389/fncom.2019.00083.

➢Sheliemina, N., 2024. The use of artificial intelligence in medical diagnostics: Opportunities, prospects and risks. *Health Economics and Management Review*, 5(2), pp.104-124.

➢Kaggle link of Project: https://www.kaggle.com/code/connorc/deep-learning-ica-brain-tumour-classification