****

**FACULTY:** Applied and Computer Sciences

**SUBJECT:** Business Analysis 3.2

**SUBJECT CODE:** AIBUY3A

**LECTURER:** Mr. Malefetsane Matsela

**YEAR:** 2025

|  |  |
| --- | --- |
| **ASSESSMENT DATE:** | 16TH OCTOBER, 2025 |

|  |  |  |
| --- | --- | --- |
| **STUDENT NUMBER** | **SURNAME AND INITIALS** | **SIGNATURE** |
| 223515760 | CHIGBU CCI |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

**DECLARATION**

We, declare that:

* The contents of this project represent our own unaided work, and that the project has not previously been submitted for academic examination towards any qualification.
* We have not engaged in plagiarism and use of unauthorized materials within our work.
* All resources used in this work has been appropriately cited and acknowledged.
* This work represents our own opinions and not necessarily those of the Vaal University of Technology.

|  |  |  |  |
| --- | --- | --- | --- |
| **Student No.** | **Surname & Initials** | **Date** | **Signature** |
| 223515760 | Chigbu CCI |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Table of Contents**

DOCUMENTATION ASPECTS

1. **brAIniac’s Relevance To The Theme**

The diagnosis of brain tumors is one of the most critical tasks in healthcare. Despite its significance, it remains yet still a time consuming activity requiring lots of resources and expertise to move from differentials to diagnosis.

Medical practitioners often rely on manual inspection of MRI scans a process, which is prone to human errors even in developed countries. Our proposed solution leverages the BRISC 2025 dataset – a high quality, vetted dataset containing multiple MRI scans already vetted by experts in radiology and neurology, to build a solution to assist and support healthcare in **Chris Hani Baragwanath Hospital** for tumor detection and classification.

The solution adopts deep learning for image classification to identify the type of brain tumor (Glioma, Meningioma, Pituitary Adenomas or No Tumor) This AI solution also aligns with South Africa’s goal of leveraging the power of AI technologies to create innovative ideas and solutions in the healthcare industry. This will reduce diagnostic workload and provide for faster response from diagnosis to corrective actions for such a life threatening condition.

1. **Business Objectives**

* To reduce time required for diagnosis
* To improve accuracy of brain tumor diagnosis by leveraging BRISC 2025 dataset
* To improve patient’s prognosis by early detection and allowing for efficient clinical decision making
* To reduce healthcare costs for both patients and hospitals without sacrificing quality decision making
* To demonstrate how AI solutions like brAiniac can automate medical processes
* To promote the use and adoption of AI solutions like brAiniac to more remote and low-resource locations in South Africa

1. **Business Success Criteria**

* To achieve more than 80% overall test accuracy when classifying brain MRI scans into the different categories; Glioma, Meningioma, No tumor and Pituitary.
* To process 1000 test images and generate a confusion matrix and a classification report in less than 5 minutes on limited CPU power ( <= 16 GB ram)
* To be able to predict on a single brain MRI scan and classify it while showing a confidence score for each diagnosis in less than 2 seconds
* To achieve more than 80% user adoption from participating radiologists and neurologists from Chris Hani Hospital

1. **Business Background**

With increasing demand for better quality of life, healthcare has been a paramount topic. Due to this necessitated demand, South Africa faces a scarcity of key specialists in (neurology and radiology) as seen in this article from UCT news of 2021 stating that there are 150 neurologists and only about 35 work in public hospitals like Chris Hani Hospital. These shortages together with equipment constraints contribute to MRI backlogs that delay time critical diagnosis as sited in an article from HEALTH-E News stating, “About 3500 patients are currently awaiting MRI scans, underscoring a severe diagnostic backlog”.

We at **QUANTUM.AI** having recognised these constraints the public health sector faces in South Africa, have developed a practical AI decision – support tool that analyses brain MRI scans to detect and classify tumors while reducing workloads for clinicians. By automating parts of the diagnostic process with deep learning, **brAIniac** provides a fast, consistent and diagnostic assessment with confidence scores thus helping medical experts prioritize cases and accelerate time taken to medical interventions.

1. **Requirements**

**Functional requirements (what a system must do):**

* The system will accept brain MRI scans in digital image formats such as JPG
* The system must normalize and pre-process images before analysing
* The system must predict each MRI image into one of the following classes of [glioma, meningioma, no\_tumor, pituitary] as well as return the confidence score
* The system will generate summary reports on test data

**Non-Functional requirements (how well the system must do it):**

* The system will handle at least 30 MRI scans per hour concurrently
* The system will maintain consistent performance across different MRI scan images
* The system will support expansion to other hospitals in the future
* The system must be highly available about 99.99% of the time
* The system will have a user friendly dashboard for non-technical users to navigate easily

1. **Constraints**

* The system analyses brain MRI images only. Other medical imaging such as PET, CT and X-Rays as well as other body regions are out of scope.
* Classification is currently restricted to four classes [glioma, meningioma, no\_tumor, pituitary]
* Model performance is validated only on data similar to BRISC 2025
* Testing of the system will be done on a limited compute workstation (less than 16 GB RAM)

1. **Risks**

|  |  |  |  |
| --- | --- | --- | --- |
| **RISK** | **LIKELIHOOD** | **IMAPCT** | **MITIGATION** |
| Misclassification of tumors | Medium | High | Ensure all images are normalized and sized appropriately  If confidence score is below 80% then probe for further review |
| Over-reliance | Medium | High | Disclaimer for decision- support only  Transparent reporting system along with confidence score prediction to prevent blind trust |
| Privacy data breach | Low | High | Comply with South Africa’s POPIA act and principles of least privileged to collect information about patient data  Clearly define what data is retained  Implementing RBAC when being deployed |
| Poor User Adoption | Medium | High | 2 week pilot operation to teach about the system  Appoint a champion from Chris Hani hospital  Provide a How to use Documentation |
| Technical failure | Medium | High | Host solution on cloud services with high availability incorporated in their SLA as well as other services such as the fault domains and availability zones |

1. **Tools And Techniques**
2. **Programming Language**

* Python – the primary language due to its wide range of AI libraries, large community support for various features and ease of deployment

1. **Machine Learning / Deep Learning techniques**

* Supervised Image classification with Convolutional Neural Networks using Keras Sequential API
* Data Augmentation (zooms, rotations, shifts) on appropriate data
* Normalization of all image data
* Dataset split and loading (90% for training and 10% for validation)

1. **Python Libraries Used**

* os- for robust file/path handling
* random – for reproducible sampling
* numpy – for numerical array
* tensorflow – for input pipelines and random seeding
* matplotlib – for visualizing single image (prediction , confidence and actual class)
* Keras- for model definitions, layers, training prediction and metrics tracking
* sklearn - for evaluation reports such as confusion matrix and classification report

1. **Development and collaboration tools**

* Meetings / Whatsapp – to discuss the project and share ideas and resources
* Git/ GitHub – for version control, code collaboration and project management and task tracking
* Visual Studio Code – IDE for developing brAIniac code base and pushing code fragments to Github

1. **Support tools**

* Grammarly – for formatting documentation
* Microsoft Word – for documenting project
* Canva – for creation of poster

1. **Problem Definition**

Diagnosing brain tumors with MRI scans is a very important, yet time and resource consuming task. In South Africa and other developing areas, there are not enough medical experts to fulfil this task in a timely manner. In public hospitals such as Chris Hani hospital, the problem is even more emphasized with the few experts they are able to spare being over tasked which can lead to mistakes, delays in treating patients and even missed findings.

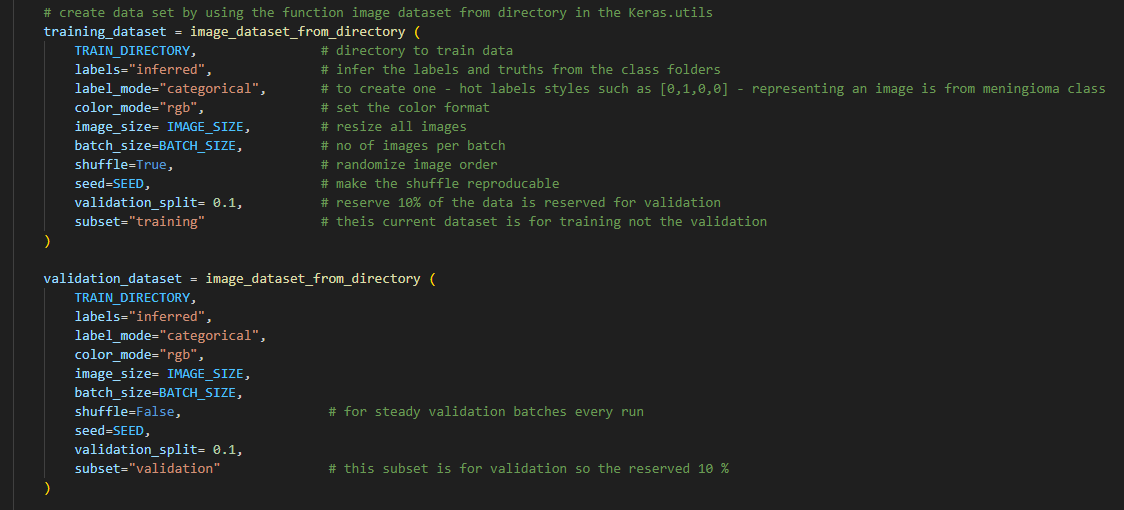
This problem is very critical and should adequately be addressed as brain tumors can result to death or poor prognosis for untimely interventions. Early detection of brain tumors and improved accuracy of categorization of tumor type can greatly improve a patient’s prognosis by bringing timely clinical decision making to them. Our proposed solution, brAIniac uses the BRISC 2025 dataset and AI ML/DL techniques to analyse and classify brain MRI images. This solution will support specialists reduce their workload and improve decision-making.

1. **Poster**

**THEORETICAL ASPECTS**

1. **Machine Learning Approach**

We used supervised multiclass image classification of brain MRI scans grouped into 4 classes [glioma, meningioma, no\_tumor and pituitary]. The labelled data are loaded from respective class folders with a 90/10 split for train/validation data. By using the Keras.utils function of image\_dataset\_from\_directory, we mapped each image to its corresponding class and created a one-hot label for each image.



1. **Data**

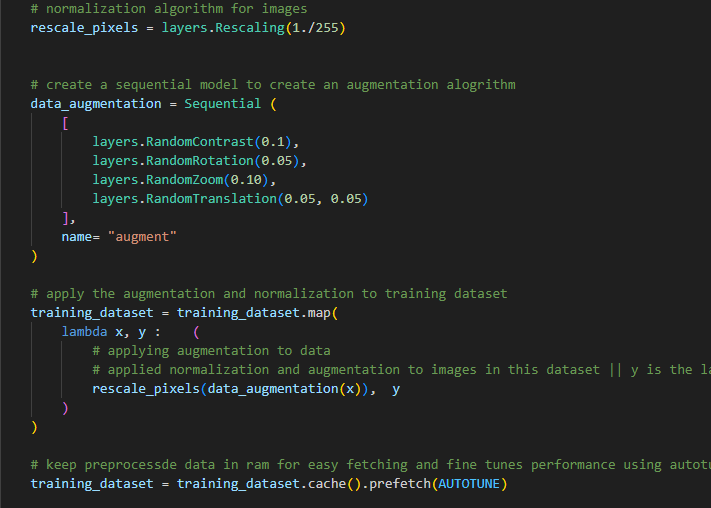
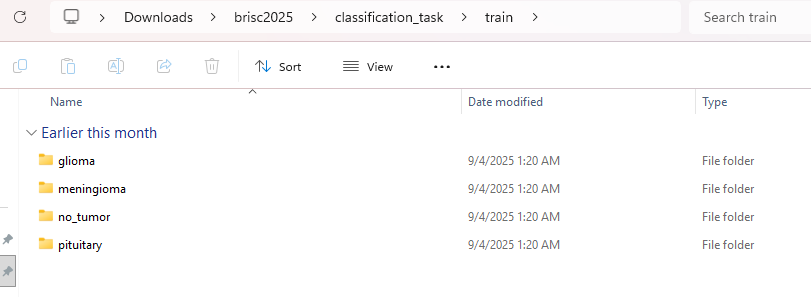
brAIniac employed the use of BRISC 2025 dataset sourced from KAGGLE dataset library and focused on the classification\_task. The folder structure is train/ and test/, each with four class folders: glioma, meningioma, no\_tumor and pituitary.

In total, there are 6000 JPG images (5000 train images and 1000 test images) each sized at 512 X 512 pixels.

In codebase, all images were resized to 224 X 224 RGB to improve efficiency and normalized to [0, 1]. Class labels are inferred from the folder names and encoded categorically.

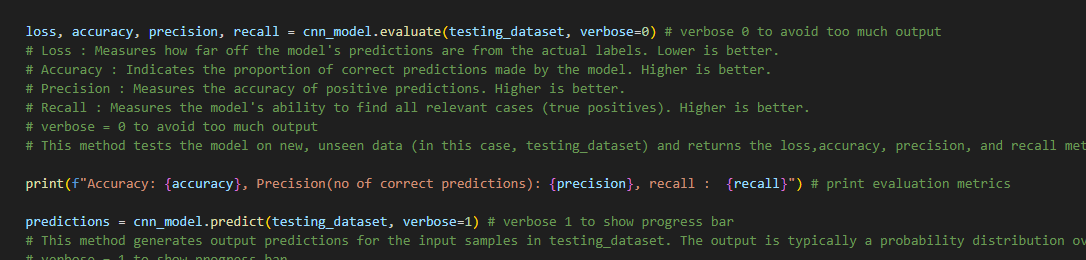
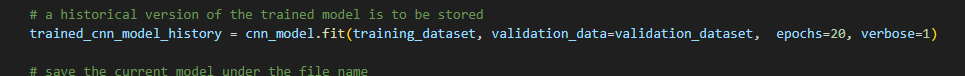
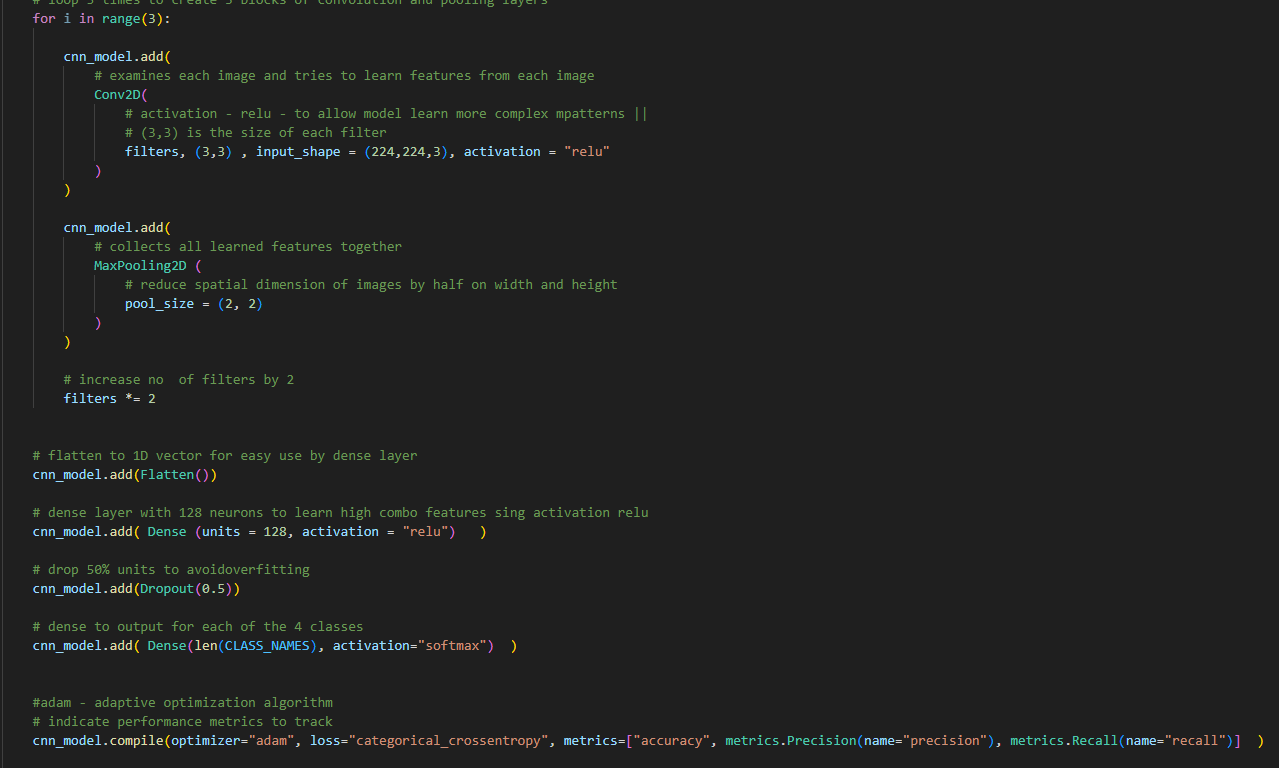
A 10% validation split was created from the training set to monitor training and the separate test set is used for final evaluation.

To improve robustness of training, we applied data augmentation on training set only such as small contrasts, rotations, zooms and translations; validation set and testing set were only normalized



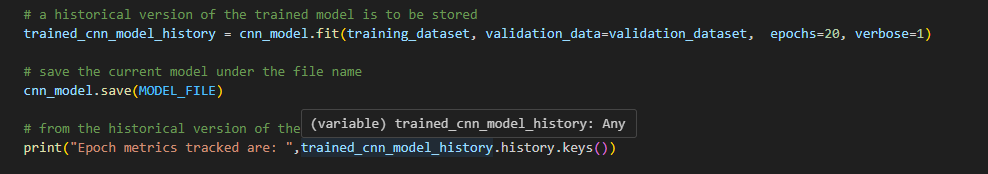
1. Model

brAIniac uses a Convolutional Neural Network with Keras Sequential API. We trained 20 epochs using the Adam optimization and categorical cross-entropy while tracking accuracy, precision and recall. For evaluation, we run the saved trained model on the separate test data and report overall metrics including the confusion matrix and the classification report to see performance per class.



1. **Time Series Analysis On Data**

Although the clinical data are images and not sequences, we still analyse a time series of training signals across epochs. During training, model records per each epoch loss, accuracy and recall for both training and validation. Treating these as a time series, it allows us to spot overfitting (when training increases and validation decreases or stalls), choose sensible stopping points before testing.



1. **Solution Techniques**

* We normalized pixel images by rescaling from the RGB [0, 255] to [0, 1] so the networks trains stably and converges faster
* We augmented only the train data to expose our model to natural variability without changing labels or feature thus improving generalization
* The convolution blocks increase the number of filter in each block from 32 to 64 and to 128 to allow the model learn from simple edges and shapes to complex tumor patterns
* A dropout of 50% was applied to avoid overfitting
* We implemented efficient input pipeline to keep the CPU fed and reducing noise

1. **Natural Language Processing**

No nlp used

1. **Deep Learning**

We implemented a 2D convolutional neural network using Keras sequential API. Images are first resized, normalised and slightly augmented (training data only) before training. Our sequential model allows building layers for examining the images.

Convolution layers for …

Pooling layers for …

We created 3 blocks of the convolution and pooling layers so as to …

The model further has a flatten layer for..

And a dense layer with 128 units (meaning) and activation style relu for …

It also has a dropout layer to avoid overfitting

And another dense layer that leads to the output layer for prediction based on the classes

1. **Other Features**

Predict single image part

REFERENCES

1. Brain mri is critical - <https://www.mdpi.com/2673-4117/6/5/82>
2. Tumor classification is time consuming-
3. SA wants to leverage ai in healthcare - <https://www.health.gov.za/wp-content/uploads/2020/11/national-digital-strategy-for-south-africa-2019-2024-b.pdf>
4. Scarcity of professionals - <https://www.news.uct.ac.za/images/userfiles/downloads/media/2021_11_19_Neurologists.pdf>
5. MRI BACKLOGS - <https://health-e.org.za/2025/06/30/patients-left-in-limbo-as-mri-services-collapse-in-gauteng-mri/?utm_source=chatgpt.com>
6. Brisc 2025 - <https://www.kaggle.com/datasets/briscdataset/brisc2025>