2023

Detection of Brain Tumour MRI Images using Convolutional Neural Networks and Traditional Machine Learning Algorithms

A Project in Machine Learning in Synthetic Biology

Chigozie Chibueze Nkwocha
UNIVERSITY OF TARTU, ESTONIA

Table of Contents

Abstract

Brain tumours are caused by an imbalance in cell growth and death. Their detection and diagnosis are difficult, time-consuming and mostly involve trained medical experts to run medical tests and analyse high-resolution image scans. For early detection and treatment, an automatic and fast system is needed. Various computational methods have been applied to detect and classify brain MRI scan images, and machine learning is one of them. Here, two traditional machine learning models: Logistic regression (LR) and K-nearest neighbours (KNN) as well as a convolutional neural network (CNN) were used to classify 3624 brain MRI scans into four classes: pituitary, glioma, meningioma tumour and healthy brain images. From the results, CNN outperformed logistic regression and K-nearest neighbours, with overall accuracies of 90.8%, 76.8%, 72.2% and F1-scores of 91%, 75.3%, and 71.9%, respectively. From their in-class performance, all three models had difficulty differentiating glioma from meningioma tumours.

Introduction

Brain tumours are the abnormal outgrowth of brain cells (Pedada *et al.*, 2023) that arise from an imbalance in the number of cell growth and death. They can exist as benign or malignant where malignant tumours are cancerous brain tumours that are characterised by uncontrollable cell growth and division evading the whole cell (Kareem *et al.*, 2023).

Their detection and diagnosis are difficult and laborious and often involve the expertise of medical personnel who runs a series of medical tests and analyses image scans from high-resolution devices (Borole and Nimbhore, 2015; Saeedi *et al.*, 2023). As a result, an automatic and fast technique is needed for early detection and treatment. Various advanced methods such as machine and deep learning classification models have been applied to automate the process through the use of raw pixels from image scans. One such model is the Convolutional Neural Networks (CNNs) which can segment image pixels by automatically extracting them (Pedada *et al.*, 2023; Saeedi *et al.*, 2023).

Aim

To classify brain tumour MRI scan images into four classes: *no tumours*, *glioma tumours*, *meningioma tumours* and *pituitary tumours* using traditional and modern machine learning (CNN) models.

Dataset

The dataset was obtained from the research of Saeedi *et al.*, 2023 which was downloaded from Kaggle. It contains 3,624 MRI scans of a healthy brain (500), glioma tumour (926), meningioma tumour (937) and pituitary tumour (901) images. The dataset was slightly imbalanced toward the healthy tumour images.

Methodology

Models

The classification models used were chosen from the models used by Saeedi *et al.*, 2023 with slight modifications. Two traditional machine learning models: K-nearest neighbours (KNN) and Logistic regression (LR), were used.

Data Preparation

Images were rescaled to a 0,1 range and resized to 150 by 150. For the machine learning models, pixels from only grayscale images were extracted and used for classification while RGB images were used for CNNs. Due to the high dimensionality (22,500) of the features and to reduce computational time, the Principal Component Analysis (PCA) feature reduction technique was applied (Jollife and Cadima, 2016).

Model Development and Evaluation

Image data were split into train (90%) and test (10%) sets. For the traditional machine learning models, parameters with optimal performance based on accuracy were obtained via a 5-Fold cross-validation technique. A 5-depth sequential CNN model architecture was used, with max-pooling, dense layers as well as dropout layers to control overfitting.

Image flipping and rotation data augmentation techniques were used to generate new copies of training set images. To solve the class imbalance problem between the healthy tumour images and tumour images, more images of healthy images were generated during data augmentation (for only Logistic regression and K-Nearest Neighbour models).

Model performance was evaluated on the holdout test data based on the accuracy and F1-score (a combination of recall and precision).

Calculations

$$Accuracy = \frac{(TP+TN)}{TP+TN+FP+FN}$$

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

$$Precision = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN}$$

TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative

Results

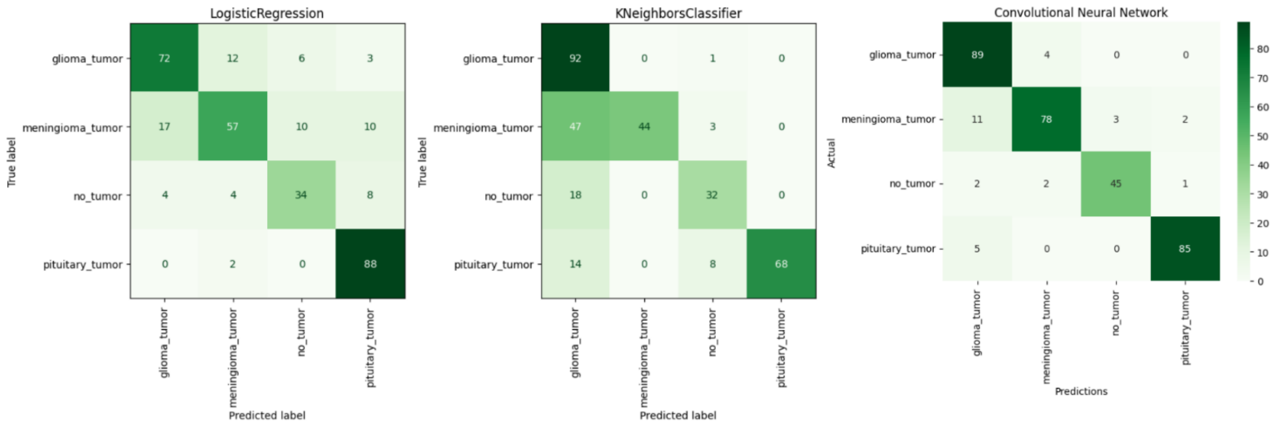

Figure 1: Confusion Matrix for CNN, LR and KNN

Table 1: Model Performance Metric (Overall)

Classification Model	Accuracy	F1-Score
KNearest Neighbours	0.7217	0.7191
Logistic Regression	0.7676	0.7533
Convolutional Neural Network	0.9083	0.9100

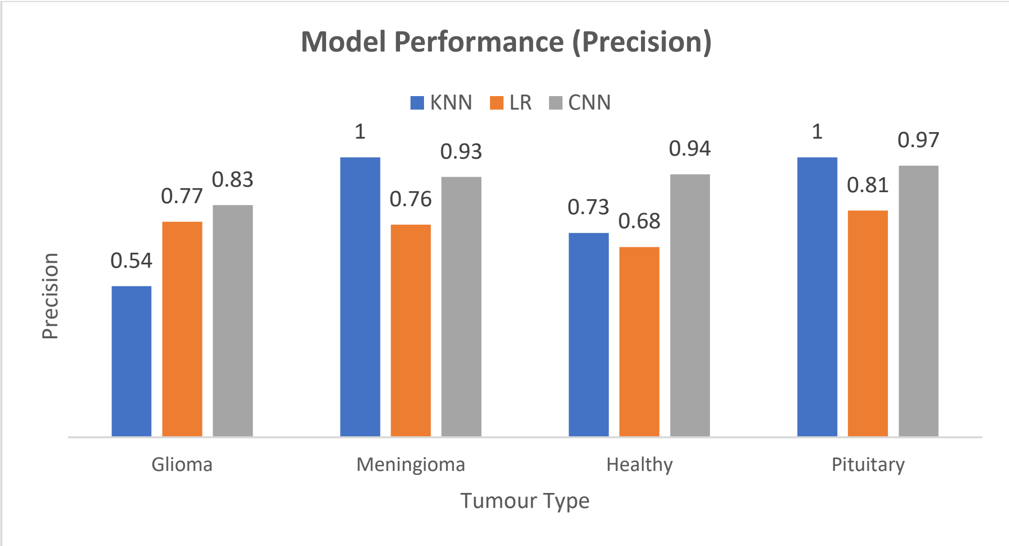

Figure 2: Class-level model performance (precision)

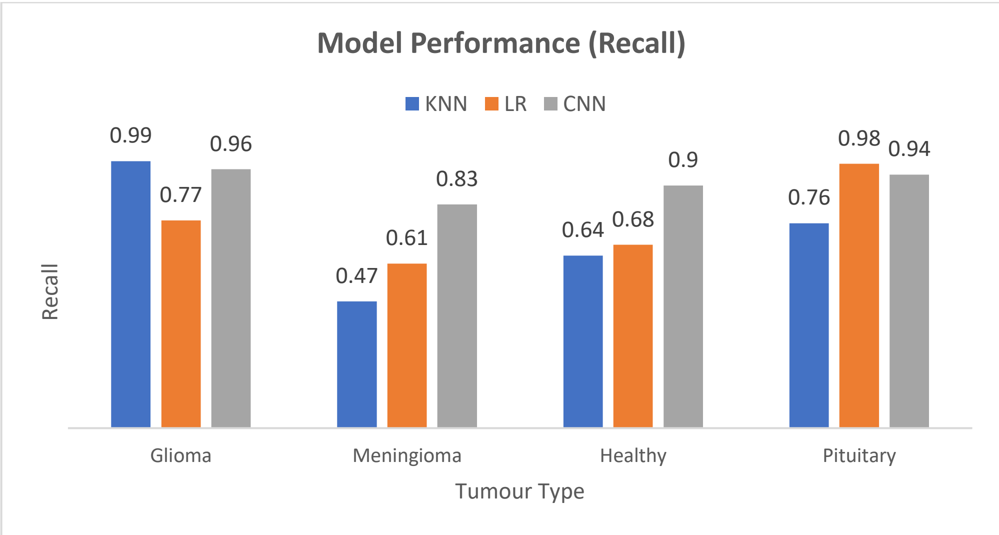

Figure 3: Class-level model performance (Recall)

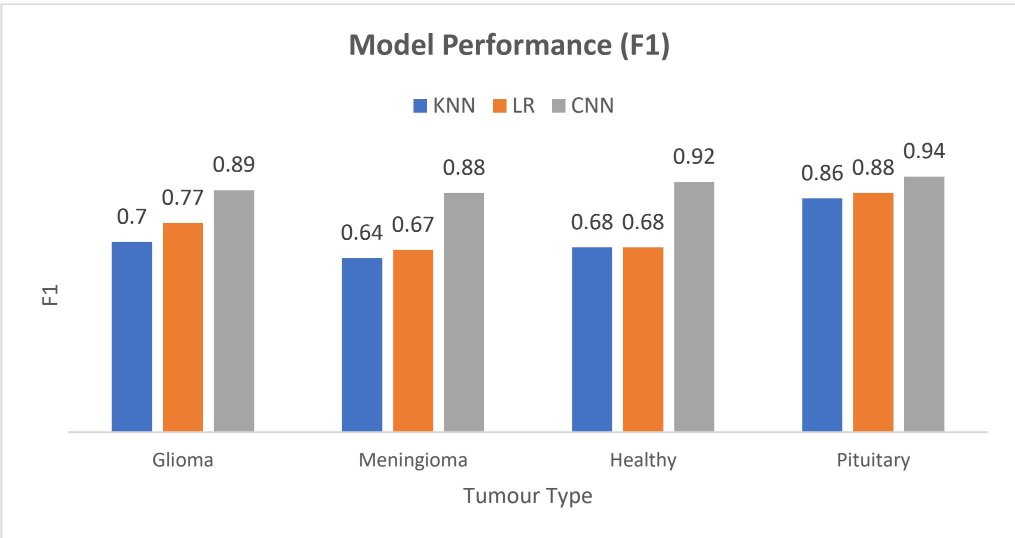

Figure 4: Class-level model performance (F1)

Figure 1 shows the confusion matrix of the classification models for all four tumour classes. The confusion matrix shows the correctly and incorrectly predicted images by a classification model. It gives information about the true positives, false positives, false negatives and true negatives for each tumour class. From Figure 1, we see that CNN performed better than the traditional machine learning models in most of the classes (overall).

For glioma, KNN correctly detected 92 out of the 93 glioma images in the test data more than the other two models (89 by CNN and 72 by LR). For meningioma, out of the 94 images, 57 of them were correctly classified by LR, 44 by KNN and 78 by CNN. For pituitary, LR slightly performed better than CNN at correctly identifying images in this class. 88 and 85 out of the 90 images were correctly identified by the two models respectively. In predicting healthy image scans correctly, CNN tops the other two models: 45 out of the 50 healthy scan images were correctly classified, 32 and 34 by KNN and LR respectively.

In terms of the misclassification rates by the three models for each class, KNN seems to have more than LR with CNN having less on average. Huge misclassification can be seen between glioma and meningioma images by KNN. 47 out of the 177 images (27%) predicted to be glioma tumours were misclassified by KNN. Also, from Figure 1, it can be observed that all three models had difficulty distinguishing between glioma and meningioma tumour images.

Table 1 shows the overall performance evaluation metrics: accuracy and F1-score for each model for all four classes. Accuracy is the fraction of images correctly identified by a model in all class labels used. F1 is a metric that combines two metrics: precision and recall. Precision is given by the percentage of images of a class correctly identified by a model for all its predictions while recall is the percentage of images in a particular class that were correctly classified by a model. From Table 1, it can be observed that CNN performed far much better than traditional machine learning models, overall. On average, about 91% of images in each class were correctly detected by CNN, that a 9% misclassification error on average for all tumour types. However, for LR and KNN, their average accuracies are below 80% (77% LR and 72% KNN). Similarly, when we combine their precisions and recall, given by the f1-score, it can be observed that CNN outperforms the other two machine learning models.

Discussion

This project aimed to develop a CNN and two machine-learning models to classify 3,264 MRI brain scans into three tumour types (glioma, meningioma and pituitary) and one healthy brain. From the result, we see that CNN is the best model overall in terms of the metrics evaluated. Its performance over machine learning models could be attributed to its ability to automatically extract and segment image pixels (Ari and Hanbay, 2018). Similarly, the poor performance of LR and KNN could be because of the limitations where its predictions are obtained from linear combinations of the independent variables while KNN on the other hand suffers from the curse of dimensionality owing to the 963 principal components extracted using PCA.

The performance (0.9083) of the CNN model is relatively comparable with the accuracy scores obtained by (Saeedi *et al.*, 2023) where they obtained a test accuracy of 0.9345 for a 2D-CNN model and 0.9093 for CNN with an autoencoder. In their paper, the images were resized to an 80 by 80 size and used different CNN architectures.

In conclusion, brain cancer is a deadly disease caused by an imbalance in brain cell growth and death. The early detection and diagnosis of brain tumours will enable early treatment and also hence save lives. However, this requires the presence of expert radiologists and doctors to diagnose it and hence can be time-consuming. Deep learning models, especially CNN, offer a promising way to reduce diagnostic time and allow these experts to focus on brain tumour MRI images that are difficult to distinguish.

References

Ari, A. and Hanbay, D. (2018) 'Deep learning based brain tumor classification and detection system', *Turkish Journal of Electrical Engineering and Computer Sciences*, 26(5), pp. 2275–2286. Available at: https://doi.org/10.3906/elk-1801-8.

Jollife, I.T. and Cadima, J. (2016) 'Principal component analysis: A review and recent developments', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065). Available at: https://doi.org/10.1098/rsta.2015.0202.

Kareem, S.W. *et al.* (2023) 'Comparative evaluation for detection of brain tumor using machine learning algorithms', *IAES International Journal of Artificial Intelligence*, 12(1), pp. 469–477. Available at: https://doi.org/10.11591/ijai.v12.i1.pp469-477.

Pedada, K.R. *et al.* (2023) 'A novel approach for brain tumour detection using deep learning based technique', *Biomedical Signal Processing and Control*, 82(December 2021). Available at: https://doi.org/10.1016/j.bspc.2022.104549.

Saeedi, S. *et al.* (2023) 'MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques', *BMC Medical Informatics and Decision Making*, 23(1), pp. 1–17. Available at: https://doi.org/10.1186/s12911-023-02114-6.

Vipin Y. Borole, Sunil S. Nimbhore, D.S.S.K. (2015) 'Image Processing Techniques for Brain Tumor Detection: A Review', *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS) Web*, 4(5), pp. 28–32. Available at: http://www.ijettcs.org/Volume4Issue5(2)/IJETTCS-2015-10-01-7.pdf.