Improving the Accuracy of Brain Tumour Classification and Object Detection

James Wright
University of Surrey
Guildford, GU2 7XH
jw01324@surrey.ac.uk

Zeya Rabani
University of Surrey
Guildford, GU2 7XH
zr00083@surrey.ac.uk

Harry Payne
University of Surrey
Guildford, GU2 7XH
hp00357@surrey.ac.uk

Dylan Parker
University of Surrey
Guildford, GU2 7XH
dp00405@surrey.ac.uk

Abstract

In this study, we classified the type of brain tumour a patient has based on various models trained from T1-weighted MRI slices. The three types of brain tumour focused on in this study are meningioma, glioma, and pituitary because these types of brain tumour are the most common: Gliomas (50.4%), Meningiomas (20.8%), Pituitary adenomas (15%) [1]. We experimented with different techniques and models to improve the accuracy of classification of the tumour and detection of the location of the tumour (object detection). Improved classification means that patients are more likely to receive the correct treatment for their specific tumour. Our goal is to implement a classification model with a high accuracy (over 90 percent), when classifying unseen data, as well as successfully implementing an object detection model that can detect the location of the patient's tumour.

1. Introduction

Classifying brain tumours is not an easy task, and if attempted by humans is prone to error. We found that AI based diagnosis was 94.6 percent accurate compared to 93.9 percent when identified by human doctors [2]. Hence, having a sophisticated AI that can more accurately identify specific brain tumours will lead to a more accurate and reliable diagnosis. This would mean that the patient can be treated more effectively. Magnetic Resonance Imaging (MRI) scans of patients' brains have been chosen as our dataset because they are the standard imaging technique for diagnosing brain tumours, although Computerised Tomography (CT) scans can also be used to identify brain tumours [3]. Feature extraction and classification are key to identifying brain tumours. The models we will create assume the patient has undergone procedures to identify if they have a brain tumour, and is positive, so we will only be classifying the type. Each section within this report explores important technologies, data, and research that has been conducted previously, allowing for a fuller picture of the subject, alongside our findings [15], [17], [18].

2. Background & Literature Review

2.1 Brain Tumour classes

A brain tumour is described as a growth of cells in the brain that multiplies in an abnormal, uncontrollable way [4]. An identified brain tumour can be classified through four different grades, with two main types. Their grades are based upon the rate of tumour growth and likelihood of return. The main types of brain tumours are as follows; non-cancerous (benign) and cancerous (malignant). Benign brain tumours are low grade, typically grade 1 or 2. Whilst Malignant brain tumours are high, typically grade 3 or 4 [4].

Currently, there are nearly ten types of brain tumours. Three types of brain tumour include Glioma, Meningioma, and Pituitary.

- Glioma – brain tumours that begin in the glial cells, supporting cells of the brain and spinal cord. Glioma tumours are found to be both benign and malignant. There are three types of glioma cells each with different tumours associated with them [5].

- Meningioma – tumours that start in the layers of tissue (meninges) that cover the brain and spinal cord. Most meningiomas are benign. Around twenty-one percent of brain tumours diagnosed between 2006 and 2010 were meningiomas [6].

- Pituitary – also called endocrine tumours. These are brain tumours that begin in the pituitary gland, found behind the eyes. The pituitary gland controls the production and releasing of hormones into the bloodstream. Pituitary tumours can be divided further into two groups depending on whether they manufacture hormones or not. Those that make hormones are called secreting, or functional tumours. Whilst those that do not are non-secreting, or non-functional. Around eight percent of brain tumours diagnosed between 2006 and 2010 were pituitary. These types of tumour are most often benign [7].

2.2 Identifying Brain Tumours

There are many ways of identifying brain tumours within a patient. The chosen method of identification usually is dependent on the health, vulnerability, and suspicions. The two primary ways of identifying or testing a patient for brain tumours is as follows: Magnetic Resonance Imaging (MRI), Blood Tests, and Tissue Sampling [8].

- MRI – uses magnetic fields to produce detailed images of the body, or its organs. It can be used to identify and measure the tumours size. A non-toxic dye, called a contrast, is fed or injected into the patient before the scan occurs. MRIs are the preferred approach to identifying tumours. Some variants of an MRI include Intravenous

Gadolinium-enhanced MRI, functional MRI, Magnetic Resonance Spectroscopy, and a technique called diffusion weighted imaging.

- Blood Tests – blood is extracted from a vein. The extracted blood then goes through a series of tests. These include full blood count, urea and electrolytes, live function tests, and tumour markers test. Bloods test are used to identify tumours due to the minimal risk associated with them [9].

- Tissue Sampling – although this can be used as a separate method of diagnosis, it is usually used alongside MRI scans in the final stages of diagnosis. This includes the removal of a small amount of tissue for examination under a microscope. This is known to be the only definitive way to identify a brain tumour with complete accuracy.

2.3 Existing Algorithms

K-Nearest Neighbours (KNN) is an algorithm frequently used for classification. It identifies images through routinely comparing a single piece of data from a set to each other piece of data in that same set, each time comparing distances and placing them in ascending order. It then obtains the first K (a chosen integer) entries and returns the mode of these labels. One of the major disadvantages of KNN is that as the dataset it is using becomes larger, it significantly reduces in speed [10].

Another algorithm that can be used for classification is Supervised Machine Learning (SVM). The SVM classifier transforms the data provided and locates boundaries between them. This is achieved through a technique called the 'kernel trick' which allows complex relationships to be created between images. However, this is at the expense of computational power and time [11].

Logistic Regression is one final algorithm that is geared towards categorically dependent data. It predicts the probability of a dependent variable belonging to a specific class, where it would typically be difficult to define. This forces only a set number of outcomes, based on the number of classes the dataset contains [12]. An example of this is shown in Figure 1.

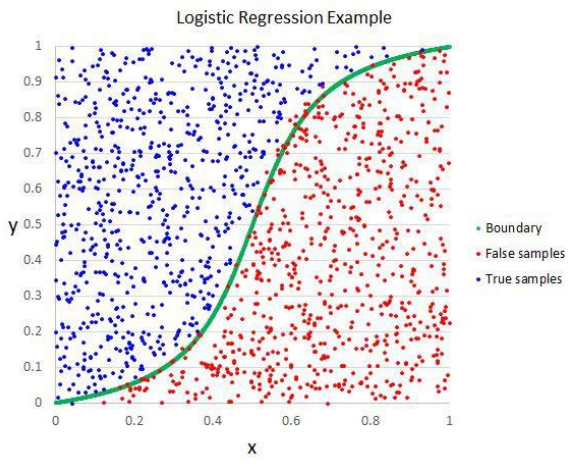

Figure 1: Logistic Regression Example [12]

One of the primary and most effective algorithms that exists for object detection is Region Convolutional Neural Network (R-CNN). This algorithm uses selective search by generating near two thousand bounding boxes for image classification. For each bounding box obtained, image classification occurs through CNN. This compares features of previously trained images to the image in question and determines the probability of it belonging to this class or label [13]. Each bounding box is then refined using regression.

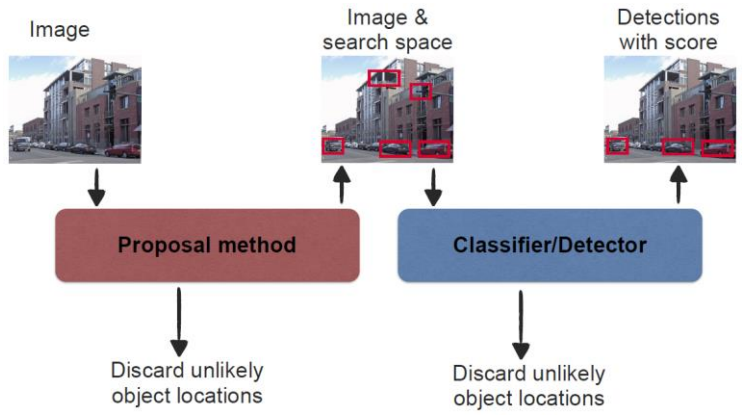

Figure 2 - Stages of R-CNN [14]

R-CNN can be developed further to allow for near real-time image detection. This was the basis for the development of Faster R-CNN. An architecture that utilizes a Region Proposal Network (RPN) to pool features from a generated backbone map. The backbone map is created by passing the image through CNN beforehand. Similarly, with Fast R-CNN and R-CNN, Faster R-CNN requires training to achieve near real-time object detection.

2.4 Existing Research

In the research paper 'Classification using Deep Learning Neural Networks for Brain Tumors' they used a Deep Neural Network (DNN) to classify a dataset of 66 MRI scans of the brain into four classes of tumors [15]. Their methodology was also split into four, from brain MRIs dataset acquisition, image segmentation using Fuzzy C-means, feature extraction using DWT and reduction using PCA technique and finally classification using DNN. All the brain MRIs were in the axial plane, T2-weighted and 256 x 256. The model performed brain tumour classification using 7-fold cross validation technique for building and training DNN of seven hidden layers structure, they also used KNN with k=1 and 3 for classification. The first three steps of the methodology were done using Matlab and the classification was done using Weka 3.9, the evaluation was measured using the average F-measure and the average areas under the ROC curve of the four tumor classes compared to other classifiers of the same term. Figure 3 and 4 show the classification rate, and other values, of each of the algorithms created within the research paper.

Algorithm	Classification rate	Recall	Precision	F-Measure	AUC (ROC)
DNN	96.97%	0.97	0.97	0.97	0.984
KNN K = 1	95.45%	0.955	0.956	0.955	0.967
KNN K = 3	86.36%	0.864	0.892	0.866	0.954
LDA	95.45%	0.955	0.957	0.955	0.983
[a]SMO	93.94%	0.939	0.941	0.963	0.939

Figure 3- Table: Performance of DNN, KNN, LDA and SMO classifiers [16]

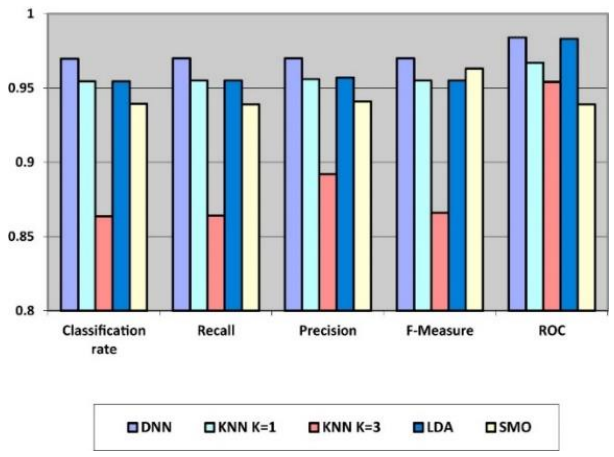

Figure 4 - Graph: Performance of DNN, KNN, LDA and SMO classifiers [16]

The results demonstrated by H. Mohsen et al. [15] show that Deep Neural Networks had the highest classification rate, precision, recall, f-measure, and ROC. This would suggest DNN would be a good choice for our model.

Although this research paper was our primary influence for the algorithm we implemented, other research papers were discussed, alongside the algorithms they used, and how the performance of those algorithms compared to the previous mentioned research paper.

One example of an alternate research papers was 'Fully automatic model-based segmentation and classification approach for MRI brain tumour using artificial neural networks' [17]. As identifiable within the title, the research papers used an Artificial Neural Network (ANN) to classify 200 MRI scans of the brain. It however was not concerned with identifying the type of tumour, but instead focuses on its location and segmenting the tumour from the image.

N. Arunkumar et al. [17] methodology was to, firstly, enrich the image using Fourier filtering. A process that reduces the high frequency components of the image. Secondly, perform image segmentation through binarization. This involves splitting a set of images into trainable and testing classes. Then finally, filter out any non-brain object within the image. This involved the use of a Histogram of Orientated Gradients (HOG), capable of achieving an accuracy of 91.48%.

The following is a Bland-Altman evaluation for the approach created by this research paper vs manual measurement. The horizontal lines are indications of the mean difference, and the limits of agreement.

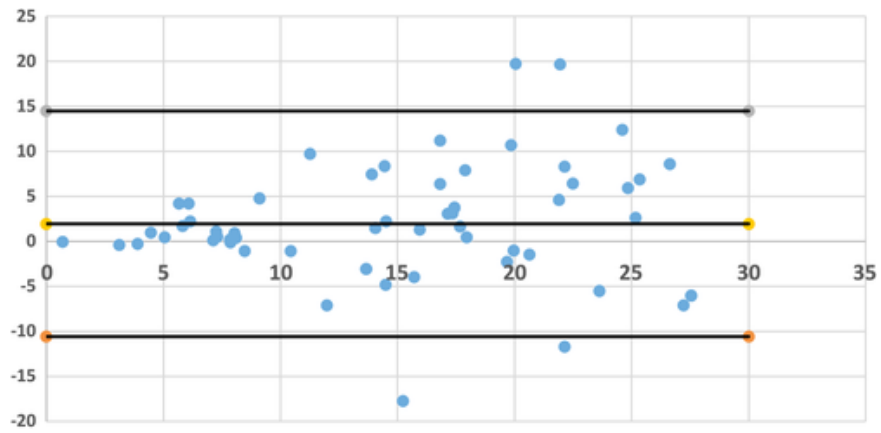

Figure 5: Altman Evaluation

The final research paper we discussed, published May of 2019, is named "Automated quantitative tumour response assessment of MRI in neuro-oncology with artificial neural networks: a multicenter, retrospective study" [18].

This paper trained an Artificial Neural Network (ANN) using 455 MRI scans of brain tumours. Whilst testing was accomplished with two datasets, one consisting of 239 MRI scans (Heidelberg) and the second containing 2034 MRI scans (EORTC-26101). The dataset contained T1-weighted and T2-weighted MRIs, that before training would have a contrast agent applied. A flow chart breakdown of the methodology explored in this research paper is shown in Figure 6.

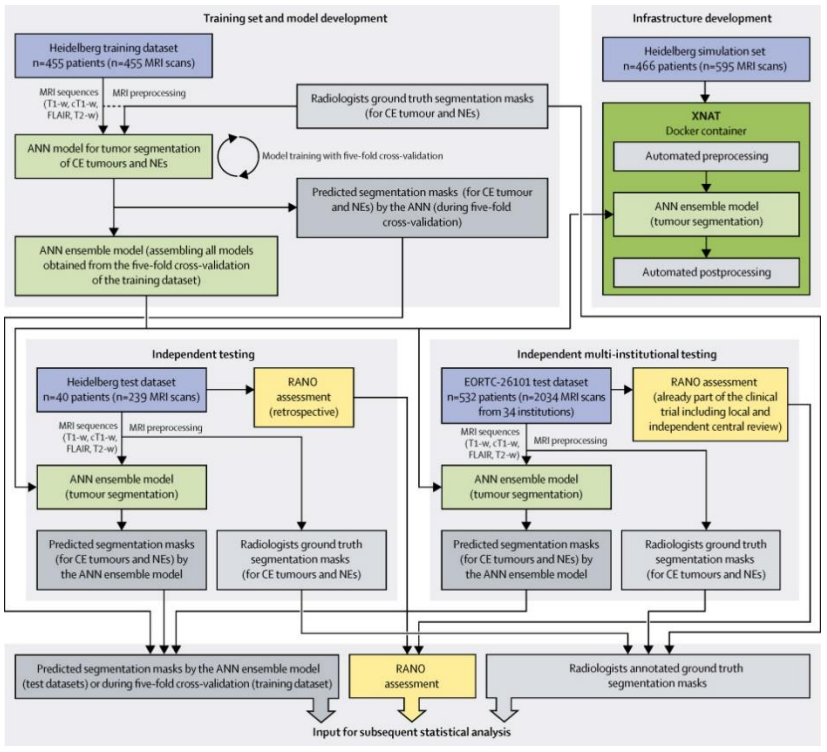

Figure 6: Artificial Neural Network Flow Chart

P. Kickingereder et al. [18] found that the ANN yielded excellent performance for accurate detection and segmentation of contrast enhanced tumours and non-enhanced volumes. Regardless of the test dataset used. The percentages of successful detection and segmentation is show in Figure 7.

	Heidelberg test dataset	EORTC-26101 test dataset
Tumour segmentation agreement		
Contrast-enhancing tumour	0·89 (0·86–0·90)	0·91 (0·90–0·92)
Non-enhancing T2-signal abnormality	0·93 (0·92–0·94)	0·93 (0·93–0·94)
Tumour volume agreement		
Contrast enhancing tumour	0·99 (0·99–1·00)	0·99 (0·99–0·99)
Non-enhancing T2 signal abnormality	0·99 (0·99–0·99)	0·98 (0·98–0·99)

Figure 7: Tumour Segmentation Table

3. Dataset

The dataset used in our model was created by Jun Cheng from the School of Biomedical Engineering [19]. It contains 3064 T1-weighted images from MRI scans of 233 different patients with three kinds of brain tumour (Figure 8). It contains the following number of types of brain tumour: 708 slices of meningioma, 1426 slices of glioma, and 930 slices of pituitary. The data is in matlab format (.mat file). Each file contains five fields for an image: label, patient ID, image data, tumour border, and tumour mask. The labels are: 1 (meningioma), 2 (glioma), 3 (pituitary). Furthermore, the dataset contains MRIs from a variety of directions, and different stages of tumour growth.

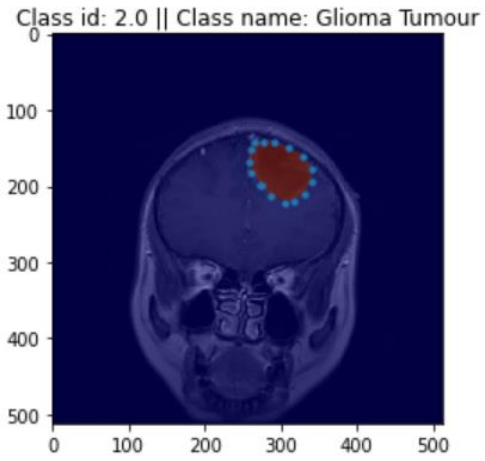

Figure 8 - Example of Images from Jun Cheng's Dataset [19]

The dataset was used in a study in 2015 focused on brain tumour classification [20]. This study has highly influenced the choices made during our study; primarily through the types of models used.

4. Models

From the research papers explored within the Literature Review, their exists some very distinguishable and frequently occurring features and methodologies. The most noteworthy being the stages of manipulation each research paper performs on the dataset to achieve the results they desire [15], [17], [18]. This involves breaking the model into multiple subsections, which include; Image Segmentation, Feature Extraction, and Dataset Classification. As these stages provide effective results, they are also present within our model.

4.1 Image Pre-Processing

4.1.1 Resizing the Dataset

For all models created, the images from our dataset needed to be reduced in size because they were taking up too much memory. We resized the images from 512 square pixels to 128 square pixels (Figure 9). This was the smallest possible dimensions we could have without affecting the features within the image itself. As well as the images being resized, we had to resize the mask data and tumour border data so that it fits our new image size. This was more complex to achieve due to the format of both the mask and tumour border datasets. To reduce the mask, we found the mean (average) for four pairs of (x, y) datapoints in the dataset – essentially a 4x4 pooling layer. Finally, each value was rounded to the nearest whole number, as the mask only reads binary (zeros and ones).

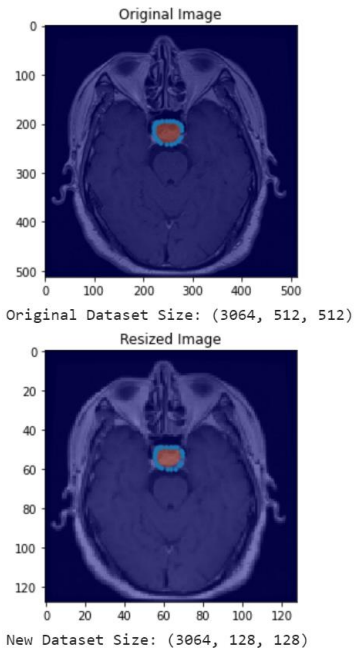

Figure 9: Resized Image Example

When resizing the borders, we had to account for two array types existing within the dataset. Each had to be handled differently. For NumPy Arrays, we divide the array by the original image size divided by the new image size. For the new size in question (128 x 128 pixels) this produces a value of four. For traditional array types, we must iterate through each element of the array and divide it by this same number individually. These new arrays can then be used to produce the downsized borders.

4.1.2 Feature Extraction

To improve accuracy, the appropriate feature extraction techniques needed to be used on the images in the dataset.

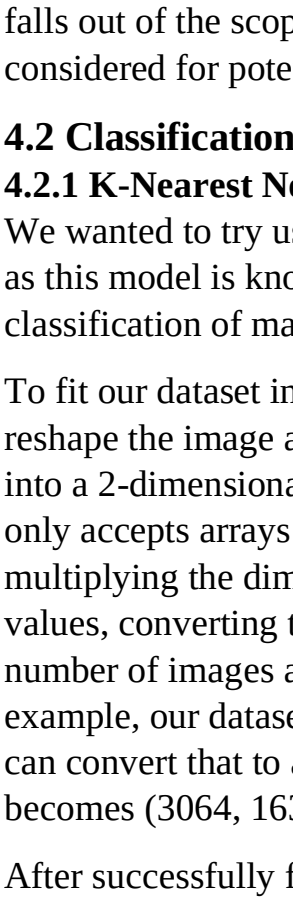

Figure 10: Image Processing

The first feature extraction technique used was to blur the image slightly to reduce the noise in the image. This is because tumours are medium to large sized in many of the slices, which means that they are not affected by the blur whereas smaller blemishes on the image become less visible.

The second feature extraction technique used was histogram equalization; this increases the contrast of the image so that the tumour area stands out. We found the model performed worse when this was implemented, and each image contrasted by a single value. In order for it to be effective for the model's accuracy, each image requires a unique contrast adjustment. Unfortunately, this would have taken too long to be implemented within a suitable time and thus falls out of the scope of our project but has been considered for potential future work (section 6).

4.2 Classification

4.2.1 K-Nearest Neighbours (KNN)

We wanted to try using KNN for the tumour classification as this model is known for being a good general model for classification of many datasets [10].

To fit our dataset into a KNN model we first needed to reshape the image array from a 3-dimensional (3-D) array into a 2-dimensional (2-D) array, because the KNN model only accepts arrays of 2-D or less. We achieved this by multiplying the dimensions related to the image pixel values, converting the pixel data into a 1-D array, with the number of images as the other dimension of the array. For example, our dataset has a shape of (3064, 128, 128) – we can convert that to a 2-D array, so (3064, 128 x 128) becomes (3064, 16384).

After successfully fitting our data to the KNN model, splitting data into seventy percent training and thirty percent testing, the optimum value of K needs to be found. To achieve this, we created a KNN model with each value of K from one to forty and calculated the error rate for each model. This created Figure 11, which suggests the correct K value for the model is one, with a mean error value of 0.05.

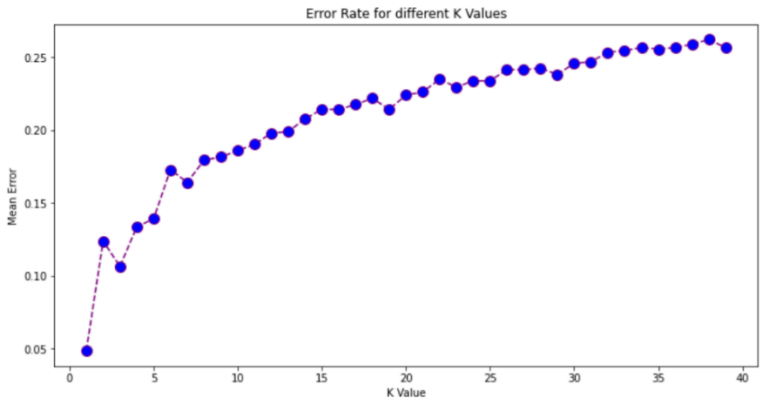

Figure 11: Error Rate for different K Values (KNN)

Adding this value or K (1) into the model gives the model an accuracy of 94.12 percent. An accuracy higher than human identification; 93.9 percent. The model is validated against the test dataset, because KNN does not require training of the model against test data. Therefore, validation on unseen data can be done on the test dataset. See Figure 13 for the confusion matrix, precision, recall, and f1-score of our model.

		Actual		
		Meningioma	Glioma	Pituitary
Predicted	Meningioma	246	26	5
	Glioma	35	529	3
	Pituitary	3	0	379

Figure 12: Confusion Matrix for KNN Model

	precision	recall	f1-score	support
1.0	0.87	0.89	0.88	277
2.0	0.95	0.93	0.94	567
3.0	0.98	0.99	0.99	382
accuracy			0.94	1226
macro avg	0.93	0.94	0.94	1226
weighted avg	0.94	0.94	0.94	1226

Figure 13: Precision, Recall, and f1-score for KNN Model

4.2.2 Convolutional Neural Network (CNN)

We wanted to try using CNN for the tumour classification as this model is known for being a good model for classifying images specifically, of which our dataset consists of.

To fit our dataset into a CNN model we must reshape the image array and the label array. For the image array we had to add an extra dimension for the colour or greyscale attribute of an image. As our image data is in greyscale, our shape was converted from (3064, 128, 128) to (3064, 128, 128, 1). Our labels were converted into one hot format; where the labels are represented a binary array to specify which class an image belongs to.

```
Using TensorFlow backend.
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_1 (Conv2D)            (None, 128, 128, 64)      640
_________________________________________________________________
activation_1 (Activation)    (None, 128, 128, 64)      0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 128, 128, 64)      36928
_________________________________________________________________
activation_2 (Activation)    (None, 128, 128, 64)      0
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 63, 63, 64)        0
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 61, 61, 128)       73856
_________________________________________________________________
activation_3 (Activation)    (None, 61, 61, 128)       0
_________________________________________________________________
conv2d_4 (Conv2D)            (None, 59, 59, 128)       147584
_________________________________________________________________
activation_4 (Activation)    (None, 59, 59, 128)       0
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 29, 29, 128)       0
_________________________________________________________________
flatten_1 (Flatten)          (None, 107648)            0
_________________________________________________________________
dense_1 (Dense)              (None, 128)               13779072
_________________________________________________________________
activation_5 (Activation)    (None, 128)               0
_________________________________________________________________
dense_2 (Dense)              (None, 3)                 387
_________________________________________________________________
activation_6 (Activation)    (None, 3)                 0
=================================================================
Total params: 14,038,467
Trainable params: 14,038,467
Non-trainable params: 0
```

Figure 14: CNN Model Summary

The CNN Model we created consists of four 2-D convolution layers with Relu as the activation function. Relu is an activation function that acts as a gate; only allowing inputs that are positive through. Any other inputs are outputted as zero. The first and last layers contain additional max pooling layers, allowing the model to learn more about features contained within the destroyed regions. Each subsequent convolution layer has an increasing number of filters, to account for the increasingly complex combinations of patterns. A flatten layer combines all dimensions within the data into a single dimension, which is then manipulated through use of multiple dense layers to output three probabilities for each tumour class. We used a batch size of sixty-four, and thirty epochs. These values were primarily chosen as it allows for the model to be run in under two minutes, whilst not having a considerable effect on the accuracy.

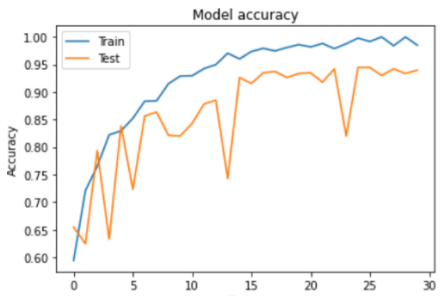

Figure 15: CNN Model Accuracy

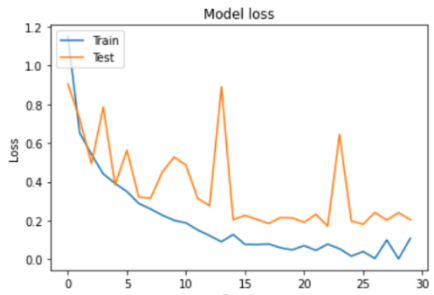

Figure 16: CNN Model Loss

Figure 15 and Figure 16 show the accuracy and loss of the trained model at each new epoch. It is clear the highest increase in accuracy, and highest decrease in loss, was achieved within the first five epochs. This was also where the model was at its most stable, both in terms of the accuracy and loss gradients. The final achieved accuracy and loss after thirty epochs was 93.96 percent, and 22.08 percent, respectively (Figure 17).

```
Epoch 30/30
1929/1929 [==============================] - 10s 5ms/step - loss: 0.1087 - accuracy: 0.9850 - val_loss: 0.2041 - val_accuracy: 0.9396
```

Figure 17: CNN 30th Epoch Accuracy & Loss

Once the model was trained, we needed to validate whether the model is not overfitted. Hence, checking if the model is accurate at classifying unseen data. To do this we separated a small amount of data from our dataset and saved this as the validation set; this data is not used in the training of the model. If the model has a high accuracy and low error (loss) when predicting the class of the unseen data, then our model has not been overfitted. This is proof that our model can classify tumour types of MRI scans of this type with high accuracy. When predicting classes of unseen data, the loss value and accuracy was 22.08 percent and 94.46 percent respectively, shown in Figure 18.

```
Loss: 22.08%
Accuracy: 94.46%
```

Figure 18: Accuracy & Loss for CNN Model

4.2.3 Support Vector Machine (SVM)

When using a Support Vector Machine (SVM) model for classification of tumour types, our image array had to be reshaped like we did for our KNN model. The image array was reshaped from a 3-D array to a 2-D array in the form (3064, 16384).

After creating the SVM model and fitting our data into the model, splitting data into seventy percent training and thirty percent testing, we achieved an accuracy value of 84.24 percent. Figure 19 and Figure 20 contain the

confusion matrix, precision, recall, and f1-score of our model.

		Actual		
		Meningioma	Glioma	Pituitary
Predicted	Meningioma	117	57	29
	Glioma	33	372	21
	Pituitary	5	0	286

Figure 19: Confusion Matrix for SVM Model

	precision	recall	f1-score	support
1	0.75	0.58	0.65	203
2	0.87	0.87	0.87	426
3	0.85	0.98	0.91	291
accuracy			0.84	920
macro avg	0.82	0.81	0.81	920
weighted avg	0.84	0.84	0.84	920

Figure 20: Precision, Recall, and f1-score for SVM Model

4.2.4 Logistic Regression

We also used a Logistic Regression model for classifying tumours, with the same reshaping of the image array needed as KNN and SVM; reshaping the data from a 3-D array to a 2-D array (3064, 16384).

After creating the Logistic Regression model and fitting our data into the model, splitting data into seventy percent training and thirty percent testing, we got an accuracy value of 84.45 percent. Figure 21 and Figure 22 contain the confusion matrix, precision, recall, and f1-score of our model.

		Actual		
		Meningioma	Glioma	Pituitary
Predicted	Meningioma	146	42	15
	Glioma	61	354	11
	Pituitary	7	7	283

Figure 22: Confusion Matrix for Logistic Regression Model

	precision	recall	f1-score	support
1	0.68	0.72	0.70	203
2	0.88	0.83	0.85	426
3	0.91	0.95	0.93	291
accuracy			0.84	920
macro avg	0.82	0.83	0.83	920
weighted avg	0.85	0.84	0.84	920

Figure 23: Precision, Recall, and f-1 score for Logistic Regression Model

4.3 Object Detection

For Object Detection we used a CNN model with reshaped x and y data. Each also contained an additional dimension. This was required to identify the values/images as greyscale, whilst reshaping allowed less demand on the RAM. Our x data consisted of the MRI Images provided by the previously mentioned dataset. The y values, also provided by this dataset, were the tumor masks of the corresponding images. The aim was to train a model that would be capable of generating its own tumour mask of an MRI Image.

The CNN model can be broken into five segments; an input layer, three hidden layers, and an output layer. The input layer contains two 2-D convolution layers, each undergone batch normalization and containing the Relu activation function. A dropout layer also exists to prevent overfitting. This is followed by the first hidden layer, which is structured the same as the input layer, only with twice the number of filters. Hidden layers two and three are also identical. Each containing a transpose layer to upscale the output from the previous layers. This is then followed by a dropout layer and two 2-D convolution layers with batch normalization and Relu, as before. The output layer is a single 2-D convolution layer containing the sigmoid activation function. This layer returns the predictions dimensions to the same format as the inputted data. A batch size of thirty-two, alongside ten epochs, were used for training. The model was able to achieve a loss of 41.98 percent, and an accuracy of 98.53 percent.

Layer (type)	Output Shape	Param #
conv2d_73 (Conv2D)	(None, 128, 128, 32)	320
batch_normalization_65 (Batc	(None, 128, 128, 32)	128
conv2d_74 (Conv2D)	(None, 128, 128, 32)	9248
batch_normalization_66 (Batc	(None, 128, 128, 32)	128
max_pooling2d_17 (MaxPooling	(None, 64, 64, 32)	0
dropout_33 (Dropout)	(None, 64, 64, 32)	0
conv2d_75 (Conv2D)	(None, 64, 64, 64)	18496
batch_normalization_67 (Batc	(None, 64, 64, 64)	256
conv2d_76 (Conv2D)	(None, 64, 64, 64)	36928
batch_normalization_68 (Batc	(None, 64, 64, 64)	256
max_pooling2d_18 (MaxPooling	(None, 32, 32, 64)	0
dropout_34 (Dropout)	(None, 32, 32, 64)	0
conv2d_transpose_17 (Conv2DT	(None, 64, 64, 64)	36928
dropout_35 (Dropout)	(None, 64, 64, 64)	0
conv2d_77 (Conv2D)	(None, 64, 64, 64)	36928
batch_normalization_69 (Batc	(None, 64, 64, 64)	256
conv2d_78 (Conv2D)	(None, 64, 64, 64)	36928
batch_normalization_70 (Batc	(None, 64, 64, 64)	256
conv2d_transpose_18 (Conv2DT	(None, 128, 128, 32)	18464
dropout_36 (Dropout)	(None, 128, 128, 32)	0
conv2d_79 (Conv2D)	(None, 128, 128, 32)	9248
batch_normalization_71 (Batc	(None, 128, 128, 32)	128
conv2d_80 (Conv2D)	(None, 128, 128, 32)	9248
batch_normalization_72 (Batc	(None, 128, 128, 32)	128
conv2d_81 (Conv2D)	(None, 128, 128, 1)	33

Total params: 214,305
Trainable params: 213,537
Non-trainable params: 768

Figure 21: Object Detection Model Summary

The CNN model generates a (128,128,1) array that contains probabilities. These probabilities determine the likelihood of each pixel being a tumour, or part of a tumour. Higher probabilities are red, whilst low probabilities remain dark blue. These outputs have been squeezed to allow them to be plotted.

The outputs of our model had difficulty detecting early stage tumours. This was likely due to the values the tumour had in relation to the surrounding brain. Which resulted in the model offering low probabilities of the tumour and surrounding brain being tumours (Figure 25). However, the tumours at later stages are larger and contain identifiable characteristics which the model was able to use for better tumour masking during prediction (Figure 24).

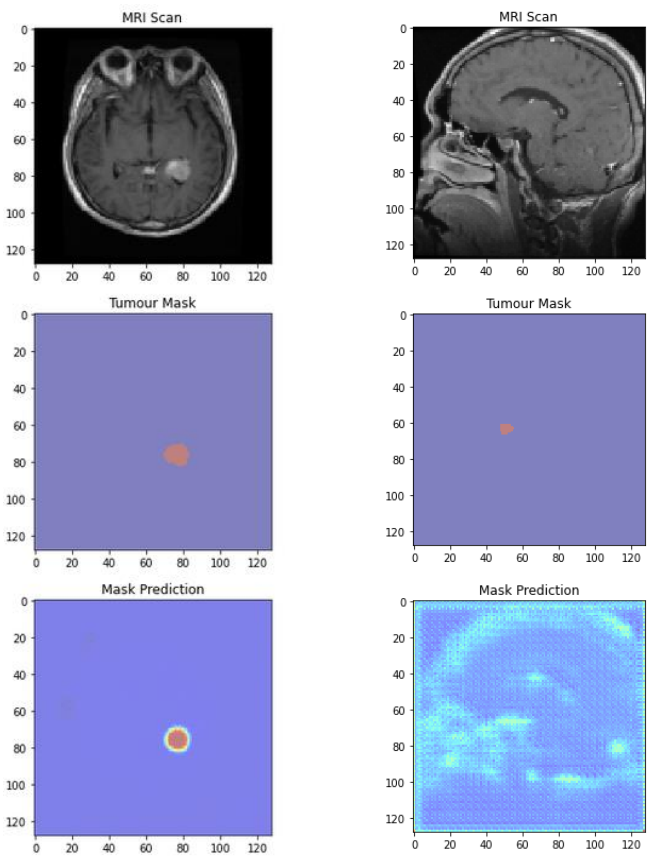

Figure 25: Plots of MRI Scan, Tumour Mask, Successful Mask Prediction

Figure 24: Plots of MRI Scan, Tumour Mask, Unsuccessful Mask Prediction

5. Evaluation

The classification model that gave us the best accuracy was Convolutional Neural Network (CNN), with an accuracy of 94.46 percent on unseen data. K-Nearest Neighbours (KNN) was the second most accurate model, with an accuracy of 94.12 percent on unseen data. Whereas Support Vector Machine (SVM) and Logistic Regression had very similar, but lower accuracies with 84.24 percent and 84.45 percent, respectively (Figure 26, Figure 27).

Model	Accuracy (%)
CNN	94.46%
KNN	94.12%
SVM	84.24%
Logistic Regression	84.45%

Figure 26: Table of Classification Model Accuracies

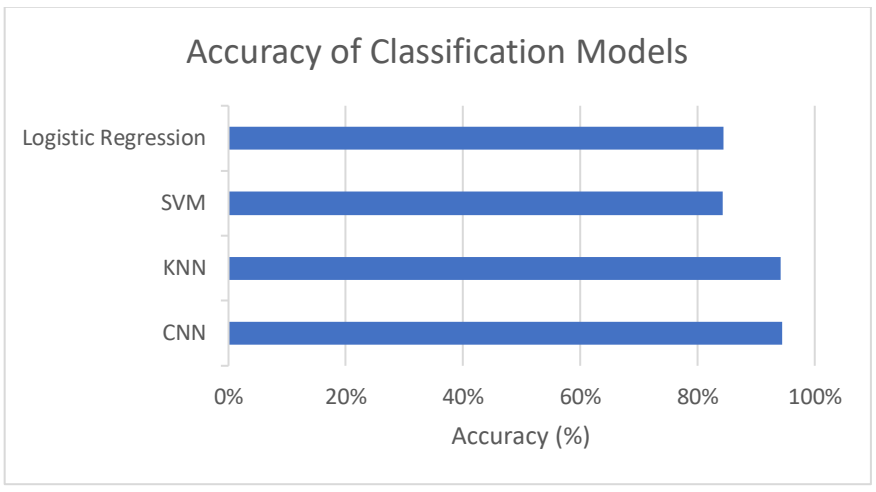

Figure 27: Chart of Classification Model Accuracies

In the research paper 'Classification using Deep Learning Neural Networks for Brain Tumors', the accuracy of H. Mohsen et al [15] KNN Model was, at its highest, 95.45 percent. In comparison, our KNN model falls 1.33 percent behind in accuracy. They also used a deep neural network that included seven hidden layers and utilized 7-fold cross validation. This allowed it to achieve 96.97 percent accuracy, the highest accuracy they achieved through use of their models. Our CNN model contains a reduced number of layers and still achieved 94.46 percent accuracy; 2.51 percent lower. The reduced number of layers allowed us to create a much more time restrained system, compared to the one generated in the research paper [15].

The object detection CNN model had an accuracy of 98.53, an impressive accuracy in comparison to the research paper 'Fully automatic model-based segmentation and classification approach for MRI brain tumour using artificial neural networks [17]. They achieved 91.48 percent accuracy through use of a Histogram of Gradients (HOG). However, they achieved much higher consistency with identification of their tumours across sizes and colouration.

6. Conclusions and Future Work

Our study set out to try to implement a model that would increase the accuracy of classification for tumours (Gliomas, Meningiomas, and Pituitary adenomas); this goal was achieved. We created four models; CNN, KNN, SVM, and Logarithmic Regression, of which performed admirably in accuracy of classification with low loss. We found with classification, producing a faster model with slightly decreased accuracy boded better for the majority of cases. This would allow patients to be treated faster, potentially increasing the likelihood of a positive outcome for the patient.

We also set out to implement an object detection model that correctly identifies a tumour in an MRI scan of the brain. This was completed through training a CNN model to mask regions of the brain suspected of being tumours. However, we obtained mixed results. The predicted masks accuracy was dependent on the size and colouration of the tumours seen within the MRI scans.

Combination of both Convolutional Neural Networks to produce a singular neural network is something we would like to achieve in future development. This would allow a mask and label to be generated simultaneously for an inputted MRI scan. Improving the consistency and detection rate of the object detection CNN is also something that would be seen in future work. We would also rework, through use of a model or otherwise, the contrast feature extraction to allow our model to reach its maximum potential.

7. Contributions

7.1 Poster

- **James Wright:** Abstract, Introduction, Dataset, Models, Evaluation, Conclusions and Future Work, Contributions, References, Converting References to IEEE format.
- **Zeya Rabani:** Introduction, Background & Literature Review, Contributions.
- **Harry Payne:** Background & Literature Review, Contributions.
- **Dylan Parker:** Background & Literature Review, Dataset, Models, Evaluation, Conclusions and Future Work, Contributions, References, Proof Reading.

7.2 Coding

- **James Wright:** Resizing Images, Pre-Processing Images, KNN, SVM, Logistic Regression, Adjusting Classification CNN Model, CNN Model Validation.
- **Zeya Rabani:** Classification CNN Model.
- **Harry Payne:** Importing Dataset from GitHub, Formatting & Reshaping Arrays, Classification CNN Model.
- **Dylan Parker:** Resizing Mask & Border, Displaying Standard & Resized Images, Classification CNN Model, Object Detection CNN Model, Classification and Mask Prediction.

8. References

1. *BJ. Park, HK. Kim, B. Sade, JK. Lee, (2009). "Epidemiology". In JH Lee (ed.). Meningiomas: Diagnosis, Treatment, and Outcome. Springer. p. 11. ISBN 978-1-84882-910-7. Available: https://link.springer.com/chapter/10.1007/978-1-84628-784-8_2*

2. *NHS. (2020, May. 23) AI Trained to spot brain tumours faster than humans [Online]. Available: https://www.sciencefocus.com/news/ai-trained-to-spot-brain-tumours-faster-than-humans/*

3. *M. Iv, BC. Yoon, JJ. Heit, N. Fischbein, M. Wintermark (January 2018). "Current Clinical State of Advanced Magnetic Resonance Imaging for Brain Tumor Diagnosis*

and Follow Up". Seminars in Roentgenology. 53 (1): 45–61. Available: https://www.sciencedirect.com/science/article/abs/pii/S0037198X17300809?via%3Dihub

4. *NHS. (2020, Feb. 3) Brain Tumours [Online]. Available: https://www.nhs.uk/conditions/brain-tumours/*

5. *Cancer Research UK. (2019, Oct. 30) Glioma [Online]. Available: https://www.cancerresearchuk.org/about-cancer/brain-tumours/types/glioma-adults*

6. *Cancer Research UK. (2019, Oct. 30) Meningioma [Online]. Available: https://about-cancer.cancerresearchuk.org/about-cancer/brain-tumours/types/meningioma?_ga=2.25399997.1857704504.1588279625-128960178.1588279625*

7. *Cancer Research UK. (2019, Oct. 30). Pituitary [Online]. Available: https://about-cancer.cancerresearchuk.org/about-cancer/brain-tumours/types/pituitary-tumours?_ga=2.121696779.1857704504.1588279625-128960178.1588279625*

8. *Cancer.Net. (2019, Mar.). Brain Tumour: Diagnosis [Online]. Available: https://www.cancer.net/cancer-types/brain-tumor/diagnosis*

9. *Cancer Research UK. (2019, Apr. 23). Blood Tests [Online]. Available: https://www.cancerresearchuk.org/about-cancer/cancer-in-general/tests/blood-tests*

10. *O. Harrison (2018, Sept. 9). Machine Learning Basics with the K-Nearest Neighbors Algorithm. towards data science [Online]. Available: https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761*

11. *KDnuggets (2017). What is a Support Vector Machine, and Why Would I Use it? [Online]. Available: https://www.kdnuggets.com/2017/02/yhat-support-vector-machine.html*

12. *R. Brid. (2018, Oct. 17). Logistic Regression. Medium [Online]. Available: https://medium.com/greyatom/logistic-regression-89e496433063*

13. *Brandon. (2016, Aug. 18). How do Convolutional Neural Networks work?. Brohrer.github.io [Online]. Available: https://brohrer.github.io/how_convolutional_neural_networks_work.html*

14. *S.H. Tsang. (2018 Aug. 31). Review R-CNN Object Detection. Medium [Online]. Available: https://medium.com/coinmonks/review-r-cnn-object-detection-b476aba290d1*

15. *H. Mohsen, E.A. El-Dahshan, E.M. El-Horbaty, A.M. Salem (June 2018). Classification using deep learning neural networks for brain tumors. Future Computing and Informatics Journal [Online]. 3 (1): 68–71. Available: https://www.sciencedirect.com/science/article/pii/S2314728817300636*

16. *H. Mohsen, E.A. El-Dahshan, E.M. El-Horbaty, A.M. Salem. "Brain tumor type classification based on support vector machine in magnetic resonance images", Annals Of "Dunarea De Jos" University Of Galati, Mathematics, Physics, Theoretical mechanics, Fascicle II, Year IX (XL), No. 1; 2017.*

17. *N. Arunkumar, MA. Mohammed, SA. Mostafa, DA. Ibrahim, JJPC. Rodrigues, VHC. de Albuquerque. (2018, October) Fully automatic model-based segmentation and classification approach for MRI brain tumor using artificial neural networks. Concurrency Computat Pract Exper. [Online] Available: https://doi.org/10.1002/cpe.4962*

18. *P. Kickingereder, F. Isensee, I. Tursunova, et al. (2019, May) Automated quantitative tumour response assessment of MRI in neuro-oncology with artificial neural networks: a multicentre, retrospective study. Lancet Oncol, 20. [Online] Available: https://www.sciencedirect.com/science/article/pii/S1470204519300981*

19. *J. Cheng. (2017, Apr. 02) Figshare: Brain Tumor Dataset [Online]. Available: https://figshare.com/articles/brain_tumor_dataset/1512427*

20. *J. Cheng, W. Huang, S. Cao, R. Yang, W. Yang, Z. Yun. (2015, October). Enhanced Performance of Brain Tumor Classification via Tumor Region Augmentation and Partition. PLOS ONE [Online] Available: https://doi.org/10.1371/journal.pone.0140381*