Comparative Analysis of Brain Tumor Classification Using SVM and K-NN Algorithm

Steven C. Yoseph[1], Valen[1], Kaizerio E. Wibowo[1], Renaldy Fredyan[1], Muhammad Amien Ibrahim[1].

[1]*Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, 11480, Indonesia*

Abstract - Brain is a vital component of the nervous system. Therefore, brain tumor classification presents a unique challenge in image classification, particularly in the medical field. In recent years, artificial intelligence and deep learning, especially in the medical domain, have been advancing rapidly. These developments are vital in biomedical engineering because of the delicate nature of health-related issues. Deep learning is frequently applied in medical classifications, and accurate, swift diagnosis is crucial for classifying brain tumors using MRI. This research seeks to create and develop an efficient system for classifying brain tumors that saving time, and also helping radiologists in making accurate diagnoses. We use two different models to achieve optimal accuracy.

Keywords Brain tumor · Image Classification · Magnetic Resonance Imaging (MRI) · Deep learning

1 Introduction

Tumors are groups of abnormal cells that form lumps or growths. Brain tumors are one of the most dangerous and deadly tumors. For humans, the brain is the most complex organ because the brain has an important role in controlling functions in our body, such as alertness, movement, sensation, thinking, speech, and memory. The tumor can cause our brain unable to carry out these functions properly. Brain tumors can occur due to genetic changes or mutation in brain cells. The growth of abnormal cells that occur can cause damage to important brain tissues and develop into cancer. There are two categories of brain tumors: benign tumors grow at a slower rate and do not metastasize to other tissues, while malignant tumors can spread to other parts of the body and progress quickly. The spreads will form new tumors that are destructive. The high rate of malignant tumor spread requires early diagnosis and accurate classification. Therefore, humans need tools to speed up classification time and to improve diagnosis accuracy.

Early diagnosis of glioma is important in finding ways to threat it. Medical imaging techniques such as Single-Photon Emission Computed Tomography (SPECT), Magnetic Resonance Spectroscopy (MRS), Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Positron Emission Tomography (PET) are used to provide useful information regarding shape, size, location, and brain tumors metabolism. MRI is famously known as a standard technique for identifying brain tumors because it can view soft tissue well. MRI is an imaging technique that uses radio frequencies that target tissue to create an image generated by a strong magnetic field. MRI must be done repeatedly with changing excitations that will be used to produce contrast images of different tissue types and yield valuable structural information and enable diagnosis.

There are 4 MRI standards that are used to detect glioma, namely T1-weighted MRI (T1), T2-weighted MRI (T2), T1-weighted MRI with gadolinium contrast enhancement (T1-Gd) and Fluid-Attenuated Inversion Recovery (FLAIR)[1]. When performing an MRI, there will be around 150 2D image slices produced to represent the 3D brain volume. MRI must be performed by an expert, which is a difficult and time-consuming task. Human observations can cause classification errors, therefore automatic or semi-automatic classification techniques are needed to help determine the type of tumor.

The presence of Artificial Intelligence (AI) is very helpful in reducing error rates, especially in the field of radiology. The machine learning method used is deep learning. Deep learning methods is being used by radiologist to detect and classify tumors quickly without the need for surgical intervention. The algorithm employs a subset of deep learning that is a Convolutional Neural Network (CNN. which is very common in "Medical Imaging Problems". Several models used in brain tumor classification are K-NN, SVM, etc. Model not only helps classify brain tumors but has also succeeded in classifying other medical diseases. The model will be trained first with the MRI dataset and compare the performance of one with another. However, classification of brain tumors is very difficult due to morphological structure changes, unclear tumor images, and irregular lighting effects. Therefore, radiologists need good brain tumor classification techniques to support their decisions. Over time, new methodologies were created to overcome the problems of brain tumor classification approaches.

2 Related Works

Priyanka, Balwindersingh is focusing on surveying the brain-tumor detection algorithms that so far detect the tumor's location[3]. There are many ways that it can be used to classify brain MRI images of both normal and abnormal. Our main objective and concentration are in techniques that can be used to do an image segmentation that is useful to detect tumors in the brain. Image segmentation itself was a digital image partition process into several segments.

R.J. Ramteke and Khachanemonali Y presented a method for automatically classifying medical images into two groups: normal and abnormal. This approach depends on image features and the automated detection of abnormalities[3]. For drawing classification there is a technique called K-NN (K-resistant) that is the most conceptual and comprehensive technique that can produce good classification accuracy. K-NN algorithms depend on the use of a function to measure distance and define majority sounds within a nearby neighbor number k (k-architectures), using a method of calculating distance from Euclidean. As for SVM (support vector machines), it is aimed and applied to the classification of brain images using the resulting features. The SVM has much higher reproductive capabilities and a much faster convergence.

Khushboo Singh and Satyaverma suggested sophisticated classification methods utilizing SVM, which is proposed and applied to classifying images of the brain using features already obtained[3]. The SVM is an artificial neural network (ANN) used for classification elevate learning. The characteristic of the SVM is its ability to solve classification problems through convex quadratic programming. SVM is trying to minimize battery on error generalization. The co-crystallized matrix (GLCM) is used to extract excretion from the data detected by the tumor. The co-logical matrix level gray is considered one of the most important techniques for obtaining statistical data that will be used for further classification, which is used in the study[4].

The remain of this paper is structured as follows: Section 2 provides a concise overview of brain tumor classification methods. Section 3 delves into the application of machine learning-based techniques. Finally, the Results section presents a comparative analysis of the outcomes from all models.

3 Method

The method used in this classification consists of several steps such as image pre-processing, classification, and evaluation. Fig 1

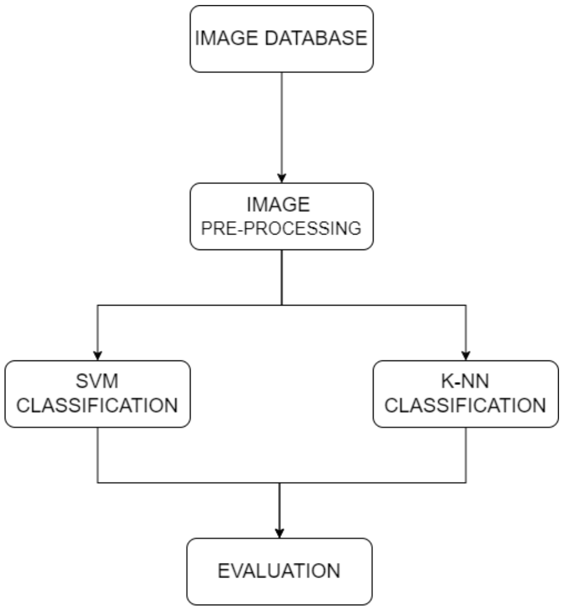

Fig. 1

3.1 Image Pre-processing

In the image pre-processing process, there are 3 main stages, namely data augmentation, train-test split, and normalization.

3.1.1 Data Augmentation

In data augmentation, we use the imread function from OpenCV library to read the image data and return them as an array. After that, we resize and reshape for each data in order to make the data size consistent and reduce the array dimensions.

3.1.2 Train-Test Split

In Train-Test Split, we split our data into two parts with 80% of training data and 20% of testing data. The training dataset was utilized to train machine learning models, while the testing dataset was employed to assess their performance. This data split was implemented to avoid overfitting.

3.1.3 Normalization

After splitting into train and test data, the training data will be normalized using MinMax Scaling. This normalization is useful for improving machine learning performance. This MinMax Scaling work by creating a value range of all features into the same range, usually between 0 and 1. If any feature has a higher value range, it would dominate other features and make a high bias.

MinMax Scaling formula :

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

Where x_{min} is a minimum value of the feature and x_{max} merupakan nilai maksimum dari fitur tersebut.

3.2 Classification

In this step, all features that have passed the pre-processing step will be used as an input for training and testing the classification models. We use 2 type of classification model, namely K-Nearest Neighbor (KNN) and Support Vector Machine (SVM)

3.2.1 K-Nearest Neighbor

K-Nearest Neighbor (KNN) is a straightforward classification method in which data is classified based on the majority class among its k-nearest neighbors. KNN uses the Euclidean distance formula to measure the distance between the data point being classified and its k closest data points to determine the majority class.

Eucledian distance formula :

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (2)[10].$$

where x_2 and $y_2$are the coordinates of the classified data, x_1 dan $y_1$are the coordinates of the target data, and d is a distance between classified data and target data.

3.2.2 Support Vector Machine (SVM)

Support Vector Machine (SVM) is a classification model where SVM classified data by finding the best hyperplane to divide two classes.

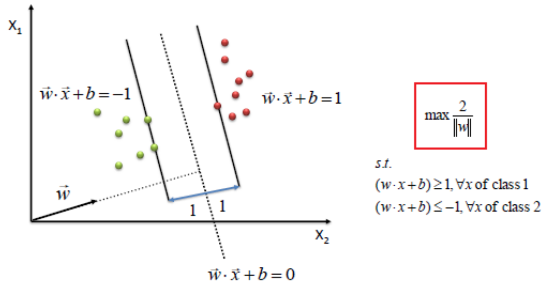

Fig. 2

3.3 Evaluation

After the dataset is classified using 2 classification models, the results will be compared and evaluated. After that, we will determine which type of classification is more optimal. The evaluated result are confusion matrix, accuracy, precision, recall, F1-score, etc. These matrices are used for evaluation because they are able to explain how accurate KNN and SVM are in classifying images into the right class.

4 Result

4.1 Dataset and Pre-processing

Public brain tumor data was used from a Kaggle source, consisting of 3,264 brain MRI slices that have been examined and certified by Dr. Shashikani N.Ubhe. From this dataset we acquire 4 types of brain tumors : 926 (Glioma), 937 (Meningioma), 500 (No Tumor), 901 (Pituitary). With a resolution of 350X350 images are in 2D volumes.

To improve model training capabilities we can use Data augmentation. Where the MR Images were resized so that the input image will be fit for SVM and K-NN model (200 X 200).

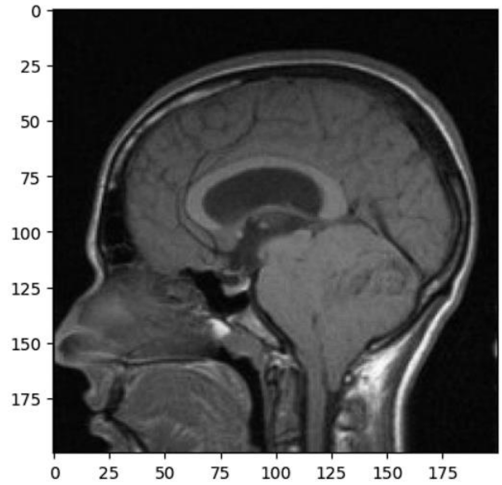

Fig. 3

Fig. 3 One of the example of coronal acquisition views after Resizing.

4.2 Experimental Results and Discussion

In the Results section, we evaluate the performance of the SVM and K-NN algorithms for classifying brain tumors using an MRI dataset that contains both tumor and non-tumor images. The dataset is divided, with 80% used for training and 20% for testing. To measure the system's accuracy, a confusion matrix is generated based on the models' true and false predictions[6][8].

Metrics	SVM	K-NN
Accuracy	97%	93%
Precision	99%	97%
Recall	99%	99%
f1-score	95%	95%

Based on the testing results, it achieved an accuracy of 97% for SVM and 93% for K-NN. SVM has archived the greatest accuracy of 97% and outperforms various measures such as precision. SVM has performed very well and achieved good results for classifying brain tumors.

5 Conclusion

n this study, a comparative analysis of brain tumor classification was performed using the SVM (Support Vector Machine) and K-NN (K-Nearest Neighbor) algorithms. The main objective of this research is to develop an efficient system for classifying brain tumors, aiming to save time and aid radiologists in making accurate diagnoses. Five different models are used to achieve optimal accuracy. In this research, a model was trained using the MRI dataset. This method assists radiologists in quickly identifying and classifying tumors without the need for surgical intervention. The research results show that both classification models, namely SVM and K-NN, provide good performance in classifying MRI images of brain tumors. SVM has a high ability to classify brain images based on the resulting features, while K-NN uses a majority approach from nearest neighbors to classify data. The evaluation is carried out using a confusion matrix, in addition to metrics like accuracy, precision, recall, and F1-score. In this research, pre-processing techniques are used to prepare data before classification. Data augmentation, train-test split, and normalization are the main stages in pre-processing. After that, the features that have gone through the pre-processing stage are used as input to train and test the SVM and K-NN classification models. Pre-processing had a positive impact on all studied models because it applied image scaling and data cleaning [5]. In the evaluation, confusion matrices and other evaluation metrics are used to evaluate the performance of the two classification models. From the evaluation results, it can be determined which type of classification is more optimal in classifying brain tumor images into the right class. This research advances the development of an efficient brain tumor classification system by employing deep learning techniques alongside SVM and K-NN algorithms. This system can help radiologists make fast and accurate diagnoses, thereby speeding up time and increasing the accuracy of brain tumor diagnosis. However, this study still has several limitations. Further research is needed to test this classification model on a larger and more diverse dataset. Apart from that, this research can also be developed by integrating other, more sophisticated methods in classifying brain tumors. Thus, this study provides important insights into the development of a brain tumor classification system that can be used in medical practice to improve patient diagnosis and treatment.

References

[1] S. Arshiya Ara, G. Katti, and A. Shireen, “Magnetic resonance imaging (MRI)-A review.” [Online]. Available:

https://www.researchgate.net/publication/279471369
[2] S. E. Nassar, I. Yasser, H. M. Amer, and M. A. Mohamed, “A robust MRI-based brain tumor classification via a hybrid deep learning technique,” *Journal of Supercomputing*, vol. 80, no. 2, pp. 2403–2427, Jan. 2024, doi: 10.1007/s11227-023-05549-w.
[3] Chinnu A, “MRI Brain Tumor Classification Using SVM and Histogram Based Image Segmentation.” [Online]. Available: www.ijcsit.com
[4] N. Zulpe and V. Pawar, “GLCM Textural Features for Brain Tumor Classification,” 2012. [Online]. Available: www.IJCSI.org
[5] R. Agarwal, S. D. Pande, S. N. Mohanty, and S. K. Panda, “A Novel Hybrid System of Detecting Brain Tumors in MRI,” IEEE Access, vol. 11, pp. 118372–118385, 2023, doi: 10.1109/ACCESS.2023.3326447.
[6] S. Mohsen, A. M. Ali, E. S. M. El-Rabaie, A. Elkaseer, S. G. Scholz, and A. M. A. Hassan, “Brain Tumor Classification Using Hybrid Single Image Super-Resolution Technique With ResNext101_32× 8d and VGG19 Pre-Trained Models,” IEEE Access, vol. 11, pp. 55582–55595, 2023, doi: 10.1109/ACCESS.2023.3281529.
[7] S. Ahmad and P. K. Choudhury, “On the Performance of Deep Transfer Learning Networks for Brain Tumor Detection Using MR Images,” IEEE Access, vol. 10, pp. 59099–59114, 2022, doi: 10.1109/ACCESS.2022.3179376.
[8] B. S. Alemu, S. Feisso, E. A. Mohammed, and A. O. Salau, “Magnetic resonance imaging-based brain tumor image classification performance enhancement,” Sci Afr, vol. 22, Nov. 2023, doi: 10.1016/j.sciaf.2023.e01963.
[9] U. Forest Service - Rocky Mountain Research Station, “kNN vs. SVM: A Comparison of Algorithms.”
[10] W. Lestari and S. Sumarlinda, “Implementation of K-nearest neighbor (KNN) and SUPORT vector machine (SVM) for Clasification Cardiovascular disease,” INTERNATIONAL JOURNAL OF MULTI SCIENCE, https://multisciencejournal.com/index.php/ijm/article/view/227.
[11] U. S. Sudeep, K. N. Naidu, P. S. Girish, T. N. Nikesh, and C. Sunanda, “Brain Tumor Classification using a Support Vector Machine,” International Journal of Computer Applications, vol. 184, no. 28, pp. 15–17, Sep. 2022, doi: 10.5120/ijca2022922347.
[12] E. I. Zacharaki et al., “Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme,” Magn Reson Med, vol. 62, no. 6, pp. 1609–1618, 2009, doi: 10.1002/mrm.22147.