Jurnal E-Komtek
Vol. 10, No. 1. (2026) pp. 87-96
https://jurnal.politeknik-kebumen.ac.id/index.php/E-KOMTEK
p-ISSN : 2580-3719 e-ISSN : 2622-3066

Classification of Magnetic Resonance Imaging (MRI) Images of Brain Tumors Using Convolutional Neural Network (CNN) Methods with VGG-16 and Xception Architectures

Melania Erika Andolo[1], Herman Bedi Agtriadi[2], Meilia Nur Indah Susanti[3]

[1]Department of Computer Science Magister, Institut Teknologi PLN, Indonesia, 12430

melaniaerika2002@gmail.com

https://doi.org/10.37339/e-komtek.v10i1.2974

Published by Politeknik Piksi Ganesha Indonesia

Artikel Info
Submitted:
09-01-2026
Revised:
26-05-2026
Accepted:
30-06-2026
Online first :
30-06-2026

Abstract

Brain tumors are critical medical conditions that require early diagnosis to improve treatment outcomes. Magnetic Resonance Imaging (MRI) is widely utilized for brain tumor detection due to its ability to produce high-resolution images of soft tissues. Nevertheless, manual interpretation of MRI images presents several challenges, including time inefficiency and variability among observers. To address these issues, this study applies the Convolutional Neural Network (CNN) approach using VGG-16 and Xception architectures to classify brain tumor MRI images and to evaluate their performance comparatively. The dataset comprises 2,877 MRI images categorized into four classes: glioma tumor, meningioma tumor, pituitary tumor, and no tumor. Preprocessing stages include resizing images to 224×224 pixels and dividing the dataset into training, validation, and testing sets with a ratio of 80:10:10. Model performance is assessed using accuracy, precision, recall, and F1-score metrics. Experimental results indicate that the VGG-16 architecture achieves an accuracy of 92%, while the Xception architecture records an accuracy of 91%.

Keywords: Brain tumor; MRI; Convolutional Neural Network; VGG-16; Xception

Abstrak

Tumor otak merupakan salah satu penyakit berbahaya yang memerlukan diagnosis sejak dini agar penanganan dapat dilakukan secara optimal. Magnetic Resonance Imaging (MRI) banyak digunakan dalam mendeteksi tumor otak karena mampu menghasilkan citra jaringan lunak dengan tingkat detail yang tinggi. Namun, proses analisis citra MRI secara manual masih memiliki berbagai keterbatasan, seperti membutuhkan waktu yang relatif lama serta adanya perbedaan interpretasi antar tenaga medis. Untuk mengatasi permasalahan tersebut, penelitian ini mengusulkan penerapan metode Convolutional Neural Network (CNN) dengan arsitektur VGG-16 dan Xception dalam melakukan klasifikasi citra MRI tumor otak serta membandingkan kinerja kedua arsitektur tersebut. Dataset yang digunakan berjumlah 2.877 citra MRI yang terbagi ke dalam empat kelas, yaitu tumor glioma, tumor meningioma, tumor pituitary, dan no tumor. Tahapan preprocessing meliputi perubahan ukuran citra menjadi 224×224 piksel serta pembagian data ke dalam data latih, validasi, dan uji dengan rasio 80:10:10. Kinerja model dievaluasi menggunakan metrik akurasi, presisi, recall, dan F1-score. Hasil pengujian menunjukkan bahwa arsitektur VGG-16 memperoleh akurasi sebesar 92%, sedangkan arsitektur Xception mencapai akurasi sebesar 91%.

***Kata-kata kunci**: Tumor otak; MRI; Convolutional Neural Network; VGG-16; Xception*

1. Introduction

A brain tumor is a mass of brain cells that grow abnormally. Most tumors can spread through brain tissue but rarely spread to other areas of the body. However, benign brain tumors can destroy and compress other normal brain tissue, causing paralysis [1]. According to data from the National Cancer Institute, the mortality rate caused by brain tumors is 4.4 per 100,000 individuals each year. This figure is based on cases from 2017–2021 and deaths in 2018–2022, which is 4.3–4.4 per 100,000.

To detect brain tumors at an early stage, magnetic resonance imaging (MRI) is considered the best option due to its high sensitivity. MRI can also provide accurate information due to its high sensitivity [2]. A medical imaging method known as magnetic resonance imaging (MRI) uses magnetic fields and radio waves to produce detailed images of human tissues and organs in three or two dimensions [3]. Due to its ability to display soft tissue with high resolution, MRI has become the primary modality in the diagnosis of brain tumors. However, radiologists face a number of significant problems when they interpret MRI images manually. Interobserver variability can range from 15 to 25 percent. This is especially true for tumors smaller than 2 cm. Additionally, the manual analysis process takes thirty to forty-five minutes per case, which is problematic in healthcare facilities with many employees. In everyday clinical practice, distinguishing low-grade tumors from normal tissue is also often a problem. Deep learning and advances in artificial intelligence (AI) may be able to solve this problem [4].

The most common data mining method is classification. The data classification process involves learning and classification. Classification algorithms are used to analyze training data during learning. During the classification process, data is tested by estimating the accuracy of classification rules. In practical terms, image classification means analyzing image data and returning labels for categorization. The purpose of classification is to distinguish between types of categories and labels that have been created [5]. One of the categories used in the classification in this study is Convolutional Neural Network (CNN).

Convolutional Neural Networks (CNN) consist of two main components: classification and feature extraction. Feature extraction consists of several convolutional layers, followed by activation functions and max-pooling. Classifiers usually consist of fully connected layers. [6]. Convolutional neural networks can analyze features unsupervised, which distinguishes this method from other machine learning methods. They can also classify with a high degree of accuracy because they can handle changes in the input image such as rotation, scale, translation, and scale while reducing the number of free parameters. Image input, convolution, relu, maxpool, fully connected, and softmax are components of convolutional neural networks [7]. CNN has many architectures, such as VGG16, MobileNet, Xception, and InceptionV3, among others **[8]**. In this study, we chose to use the VGG16 architecture, which is one of the best CNN algorithm architectures and is considered one of the best computer vision models. For object classification such as images, the VGG-16 architecture is commonly used because it has many parameters. This architecture has the ability to produce strong and appropriate features. The Xception architecture was used as a comparison in this study because, with its separable depthwise convolution, Xception is known to be capable of extracting features from images well. It has been proven effective in various image classification tasks **[9]**.

The purpose of this study is to determine the accuracy level of applying the Convolutional Neural Network (CNN) architecture model to detect brain tumors in magnetic resonance imaging (MRI) images. Furthermore, this study will evaluate the accuracy level of applying the VGG16 and Xception architecture models to MRI images.

2. Method

2877 data were categorized into four classes: glioma tumors, meningioma tumors, pituitary tumors, and no tumors based on brain tumor MRI image data taken from Kaggle. In the next step, data preprocessing was carried out. Some images were reduced in size to 224 x 224 pixels. Next, the data set was divided into three subsets with a ratio of 80% training data, 10% test data, and 10% validation data. The structure of a Convolutional Neural Network (CNN) model usually consists of several layers. Some layers typically include a pooling layer, a dropout layer, a convolution layer, and a fully connected layer.

2.1 Results of Convolutional Neural Network Model Formation

CNN typically consist of several layers. These usually include convolutional layers, pooling layers, dropout layers, and fully connected layers. To begin modeling the system, this study analyzed object image data. Images are the initial input for the Convolutional Neural Network (CNN). The data used consists of images of brain tumors classified into four categories: glioma tumors, meningioma tumors, pituitary tumors, and no tumors. These images are 224 x 224 pixels in both architectures used. The training data is used to train the model during the evaluation process. If the training results are unsatisfactory, the hyperparameters, data samples, and Convolutional Neural Network layers must be changed. If the model's accuracy improves after adjustment, the test data will be processed to determine the correct classification type.

2.2 Training Stages with the Convolutional Neural Network Algorithm

[10] To explain the stages of the Convolutional Neural Network process, the following steps were taken:

a. Data Collection. This data was collected from Kaggle and divided into four tumor categories: Glioma, Meningioma, Pituitary Tumor, and No Tumor.

b. Calculating Brain Tumor RGB Values. The first step in this process is to convert the image into a numerical format (numbers) to determine the values of the red, green, and blue (RGB) color components. This process is very important for analyzing and understanding the color distribution in the image.

c. Feature Learning. Feature learning has layers that help convert input into usable features. Numbers in vector form, convolutional layers, and extraction layers are among these input features.

   1) Convolutional Layer

      In the convolution process, kernels and steps are used. In this case, the evolution process combines two different matrices to create a new matrix. The goal is to perform convolution on image data or extract features from image input.

   2) Rectified Linear Unit (ReLU) Activation Function

      In this layer, the feature map values are changed at certain intervals according to the ReLU activation function used. By using this feature, we can apply the meaning of the role $f(x)=max(0,x)$ threshold with a zero input image pixel value. In other words, if the pixel value of each image pixel is less than zero, the image will become zero.

   3) Pooling Layer

      Pooling reduces the size of each feature map stack. This is typically done with a 2x2 filter, which is applied in two steps and acts on each slice of the input.

d. Classification

   Each neuron with previously extracted characteristics, such as curvature, full connectivity, and softmax activation, can be classified more easily by this layer [11].

1) Flatten

The flatten matrix operation converts it into a one-dimensional vector. This converts the map features obtained from the previous layer into a one-dimensional vector, which allows the map features to be classified with a fully connected layer and softmax.

2) Fully Connected Layer

Fully connected layers are commonly used when applying dimensionality reduction, which allows data to be grouped into linear form. The one-dimensional data of each neuron in the convolutional layer must be converted before being transferred to the fully connected layer. In this way, the actual and unalterable information from the data is lost, and only the fully connected layer can be described at the end of the network.

3) Softmax

In the output layer, the Softmax activation function is used. This is because the output layer has many similarities with the fully connected layer, which distinguishes the two layers. This shows that the ReLU activation function is used in the fully connected layer and the Softmax activation function is used in the output layer.

2.3 *Evaluation* Model

Model evaluation is used to ensure that the model has reliable performance. Its performance is assessed using accuracy, precision, recall, and F1 score metrics [12].

$$Akurasi = \frac{Total\ Data}{Jumlah\ Prediksi\ Benar} \quad (2.1)$$

Accuracy is a good basic metric for measuring model performance. Accuracy is a measure of how many correct predictions a model makes on a complete test data set. If the data set is unbalanced, accuracy will be lower. Meanwhile, to measure the level of prediction precision, you can use the equation :

$$Presisi\ Prediksi = \frac{TP}{TP+FP} \quad (2.2)$$

Predictions are usually made using precision calculations. In addition, you can also see the system's performance in more detail, including the recall or sensitivity of the system to classes. Recall calculations can be done using the equation :

Recall = $TP+FNTP$ **(2.3)**

The F1 score is the harmonic mean of precision and recall, which provides a combined view of both metrics, with the equation :

$$F1 - Score = 2 \times \frac{Precision\ x\ Recall}{Precision + Recall} \quad (2.4)$$

The best F1-Score is 1.0, and the worst score is 0. In other words, a good score indicates that the classification model we have created has good precision and recall.

2.4 Research Design

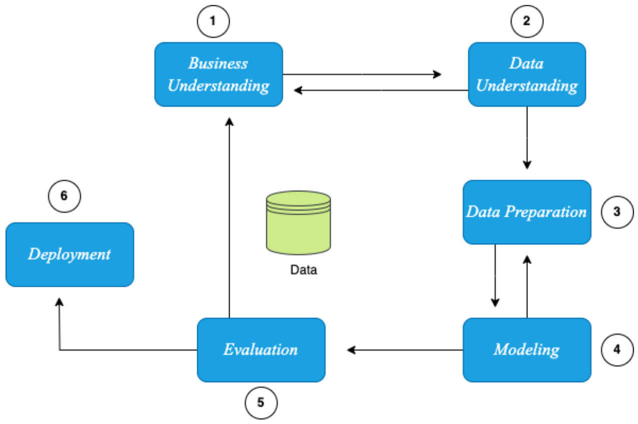


Figure 1. Research Design

The first step is to understand the business, which analyzes the problems to be discussed in the data collection process. The data to be used in this study will be collected using MRI image data generated by Kaggle, totaling 2877 data. After the data is collected, the data preparation process consists of a pre-processing process, which means resizing images and splitting the data into instructions, tests, and validations. After the pre-processing process is complete, the data is created using the CNN (Convolutional Neural Network) technique using the Xception and VGG16 architecture models. After the data is processed, performance testing must be carried out to ensure that it achieves the expected objectives. The process continues to deployment after evaluation and the model is declared correct. After the use of machine learning and technology, the deployment stage ends with the writing of a thesis.

3. Results and Discussion

With the number of brain tumor patients increasing every year, challenges are emerging, such as high consultation costs and a shortage of experienced medical personnel. In the medical field, the use of technology is becoming increasingly important to address these issues. Deep learning is one promising solution. Deep learning can help doctors and radiologists make better diagnoses by looking at medical images.

a. Training Data

Training Data is a collection of data gathered through Kaggle, used to train and build models.

b. Preprocessing

The purpose of the preprocessing stage, which is the initial stage in the structure of this research system, is to improve the images produced from MRI results into better images than those produced in the previous stage. This will make the next step easier to perform. This stage includes resizing to change the image size and data separation to divide the data into two or more parts.

c. Convolutional Neural Network Modeling

The VGG16 model is a CNN architecture that has 16 trained layers, consisting of 13 convolution layers with a kernel size of 3×3 and 3 fully connected layers. Meanwhile, Xception is used as a comparison model with a separable depthwise convolution approach, which begins with a standard convolution layer and continues with 36 separable depthwise convolution modules. Test results show that VGG16 achieves an accuracy rate of 92%, while Xception achieves an accuracy rate of 91%. The Precision, Recall, and F1-Score values for both architectures also show good performance.

d. CNN Model Training Chart

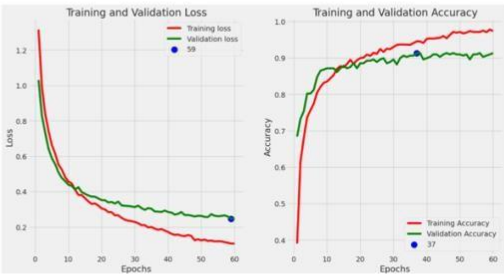
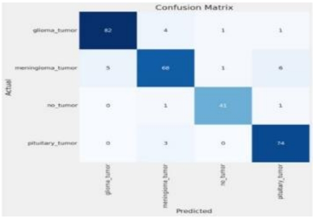


Figure 2. VGG16 Training Graph

Figure 3. Xception Training Graph

Figure 2 shows the accuracy and loss graphs of the VGG16 model, and Figure 3 shows the accuracy and loss of the Xception model obtained after the training iteration process. The values shown on the X-axis indicate the number of epochs, and the Y-axis indicatesthe accuracy level, which is the percentage of correct predictions out of the total predictions. The accuracy of VGG16 is 92%, while that of Xception is 91%.

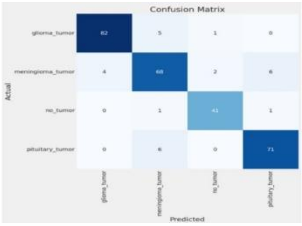
e. Classification Stages



Classification Report:

	precision	recall	f1-score	support
glioma_tumor	0.94	0.93	0.94	88
meningioma_tumor	0.89	0.85	0.87	80
no_tumor	0.95	0.95	0.95	43
pituitary_tumor	0.90	0.96	0.93	77
accuracy			0.92	288
macro avg	0.92	0.92	0.92	288
weighted avg	0.92	0.92	0.92	288

Figure 4. VGG16 Model Testing Results



Classification Report:

	precision	recall	f1-score	support
glioma_tumor	0.95	0.93	0.94	88
meningioma_tumor	0.85	0.85	0.85	80
no_tumor	0.93	0.95	0.94	43
pituitary_tumor	0.91	0.92	0.92	77
accuracy			0.91	288
macro avg	0.91	0.91	0.91	288
weighted avg	0.91	0.91	0.91	288

Figure 5. Xception Model Testing Results

For the testing process, the model was tested on test data, then tested using confusion matrix calculations, as illustrated in Figures 5 and 6 above. A total of 2877 datasets were used for testing, and the results showed that the accuracy of the confusion matrix classification test performance in object recognition on VGG16 was 92% and on Xception was 91%.

f. Interface Testing Stage

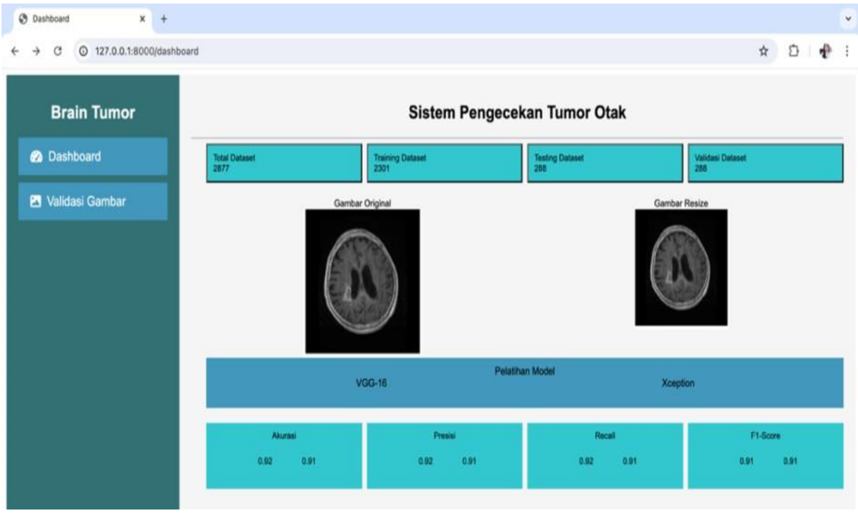


Figure 6. Dashboard Display

This dashboard page explains information about various classification processes, including image resizing, data partitioning, modeling processes, accuracy and loss graphs, and confusion matrices.

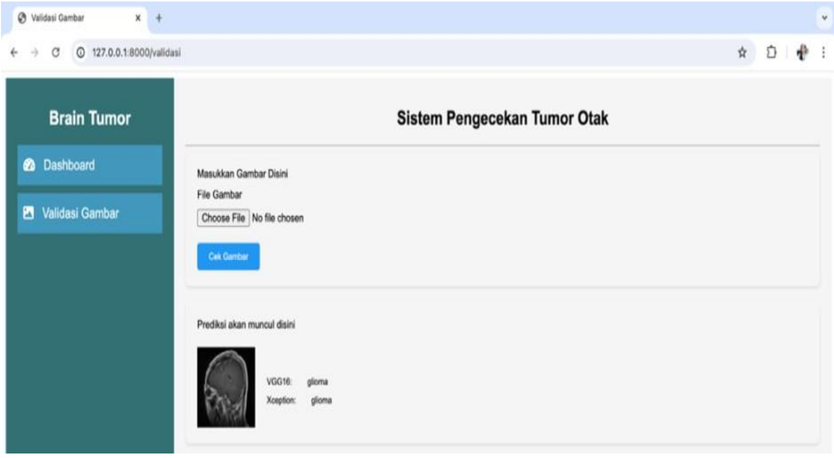


Figure 7. Validation Display

On this image validation page, enter the data (MRI image). Then, click the "check image" button, which will display the image and predict the type of brain tumor on both architectures, namely VGG-16 and Xception.

4. Conclusion

The process of classifying brain tumor MRI images using the Convolutional Neural Network (CNN) method is carried out in several main stages. The initial stage is preprocessing, which includes image resizing and dataset division. After the data is divided, the convolution process is carried out by applying a sharpening kernel to improve image sharpness. The convolution results are then processed using the ReLU activation function, followed by the addition of a MaxPooling2D layer to perform pooling.

The performance of two CNN architectures, VGG16 and Xception, is evaluated using a confusion matrix. The evaluation results show that both models produce excellent accuracy rates, with a score of 92% for VGG16 and 91% for Xception. In addition to accuracy, other evaluation metrics such as precision, recall, and f1-score were also used to support the model performance analysis. The VGG16 model obtained a precision value of 92%, a recall of 92%, and an f1-score of 91%, while the Xception model achieved values of 91% for precision, recall, and f1-score, respectively.

For further research, it is recommended to use different brain tumor MRI image datasets, including unlabeled data, to test and improve the model's generalization ability in more diverse conditions.

References

[1] L. Widya Astuti, "Ekstrasi Fitur Citra MRI Otak Menggunakan Data Wavelet Transform (DWT) untuk Klasifikasi Penyakit Tumor Otak", [Online]. Available: http://med.harvard.edu/AANLIB/.

[2] I. B. L. M. Suta, M. Sudarma, and I. N. Satya Kumara, "Segmentasi Tumor Otak Berdasarkan Citra Magnetic Resonance Imaging Dengan Menggunakan Metode U-NET," *Majalah Ilmiah Teknologi Elektro*, vol. 19, no. 2, p. 151, Dec. 2020, doi: 10.24843/mite.2020.v19i02.p05.

[3] D. Husen, "Klasifikasi Citra MRI Tumor Otak Menggunakan Metode Convolutional Neural Network," *bit-Tech*, vol. 7, no. 1, pp. 143–152, Aug. 2024, doi: 10.32877/bt.v7i1.1576.

[4] M. Z. Basri, Muhammad Gabriel Somoal, and Rafiq Setyo Aji, "Deteksi Tumor Otak pada Citra Magnetic Resonance Imaging (MRI) Menggunakan Arsitektur MobileNet dengan Optimizer Adam," *JEKIN - Jurnal Teknik Informatika*, vol. 5, no. 2, pp. 539–550, Jun. 2025, doi: 10.58794/jekin.v5i2.1394.

[5] P. Rosyani, S. Saprudin, and R. Amalia, "Klasifikasi Citra Menggunakan Metode Random Forest dan Sequential Minimal Optimization (SMO)," *Jurnal Sistem dan Teknologi Informasi (Justin)*, vol. 9, no. 2, p. 132, Apr. 2021, doi: 10.26418/justin.v9i2.44120.

[6] E. Hari Rachmawanto, "Convolutional Neural Network (CNN) untuk Klasifikasi Citra Penyakit Diabetes Retinopathy," *SKANIKA: Sistem Komputer dan Teknik Informatika*, vol. 5, no. 2, pp. 167–176, 2022.

[7] M. Setiono, "Klasifikasi Penyakit Antraknosa Citra Cabai Rawit Dengan Metode Convolutional Neural Network (CNN)," *Hal; Jl. Jembatan Merah No.84 C Gejayan Yogyakarta*, vol. 11, no. 2, 2024.

[8] I. Elzandy *et al.*, "KLASIFIKASI VARIETAS ANGGUR BERDASARKAN CITRA DAUN MENGGUNAKAN VGG16-XGBOOST," 2025.

[9] D. Prasetia, B. Hanggara, R. Naufaldihanif, A. Clark Peter Wijaya, and A. Ibrahim, "KLASIFIKASI KONDISI KEBERSIHAN SUNGAI DARI CITRA DIGITAL MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK DENGAN ALGORITMA INCEPTION V3 DAN XCEPTION," 2025.

[10] K. Azmi, S. Defit, and U. Putra Indonesia YPTK Padang Jl Raya Lubuk Begalung-Padang-Sumatera Barat, "Implementasi Convolutional Neural Network (CNN) Untuk Klasifikasi Batik Tanah Liat Sumatera Barat," vol. 16, no. 1, p. 2023.

[11] A. Yusuf, R. Cahya Wihandika, and C. Dewi, "Klasifikasi Emosi Berdasarkan Ciri Wajah Menggunakan Convolutional Neural Network," 2019. [Online]. Available: http://j-ptiik.ub.ac.id

[12] R. Rizkyatul Basir *et al.*, "EVALUASI KINERJA MODEL KLASIFIKASI PEMBELIAN SEPEDA MENGGUNAKAN ALGORITMA RANDOM FOREST," 2025, doi: 10.1007/s11749-016.