Evaluating ResNet Acrchtectures for Brain Tumor Classification Based on MRI Images

Szymon Kruszyk, Damian Ledziński and Sandra Śmigiel

June 10, 2024

Szymon Kruszyk[1,*], Damian LEDZIŃSKI, Sandra ŚMIGIEL

1. EVALUATING RESNET ARCHITECTURES FOR BRAIN TUMOR CLASSIFICATION BASED ON MRI IMAGES

1.1. Introduction

In the digital age, advancements in medical technology and informatics have opened new horizons in the diagnosis and treatment of diseases. One promising direction is the application of machine learning methods to analyze medical images, particularly MRI scans, for the detection and classification of brain tumors. Brain tumors are among the most complex and unpredictable oncological diseases. Rapid and accurate diagnosis is crucial for planning effective treatment and increasing patients' survival chances. Traditional diagnostic methods rely on the assessment of MRI images by qualified radiologists, which is time-consuming and subject to subjective interpretation. The development of machine learning technology, especially deep neural networks, has brought new possibilities for automating and supporting the diagnostic process.

This study focuses on reviewing the latest advancements in machine learning applied to MRI image analysis, identifying the most commonly used models and techniques such as convolutional neural networks (CNNs), and evaluating their effectiveness. Real-world data is analyzed, selected models are applied to specific image datasets, and their performance is assessed in the context of brain tumor diagnostics.

[1] Corresponding author: szymon.kruszyk@gmail.com

1.2. Transfer Learning in Brain Tumor Classifiation

Transfer learning has become a pivotal technique in the domain of brain tumor classification, offering significant advancements in the accuracy and efficiency of diagnostic models. This approach involves leveraging pre-trained models, initially developed on large, diverse datasets, and fine-tuning them for specific tasks such as classifying brain tumors from MRI images. The versatility and effectiveness of transfer learning have made it a popular choice in medical imaging applications.

1.2.1. Concept and Benefits of Transfer Learning

Transfer learning is a machine learning technique that allows a model trained on one problem to be adapted for another, related problem. This method is particularly beneficial when dealing with limited data, which is a common scenario in medical imaging. The main advantages of transfer learning in brain tumor classification include:

- Reduced Data Requirements: Transfer learning mitigates the need for extensive labeled datasets by utilizing knowledge gained from large, generic datasets. This is particularly useful in medical fields where labeled data can be scarce and expensive to obtain.
- Improved Training Efficiency: Fine-tuning a pre-trained model requires significantly less computational power and time compared to training a model from scratch. This efficiency makes transfer learning an attractive option for developing diagnostic tools.
- Enhanced Performance: Models benefit from pre-learned features that capture essential patterns in images, improving their performance on specific tasks such as tumor detection and classification.

1.2.2. Application in Brain Tumor Classification

In brain tumor classification, transfer learning typically involves using convolutional neural networks (CNNs) pre-trained on large image datasets like ImageNet. These models are then fine-tuned using MRI images of brain tumors. The process generally includes retraining the final layers of the pre-trained models to adapt them to the specific characteristics of MRI images.

Several studies have demonstrated the effectiveness of transfer learning in this field. For example, the use of models like AlexNet, GoogLeNet, and ResNet has shown high accuracy rates in distinguishing between benign and malignant brain tumors. In a notable study, AlexNet achieved an accuracy of 99.04% in classifying MRI images of brain tumors, showcasing the potential of transfer learning to significantly enhance diagnostic accuracy[1]

1.2.3. Methodologies and Approaches

- Fine-Tuning Pre-Trained Models: Models pre-trained on large datasets are fine-tuned with brain tumor MRI images to improve their specificity to medical imaging tasks. This approach has been successful in several studies, achieving high accuracy and sensitivity.
- Hybrid Models: Combining multiple pre-trained models to leverage the strengths of each can result in more robust classification systems. For instance, using both EfficientNet and ResNet50 for feature extraction and then combining these features has proven to be effective.[2]
- Data Augmentation: To further improve the performance of transfer learning models, data augmentation techniques such as rotation, flipping, and scaling of images are used. This helps in simulating a larger dataset, allowing the models to generalize better.

Table 1

Experimental models' parameters

	1D model	2D model
Inner electrolyte diffusion coefficient	$1 * 10^{-3}\ [mm^2 * s^{-1}]$	$0.8 * 10^{-3}\ [mm^2 * s^{-1}]$
Outer electrolyte diffusion coefficient	$1 * 10^{-3}\ [mm^2 * s^{-1}]$	$1.2 * 10^{-3}\ [mm^2 * s^{-1}]$
Solubility constant	$3.5 * 10^{-5}\ [mol^2 * dm^{-6}]$	$3.5 * 10^{-3}\ [mol^2 * dm^{-6}]$
Aggregation rate constant	$1 * 10^{1}\ [dm^6 * mol^{-2} * s^{-1}]$	$1 * 10^{1}\ [dm^6 * mol^{-2} * s^{-1}]$
Experiment time	$2 * 10^{6}\ [s]$	$86400\ [s]\ (24\ h)$

1.3. Results and discussion

Using parameters described in Table 1, two models (1D and 2D) was tested and verified against Matalon-Packter law (spacing law), width law and time law. Both models are suitable to be good description of diffusion-precipitation-agglomeration system.

1.3.1. Methodology

Fig. 1. Proces treningu i testowania
Rys. 1. Training and testing process

Data Collection

The image database used for training the model consists of 3,096 images, each with a resolution of 256x256 pixels. This dataset is derived from another database, in which data augmentation and normalization were performed, including removing duplicate data, normalizing to grayscale, enhancing image quality, and improving comparability. The images come from different patients, but the exact number is unknown due to a lack of data. The images are categorized into four classes representing different brain conditions: glioma_tumor (MRI images with glioma), meningioma_tumor (MRI images with meningioma), pituitary_tumor (MRI images with pituitary tumors), and normal (images of healthy brains). The dataset is not pre-divided into training and test sets, requiring division at the code level. MRI images are presented in various cross-sections and stages of the examination.

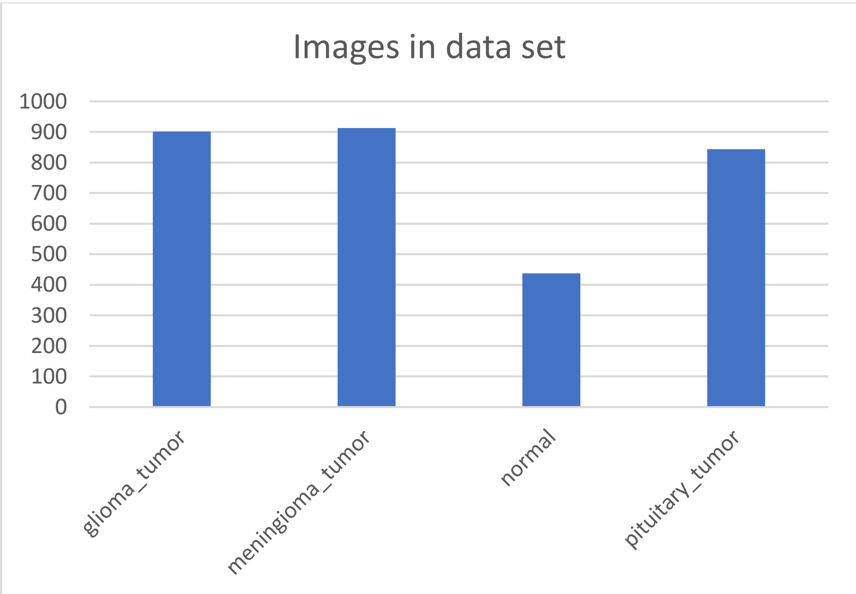

Fig. 2. Obrazy w data secie
Rys. 2. Images in data set

Model Selection

Three variants of the Residual Network (ResNet) model were selected for comparison: ResNet-50, ResNet-101, and ResNet-152. These models were chosen due to their proven effectiveness in image classification tasks and to compare three different depth models.

Training and Evaluation

All images were loaded into the models at a uniform resolution of 256x256 pixels with standard RGB color coding. Data was split into training and testing sets in an 80-20 ratio using the "train_test_split" function. Training data underwent augmentation, with additional techniques for the normal class to balance the dataset sizes.

The model was tested using the "categorical-crossentropy" loss function and the Adam optimizer with a learning rate of 0.0001, which was a balanced choice to ensure effective learning without overfitting. The experiment was conducted over 10 epochs, which was sufficient to stabilize the model while avoiding overfitting. The ResNet architecture was topped with a sequence of layers including GlobalAveragePooling2D, two Dropout layers (each with a 0.2 probability), a dense layer with 128 neurons and "relu" activation, and a final dense layer with four neurons and "softmax" activation for the classification of four types of brain tumors.

The configuration effectively combined the deep features extracted by ResNet with task-specific classification layers. The classification results were evaluated in terms of accuracy and loss on both the training and validation sets, allowing continuous monitoring of the model's progress. Additionally, the use of confusion matrices, precision, recall, and F1-score provided a comprehensive assessment of the model's performance in disease classification.

Tools and Technologies

The study employed Python programming language with TensorFlow and Keras libraries for model implementation. Data augmentation techniques such as rotations, horizontal and vertical shifts, zoom, and shearing were applied to enhance the training dataset and improve model generalization. Scikit-learn and Matplotlib libraries were used for models assessment.

1.3.2. Results

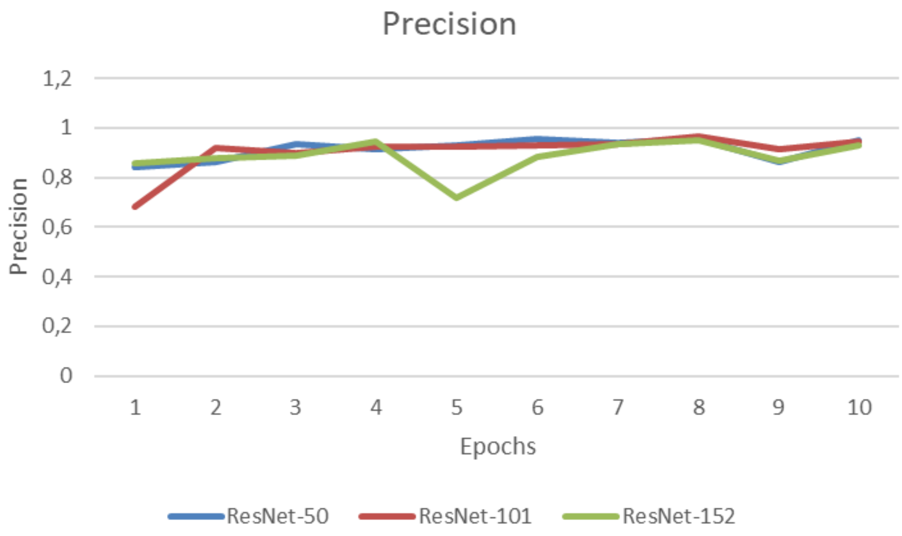

Fig. 3. Precision
Rys. 3. Precyzja

Table 2

Assessment metrics values

Metric/model	ResNet-50	ResNet-101	ResNet-152
Validation accuracy	0,95	0,95	0,93
Precision	0,95	0,95	0,93
Sensitivity	0,95	0,95	0,93
F1-Score	0,95	0,95	0,93
Avg training loss	0,09	0,09	0,09
Avg validation loss	0,29	0,3	0,42
Avg training time[s]	316,7	496,6	872,7

- ResNet-50: Achieved a validation accuracy of 0.95, with high precision and sensitivity. Training time was optimal, making it the most efficient model in this study.
- ResNet-101: Also achieved a validation accuracy of 0.95, with similar precision and sensitivity to ResNet-50. However, the training time was significantly longer.
- ResNet-152: Showed slightly lower validation accuracy at 0.93. The model exhibited increased training and validation loss, indicating overfitting and reduced generalization.

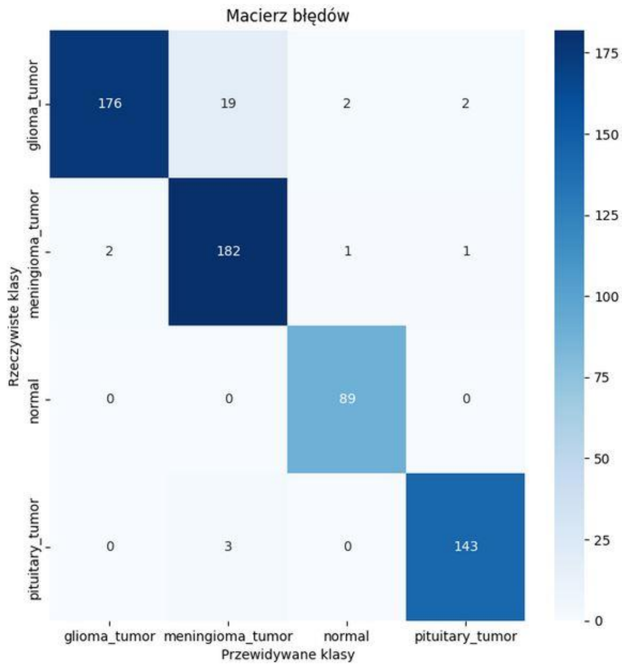

Fig. 4. ResNet-50 error matrix
Rys. 4. Matryca błędów dla ResNet-50

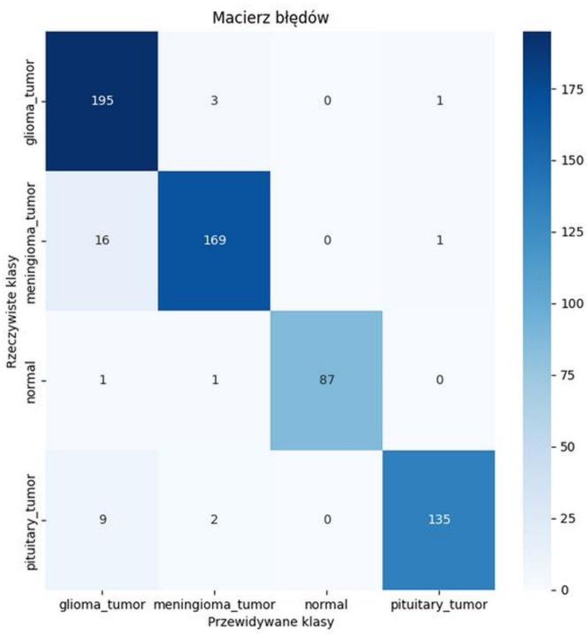

Fig. 5. ResNet-101 error matrix
Rys. 5 Matryca błędów dla ResNet-101

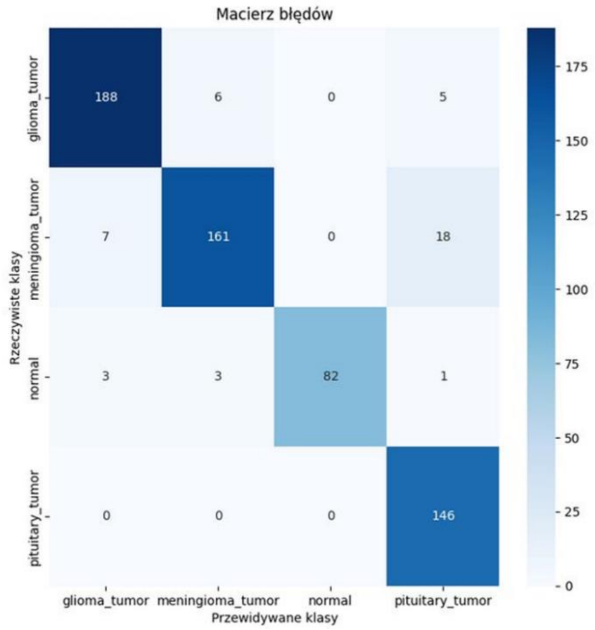

Fig.6. ResNet-152 error matrix
Rys. 6. Matryca błędów dla ResNet-152

The model shows improved training accuracy with each epoch, indicating effective learning; however, irregular validation accuracy and rising validation loss suggest potential overfitting. For instance, with the ResNet-50 model, the confusion matrix revealed high classification accuracy: 176 correct glioma tumor identifications, 182 for meningioma tumors, 89 for normal tissue, and 143 for pituitary tumors. Misclassifications included glioma tumors being mistaken for meningioma tumors 19 times and vice versa, indicating challenges in distinguishing these types. Similarly, the ResNet-101 model showed 195 correct glioma classifications with 3 errors, 169 for meningioma with 16 errors, and accurate normal and pituitary classifications, with some misclassifications suggesting possible dataset categorization issues. The ResNet-152 model had consistent training accuracy growth but fluctuating validation accuracy, with notable misclassifications in normal tissue and glioma and meningioma categories, suggesting potential dataset issues. Overall, the models exhibit strong performance but highlight the need for further optimization and dataset refinement.

1.3.3. Discusison

The study demonstrated that ResNet-50 and ResNet-101 are highly effective for brain tumor classification, achieving high accuracy and performance metrics. However, ResNet-50 stands out due to its balance of accuracy and computational efficiency. ResNet-152, while deeper, did not perform as well, suggesting that increased model complexity does not necessarily translate to better performance for this task. The findings highlight the importance of model selection based on the specific requirements of medical image analysis, such as computational resources and time constraints.

Bibliography

1. R. Mehrotra, M. A. Ansari, R. Agrawal, i R. S. Anand, „A Transfer Learning approach for AI-based classification of brain tumors”, *Machine Learning with Applications*, t. 2, s. 100003, grudz. 2020, doi: 10.1016/J.MLWA.2020.100003
2. M. Aamir i in., „A deep learning approach for brain tumor classification using MRI images”, Computers and Electrical Engineering, t. 101, s. 108105, 2022, doi: 10.1016/j.compeleceng.2022.108105.

EVALUATING RESNET ARCHITECTURES FOR BRAIN TUMOR CLASSIFICATION BASED ON MRI IMAGES

OCENA ARCHITEKTURY RESNET DLA KLASYFIKACJI GUZÓW MÓZGU NA PODSTAWIE OBRAZÓW MRI

Abstract

In the rapidly advancing field of medical image analysis, the selection of an efficient and accurate model is critical for timely and reliable diagnosis. This study aimed to evaluate the performance of different deep learning architectures by comparing three variants of the Residual Network (ResNet) model—ResNet-50, ResNet-101, and ResNet-152—in the classification of brain tumors using MRI images.

Each model was assessed on key performance metrics including validation accuracy, precision, sensitivity, and F1-Score. The results revealed that ResNet-50 and ResNet-101 both achieved a score of 0.95, outperforming ResNet-152, which scored 0.93. Although all models showed similar average training losses, an increase in validation loss with model depth suggested a decline in generalization capability for ResNet-152. Furthermore, the analysis documented a rise in training time with the complexity of the model, highlighting the greater computational requirements of the more sophisticated architectures. ResNet-50 was identified as the optimal model due to its balance between accuracy and computational efficiency, making it the preferred choice for the classification of medical images when resources and time are limited.

In summary, while all tested models displayed high performance, ResNet-50 offers the best combination of accuracy and efficiency, proving to be the most practical model for medical image classification in resource-constrained settings.

Key words: ML, Machine learning, MRI, classification, brain tumors, CNN, ResNet, Transfer Learning, TL