

# Benchmarking Deep Learning Models for Automated MRI-based Brain Tumor Detection: In-Depth Analysis of CNN, VGG16, VGG19, ResNet-50, MobileNet, and InceptionV3

Sarah Pendhari  
Thadomal Shahani Engineering  
College  
University of Mumbai  
Mumbai, India

Nazneen Pendhari  
M.H. Saboo Siddik College of  
Engineering  
University of Mumbai  
Mumbai, India

Sonal Shroff  
Thadomal Shahani Engineering  
College  
University of Mumbai  
Mumbai, India

## ABSTRACT

The early and precise diagnosis of brain tumors is paramount in the medical field, significantly impacting treatment efficacy and patient survival rates. Magnetic Resonance Imaging (MRI), a non-intrusive diagnostic tool, is extensively utilized for identifying brain tumors, eliminating the need for invasive biopsies. However, the manual interpretation of MRI scans is a challenging and laborious task due to the voluminous and complex nature of the three-dimensional images it generates. So this study harnesses the power of advanced state-of-the-art deep learning models – Convolutional Neural Network (CNN), VGG16, VGG19, ResNet-50, Mobile Net and InceptionV3 – to automate and enhance the accuracy of brain tumor detection from MRI data. The findings demonstrate a marked improvement in the detection and accuracy of brain tumors, showcasing the potential of deep learning in revolutionizing medical imaging diagnostics with accuracy scores of CNN being 97.55%, VGG16 being 97.96%, VGG19 being 97.55%, ResNet-50 being 90.20%, MobileNetV1 being 96.73% and InceptionV3 being 97.55%. This innovation not only streamlines the diagnostic process but also opens new avenues for early intervention strategies, ultimately contributing to improved patient outcomes in the fight against one of the most daunting health challenges of all time.

## General Terms

Computer Vision, Image Processing, Brain Tumor Detection

## Keywords

Medical Imaging, Brain Tumor Detection, Automated Diagnosis, Convolutional Neural Network (CNN), Deep Learning, Visual Geometry Group (VGG16, VGG19), Residual Networks (ResNet-50), MobileNet, InceptionV3

## 1. INTRODUCTION

Cancer continues to be a predominant global health challenge, with the World Health Organization reporting nearly 10 million [1] fatalities in 2020 alone. Brain tumors, a subset of cancers, are particularly alarming due to their direct impact on neurological functions. Brain tumors are classified as cancerous (also known as malignant) and non-cancerous (also known as benign). Benign tumors in the brain are non-progressive in nature and don't show growth in the other parts

of the body. On the other hand, malignancy spreads rapidly over other body parts.

Further, there are two main types of malignant tumors: Primary tumors affect the brain first and then spread over the other body parts whereas, secondary tumors are developed in other body parts and then reach to the brain [1-3]. Other kinds of brain tumors like meningioma, pituitary, and glioma too show high occurrence rate [1-3,8]. Pituitary gland and glioma tumors can either be malignant or benign. But, meningioma tumors are always malignant as they develop in the vicinity of the brain and spinal cord. Abnormal growth of cells results in Pituitary tumors lying in the pituitary glands in the vicinity of the brain, whereas in the glial cells of the brain, gliomas are formed. Tumor classification into various classes like meningioma, glioma, or pituitary is a complex process due to variation in their size, shape and intensity [1-4]. Subsequent growths due to metastasis can lead to these tumors. Their anatomical locations and various growth rates have substantial influence on neurological consequences. Hence, they require accurate and tailored methods for diagnosis and treatment. The field of medical diagnosis has a significant impact on recent developments in machine learning, namely in the domain of deep learning. To identify and classify medical imaging patterns, including brain malignancies, Conventional Neural Networks (CNNs) have been proved to be more effective [1-20]. These techniques facilitate timely identification and provide a non-intrusive alternative for traditional biopsy techniques. The objective is to improve accuracy in diagnosis of brain tumors with the help of deep learning models and to increase the classification precision. This methodology is significantly effective in detection of brain tumors by various algorithms like CNN, VGG16, VGG19, ResNet-50, MobileNet, and InceptionV3[5]. This study makes a significant impact in the field by introducing a novel approach that overcomes the past limitations and presents an efficient and dependable system for brain tumor classifications. The aim is to provide effective and practical solutions to the constraints and issues in the previous research. Six modeling methods are analyzed and assessed to determine the variations in their effectiveness.

## 2. METHODOLOGY

The whole process is performed using Keras with the Tensorflow as backend[6]. The networks are designed in the Python environment and are run by cross-library in Google Collaboratory. The detail of Collaboratory hardware technical

characteristic is as follows: The usage of GPU was a Tesla K80, compute 3.7 having CUDA cores and the CPU configured single core hyper threaded Xeon Processors 1 core, 2 threads with the RAM of 12.7GB and Disk space with 53.8GB was available. The steps accomplished in this study to achieve the goal are in Figure 1., indicating the stages of the recommended methodology described further in this paper.

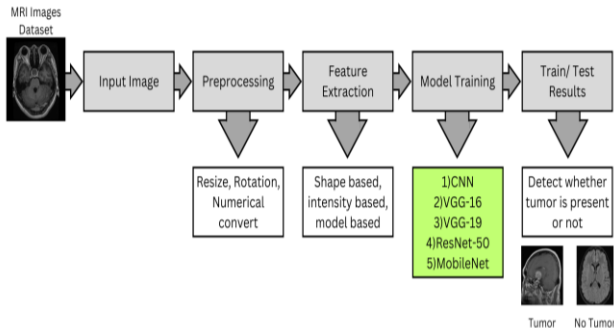


Fig 1: Proposed Methodology

## 2.1 Dataset

In this study, the brain tumor classification dataset used is available at <https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri>. It comprises 3264 T1-weighted contrast-enhanced MRI images. The brain images used from the dataset of the categories ‘no tumor’ and ‘glioma\_tumor’. A sample of data that has been used in this paper prior to preprocessing is shown in Figure 2.

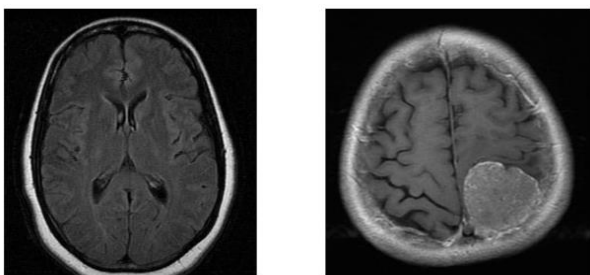


Fig 2 : No tumor and tumor images

## 2.2 Pre-Processing

The size of data set images and that of network’s input must match in order to train the network and generate predictions on new data. Therefore, the size of the images is adjusted to fit the network and then data is scaled or cropped. 0.1 is the zoom range, horizontality is the rotation, and 0.5 and 1.0 are the brightness ranges respectively, and 0.1 is the preprocessing range for rescaling, width and height shift [1,5,6,11]. Then, 80% dataset is used for training and 20% for validation [1,6].

## 2.3 Feature Extraction

In this project, convolutional neural networks (CNNs) are utilized to extract a diverse array of features from MRI images, crucial for accurate diagnosis. Initially, the CNN focuses on low-level features, like edges, contours, and textures, which are fundamental in delineating shape and boundary of brain tumors. Intermediate features like shapes, structures, and contrast variations are also identified in the later stages of the CNN. At each of the image levels, the features are more abstract, but they are more differentiated. In the case of texture, in the initial layers, the features are often represented by patches and shapes that are related to the colors, gradients, etc. It is useful when differentiating intensity, which is significant

when distinguishing tumors from the normal brain. In turn, at the medium levels of convolution, shapes, structures and differences of contrast are identified. The final layers of the CNN examine the abstract features, namely complex “non-visualizing”, and the features of high magnitude. As a result, the categories of things and the multiple scale features are determined. Characteristics related to anatomy, attributes of the tissue and the changes occurring in the image are identified. Nothing that all the characteristics are established by means of the texture and grazing is shown. The latter is explained by the importance of the textural appearance of an image on the computer display of the tomography. Thus, above was given the detailed description of how the features at all levels of the affected image and at each of the identified levels of convolution become generalized. This strategy ensures their thorough and comprehensive identification, which allows the realization and more detailed differentiation, that is, contributes to the more accurate determination of the presence or absence of tumors in the brain depending on their course.

## 2.4 CNN

The CNN model has multiple convolutional layers followed by max pooling, which allows the network to capture and abstract complex features from the brain scans. They work to detect various aspects of the images, such as edges and textures that are essential for distinguishing between healthy tissue and tumor affected tissue. Batch normalization that the model’s representation remains stable across different images with varying position of the tumor was also done. During the training process, certain neurons are randomly deactivated by dropout layers in order to reduce the risk of overfitting. Thus, enhancing the ability for the model to generalize its learning to new, unseen data.

During the training phase, which spanned 10 epochs, it was observed that there was a continual improvement in accuracy and a gradual decrease in loss on the training set, which shows efficient learning by the model. The model used was Sequential and the batch size was 32. The Adam optimizer was chosen for the model having a learning rate of 0.0001[1-8]. The learning rate is kept constant in all the model implementations. Adam, which stands for Adaptive Moment Estimation, uses the combined advantages of two popular optimizers: RMSProp and AdaGrad. Adam calculates adaptive learning rates for each parameter. An activation function of ReLu(Rectified Linear Unit) has been used. For the loss function calculation, the categorical\_crossentropy has been used. Meanwhile, the validation accuracy and loss provided an insight into the model’s performance on data it had not previously encountered.

## 2.5 VGG-16

The VGG-16-based model utilizes a pre-trained VGG16 network, a renowned architecture known for its depth and effectiveness in image classification tasks. Unlike our initial CNN model, VGG16 comes with a series of convolutional layers pre-configured and trained on a vast dataset (ImageNet). In adapting VGG16, modification of the base network to suit the binary classification task – identifying the absence or presence of a brain tumor was done. This was achieved this by appending additional layers to the pretrained VGG16 model, which was loaded with its original ImageNet weights. The model’s top layers, originally designed for 1000-class classification, were replaced with a Flatten layer to reshape the output for dense layers, and a Dense layer with 64 nodes (using ReLU activation) for further feature interpretation. A Dropout layer (0.5) randomly disables

neurons to mitigate overfitting risk and thus enhances the model's generalization capabilities.

The model's final layer is dense having the softmax activation function and is updated to output two classes, corresponding to 'tumor present' or 'no tumor'. Adam optimizer is used to compile the model and categorical cross entropy is used as the loss function to maintain consistency with our initial CNN model.

## 2.6 VGG-19

The VGG19 model is a deep neural network having 19 layers, including convolutional layers. The ImageNet dataset, which includes a diverse array of images, provides a robust foundational knowledge for feature extraction forms the base of pre-training of VGG-19. For our task, the VGG19 model is initialized with its pre-trained weights and configured without its top layers (using include\_top=False), for our binary classification problem. The input tensor is adjusted to accommodate the size of our MRI images (100x100 pixels). Once the model is loaded, freeze its layers (layer.trainable = False) to prevent them from being updated during training. This step ensures that the rich feature extraction capabilities learned from ImageNet are preserved.

Next, a Flatten layer is extra-added for 2-D to 1-D vector conversion. This is followed by a Dense layer having 64 neurons and a ReLU activation function which becomes a fully connected layer. A Dropout layer(0.5) prevents overfitting. The final layer was a Dense layer having a softmax activation function responsible to give the classification as 'tumor' or 'no tumor'. Adam optimizer and categorical cross entropy were used here as well, for a fair comparison.

## 2.7 RESNET-50

On the ResNet-50 base, a GlobalAveragePooling2D layer was added. This facilitates the dimensionality reduction of the feature maps, condensing the spatial information. This reduces the computational load and the number of parameters in the network. Then a fully connected Dense layer with 256 neurons and ReLU activation is added. This layer serves as a classifier that processes the features extracted by the ResNet-50 model. The final layer is another Dense layer having softmax activation function. Adam optimizer is used to compile this model and the categorical\_crossentropy loss function is employed to maintain homogeneity for comparison.

## 2.8 MobileNet

The MobileNet model precisely, MobileNetV1 used in our implementation is pre-trained on the ImageNet dataset. This pre-training provides an advantage, as the model has already learned a wide range of features from a diverse set of images. Using include\_top=False to discard the top classification layer, allowing us to customize the network for our task. The input shape is adjusted to match our MRI images (100x100 pixels with 3 color channels). Following the MobileNet base, a GlobalAveragePooling2D layer was added. Thus, the spatial dimensionality of the feature maps is reduced by summarizing the most critical features in every map. Next, 1024 neurons in a fully connected Dense layer with ReLU activation function[3,6]. The final layer is also Dense with a softmax activation and two neurons, designed for our task. Adam

optimizer is used to compile the model and the categorical\_crossentropy loss function is used for comparison.

## 2.9 INCEPTIONV3

InceptionV3 has a 'network within a network' approach and efficiency in handling various scales of features. It was developed by Google and pre-trained on ImageNet, and this model is adapted for binary classification of brain tumor in an MRI image. It makes use of the Inception modules which allows incorporating multiple filters within the same layer. Factorisation is used to decompose the large convolutional layers to smaller manageable operations to achieve a more robust model. This also uses a batch normalization approach so that different images can be effectively classified. Modification of InceptionV3 was done to accept 100x100 pixel images with three color channels, replacing the top classification layers to suit our binary classification task of detecting brain tumors. The GlobalAveragePooling2D a layer was also used, followed by the layer that is fully connected with a softmax activation function. Parameters were all kept the same for the sake of fair comparison.

## 3. RESULTS AND DISCUSSION

The performance metrics of the models are: precision, recall, f1-score, accuracy, specificity and the training validation and accuracy and loss graphs. Precision reflects the model's ability to correctly recognize tumors, minimizing false positives. Recall minimizes the risk of false negatives. F1-score is useful as it considers the harmonic mean of both recall and precision, so it is useful for imbalance datasets. The graphs are instrumental in understanding the learning dynamics of the model over the epochs.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$F1 \text{ score} = \frac{2 \times precision \times recall}{precision + recall} \quad (3)$$

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \quad (4)$$

$$Specificity = \frac{TN}{TN + FP} \quad (5)$$

Below is the table for the comparison for all the five models used and their respective performance measures.

**Table 1. Comparison of all the models used**

Model	Precision(%)	Recall(%)	F1-score(%)	Accuracy(%)	Specificity(%)
CNN	96.43	100.00	98.18	97.55	92.77
VGG-16	97.58	99.38	97.47	97.96	95.18
<b>VGG-19</b>	<b>97.56</b>	<b>98.77</b>	<b>98.16</b>	<b>97.55</b>	<b>95.18</b>

ResNet-50	88.76	97.53	92.94	90.20	75.90
MobileNetV1	95.24	95.24	95.28	96.73	97.53
InceptionV3	96.99	99.38	98.18	97.55	93.98

Each model exhibited distinct strengths: CNN achieved the highest recall (100.00%), excellent for identifying all tumor cases, with an overall accuracy of 97.55%. VGG-16 and VGG-19 demonstrated superior precision (above 97.5%) and high accuracy (around 97.55%), making them robust at minimizing false positives. ResNet-50 showed a unique profile with a high recall (97.53%) but lower precision (88.76%) and accuracy (90.20%), indicating a tendency towards false positives. MobileNetV1 stood out with the highest specificity (97.53%), excellent for correctly identifying non-tumor cases, and a balanced accuracy of 96.73%. InceptionV3 matched the high accuracy of CNN and VGG-19 (97.55%) and showed strong precision and recall. To further aid the contrast between each model, below are the training accuracy and validation accuracies.

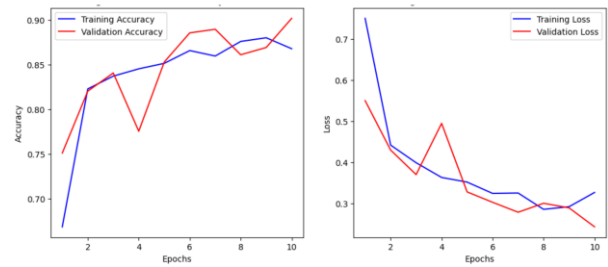


Fig 6: ResNet50 Training and validation accuracy and loss graphs

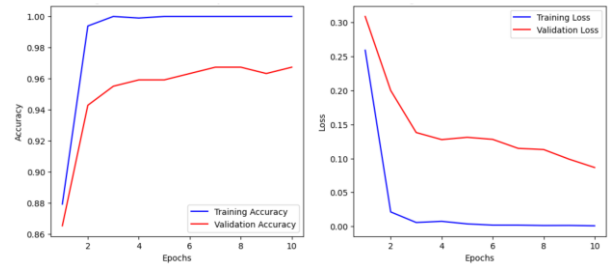


Fig 7: MobileNet Training and validation accuracy and loss graphs

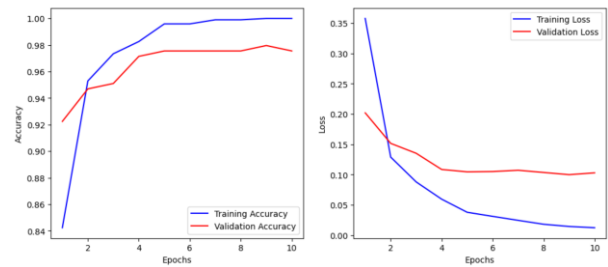


Fig 8: InceptionV3 Training and validation accuracy and loss graphs

While all models showed commendable performance in brain tumor classification, VGG-16 and VGG-19 slightly edged out others in terms of overall balance across all metrics. However, the choice of the model could depend on the actual clinical requirements, such as prioritizing the reduction of false negatives (high recall) or false positives (high specificity).

#### 4. CONCLUSION

In conclusion, this study has evaluated several deep learning models for brain tumor classification, namely CNN, VGG-16, VGG-19, ResNet-50, MobileNetV1, and InceptionV3. The VGG-16 model was the most effective, achieving an accuracy of 97.96% with a high precision value of 97.58% and specificity of 95.18%. The model's balanced performance across precision, recall, and F1-score values indicated its competence in medical image analysis tasks. Additionally, CNN, VGG-19, and InceptionV3 models were highly performing with accuracy values above 97.55%. However, the results also show that ResNet-50 and MobileNetV1 models, although slightly less efficacious, can be improved drastically with the hypertuning of parameters. For the future scope of the paper, we plan on creating an end to end system for hospitals, where a patient can store, and track all their reports and diagnosis. Also, we plan to specifically target the detection of

different tumor types i.e meningioma, pituitary and glioblastomas.

## 5. REFERENCES

- [1] Al-Azzwi ZHN, Nazarov AN. Brain Tumor Classification based on Improved Stacked Ensemble Deep Learning Methods. *Asian Pac J Cancer Prev.* 2023 Jun 1;24(6):2141-2148. doi: 10.31557/APJCP.2023.24.6.2141. PMID: 37378946; PMCID: PMC10505861.
- [2] Saeedi S, Rezayi S, Keshavarz H, R Niakan Kalhori S. MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques. *BMC Med Inform Decis Mak.* 2023 Jan 23;23(1):16. doi: 10.1186/s12911-023-02114-6. PMID: 36691030; PMCID: PMC9872362.
- [3] Ullah N, Javed A, Alhazmi A, Hasnain SM, Tahir A, Ashraf R (2023) TumorDetNet: A unified deep learning model for brain tumor detection and classification. *PLoS ONE* 18(9): e0291200. <https://doi.org/10.1371/journal.pone.0291200>
- [4] Wang, Jianyi. (2023). The Role of Machine Learning in the Detection and Classification of Brain Tumors: A Literature Review of the Past Two Years. *Computer and Information Science.* 16. 20. 10.5539/cis.v16n2p20.
- [5] Amin, Javeria & Sharif, Muhammad & Haldorai, Anandakumar & Mussarat, Yasmin & Nayak, Ramesh. (2021). Brain tumor detection and classification using machine learning: a comprehensive survey. *Complex & Intelligent Systems.* 8. 10.1007/s40747-021-00563-y.
- [6] Pandiyan, Vinishkumar & Babu, Vindhiya. (2023). Exploring AI-Powered Brain Tumor Detection with Deep Learning. 10.13140/RG.2.2.11831.296057
- [7] Nazir, M., Shakil, S., & Khurshid, K. (2021). Role of deep learning in brain tumor detection and classification (2015 to 2020): A review. *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society,* 91, 101940. <https://doi.org/10.1016/j.compmedimag.2021.101940>
- [8] Zhao, L., Ma, J., Shao, Y., Jia, C., Zhao, J., & Yuan, H. (2022). MM-UNet: A multimodality brain tumor segmentation network in MRI images. *Frontiers in oncology,* 12, 950706. <https://doi.org/10.3389/fonc.2022.950706>
- [9] Hasan, Nafiul & Aktar, Mousumi & Rana, Md. Masud & Moni, Mohammad. (2023). Machine Learning-based Biomedical Antenna for Brain Tumor Detection. 10.1109/NCIM59001.2023.10212805.
- [10] Ahamed MF, Hossain MM, Nahiduzzaman M, Islam MR, Islam MR, Ahsan M, Haider J. A review on brain tumor segmentation based on deep learning methods with federated learning techniques. *Comput Med Imaging Graph.* 2023 Dec;110:102313. doi: 10.1016/j.compmedimag.2023.102313. Epub 2023 Nov 24. PMID: 38011781.
- [11] Nazir M, Shakil S, Khurshid K. Role of deep learning in brain tumor detection and classification (2015 to 2020): A review. *Comput Med Imaging Graph.* 2021 Jul;91:101940. doi: 10.1016/j.compmedimag.2021.101940. Epub 2021 May 15. PMID: 34293621.
- [12] Alsubai S, Khan HU, Alqahtani A, Sha M, Abbas S, Mohammad UG. Ensemble deep learning for brain tumor detection. *Front Comput Neurosci.* 2022 Sep 2;16:1005617. doi: 10.3389/fncom.2022.1005617. PMID: 36118133; PMCID: PMC9480978.
- [13] R. Tamilselvi, A. Nagaraj, M. P. Beham and M. B. Sandhiya, "BRAMSIT: A Database for Brain Tumor Diagnosis and Detection," 2020 Sixth International Conference on Bio Signals, Images, and Instrumentation (ICBSII), Chennai, India, 2020, pp. 1-5, doi: 10.1109/ICBSII49132.2020.9167530.
- [14] Haq, Engr & Asad, Maimoona & Husnain, Ghassan & Anwar, Shahzad. (2023). Enhancing Brain Tumor Detection: A Machine Vision- Based Multiclass Classification Approach. *Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/Journal of Xi'an Shiyou University.* 19. 123-132.
- [15] Hammad M, ElAffendi M, Ateya AA, Abd El-Latif AA. Efficient Brain Tumor Detection with Lightweight End-to-End Deep Learning Model. *Cancers (Basel).* 2023 May 19;15(10):2837. doi: 10.3390/cancers15102837. PMID: 37345173; PMCID: PMC10216217
- [16] Kothari, S., Chiwhane, S., Jain, S., & Baghel, M. (2022). Cancerous brain tumor detection using a hybrid deep learning framework. *Indonesian Journal of Electrical Engineering and Computer Science.*doi: <http://doi.org/10.11591/ijeecs.v26.i3.pp1651-1661>
- [17] Sowrirajan, S.R., & Balasubramanian, S. (2022). Brain Tumor Classification Using Machine Learning and Deep Learning Algorithms. *International Journal of Electrical and Electronics Research.*
- [18] Y. Megersa and G. Alemu, "Brain tumor detection and segmentation using hybrid intelligent algorithms," *AFRICON 2015, Addis Ababa, Ethiopia, 2015,* pp. 1-8, doi: 10.1109/AFRCON.2015.7331938.
- [19] Tasci, Erdal & Ugur, Aybars & Camphausen, Kevin & Zhuge, Ying & Zhao, Rachel & Krauze, Andra. (2022). 3D Multimodal Brain Tumor Segmentation and Grading Scheme based on Machine, Deep, and Transfer Learning Approaches. *International Journal of Bioinformatics and Intelligent Computing.* 1. 10.61797/ijbic.v1i2.153.
- [20] Haq EU, Jianjun H, Huarong X, Li K, Weng L. A Hybrid Approach Based on Deep CNN and Machine Learning Classifiers for the Tumor Segmentation and Classification in Brain MRI. *Comput Math Methods Med.* 2022 Aug 5;2022:6446680. doi: 10.1155/2022/6446680. Retraction in: *Comput Math Methods Med.* 2023 Dec 13;2023:9848046. PMID: 36035291; PMCID: PMC9400402.