Research Article

Transfer Learning with Feature Extraction Modules for Improved Classifier Performance on Medical Image Data

Ritesh Jha [ID]**, Vandana Bhattacharjee** [ID]**, and Abhijit Mustafi** [ID]

Birla Institute of Technology, Mesra, Ranchi, India

Correspondence should be addressed to Vandana Bhattacharjee; vbhattacharya@bitmesra.ac.in

Transfer learning attempts to use the knowledge learned from one task and apply it to improve the learning of a separate but similar task. This article proposes to evaluate this technique's effectiveness in classifying images from the medical domain. The article presents a model TrFEMNet (Transfer Learning with Feature Extraction Modules Network), for classifying medical images. Feature representations from General Feature Extraction Module (GFEM) and Specific Feature Extraction Module (SFEM) are input to a projection head and the classification module to learn the target data. The aim is to extract representations at different levels of hierarchy and use them for the final representation learning. To compare with TrFEMNet, we have trained three other models with transfer learning. Experiments on the COVID-19 dataset, brain MRI binary classification, and brain MRI multiclass data show that TrFEMNet performs comparably to the other models. Pretrained model ResNet50 trained on a large image dataset, the ImageNet, is used as the base model.

1. Introduction

Transfer learning is the paradigm of learning that aims to transfer knowledge from one task to another, which is somewhat related [1]. In deep neural networks, the first layer often learns the general features, and eventually, by the last layer, the specific features are learned [2]. For example, the different nodes will learn features specific to a particular class in the last layers of a neural network. In this research work, we have studied the transferability of learning, with the variation in the number of layers trained as feature extraction modules. This can be helpful because a lot of the low-level features that have been learned from a vast amount of readily available data can be used for another task that may have less amount of useable data. Often, when the dataset on which the model is to be built is less, we opt for transfer learning. A network is trained on data like our dataset, such as image data, and the knowledge is transferred to a different task, such as the diagnosis of X-ray images. For this, some layers of the old network are retrained on this new dataset. If we retrain all the parameters in our network, then this initial training phase on image recognition is sometimes called pretraining. Then, we update all the weights and train on the target data by fine-tuning. Figure 1 depicts the general proposed methodology used in this study.

In this study, we have taken the brain magnetic resonance images (MRIs) and COVID-19 X-rays to evaluate the effectiveness of transfer learning. Magnetic resonance imaging is the most used technique for identifying brain tumors. MRI may detect these tumors with the help of an expert or a doctor's opinion. In this research, we apply transfer learning with the feature extraction modules from which the representations are input to a projection head. A softmax classifier classifies the learned features. We name our model TrFEMNet, which is effective because features from different levels of the hierarchy contribute to the final classification process. We believe that combining features from different levels of the hierarchy by a projection head is a novel contribution and results in effective outcomes, as seen from the experiments. This model is evaluated for the identification of tumors from brain MRI images or COVID-19 and viral pneumonia cases from COVID-19 X-ray images. The ResNet 50 model has been taken and pretrained on the ImageNet dataset.
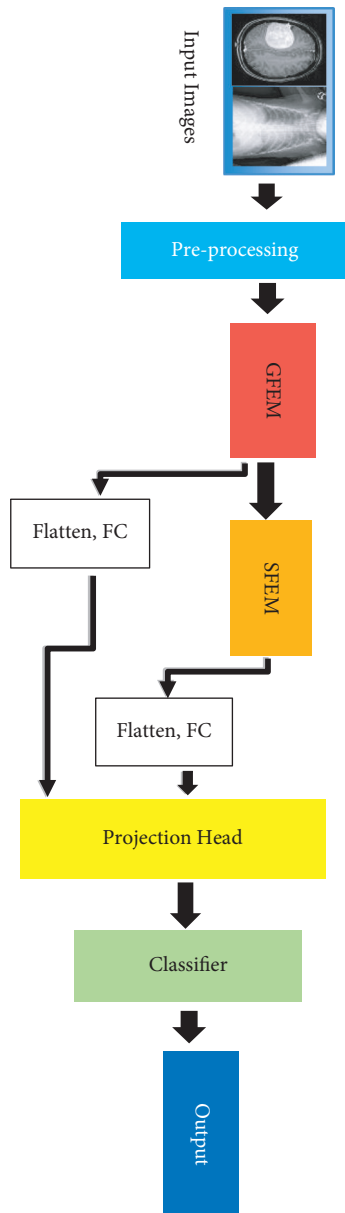
FIGURE 1: The workflow of the model is shown. In the transfer learning procedure, the input passes through GFEM (General Feature Extraction Module) and SFEM (Specific Feature Extraction Module), and the features extracted are passed through a projection head and then to the classifier to give the final output.

The highlights of this article are as follows:

(1) Preprocessing, data augmentation, and normalization of images.

(2) Presenting a mathematical formulation for transfer learning.

(3) Proposing the model TrFEMNet, with feature representation by the GFEM and SFEM, with a projection head comprising of a two-layer multilayer perceptron with nonlinearity, on the pretrained ResNet50 model

(4) Evaluating the results of experiments with Brain MRI 2-class, brain MRI multiclass, and COVID-19 X-ray dataset.

Sensitivity, specificity, F1-score, and accuracy are used to evaluate the classifiers' performance. The rest of the article is organized as follows. Section 2 presents the literature review, Section 3 presents the mathematical formulation of transfer learning, Section 4 presents the materials and methods, and Section 5 presents the results. Finally, Sections 6 and 7 present the discussion and conclusion.

2. Literature Review

Numerous studies on the brain MRI classification have been carried out. Varuna Shree et al. [3] presented a model that uses a discrete wavelet transform (DWT) for feature extraction, statistical features to reduce the number of features, and a blended artificial neural network for brain MRI classification. In this research, they obtained 98% accuracy. The authors in Ref. [4] present a content-based brain tumor detection system. They design a feature extraction framework using the VGG19 convolutional neural network (CNN) model with closed-form metric learning. The authors in Ref. [5] propose a method of predicting CT images from MRI images, showing that their method is very robust. A model for tumor classification and segmentation was presented by Ali and Davut Hanbay [6], yielding a classification accuracy of 97.18%. Sajid et al. [7] introduced the data augmentation methodology to the original dataset and then processed it with the convolutional neural network (CNN) method; the softmax function was employed in the classifier on both original and enhanced data, upon which the accuracy was 94.58%. Sachdeva et al. [8] applied the region of interest (ROI) approach to an MRI dataset. The ROI images were then segmented to eliminate tissue and color features. Then, they chose the most efficient characteristics using the genetic algorithms (GA) and went through the classification procedure. They obtained an accuracy of 94.9% in their study.

Nazir, Wahid, and Ali Khan described a new approach for automated brain tumor MRI identification and classification [9]. The data were classified into two parts: benign and malignant. In the presented paradigm, there are three basic phases. Filter methods were used to eliminate image noise in the first step. In the second stage, the features were used to derive the mean color moment of each image in the dataset. In the last stage, an artificial neural network (ANN) categorized the feature set of color moments with 91.8 percent classification accuracy. In Ref. [10], the author applied various deep learning techniques to classify Benign and malignant tumors. They achieved 98.49% accuracy using a hybrid approach of CNN and SVM.

Navid et al. [11] proposed a deep learning model with generative adversarial networks in the multiclassification of MRI images. They used the GAN on several MRI datasets, including meningioma, glioma, and pituitary tumors and then used a six-layer deep learning model to achieve an

overall accuracy of 95.60 percent. Chakraborty [12] has created a dataset on Kaggle, upon which several researchers have applied machine learning techniques and built models. Habibzadeh et al. [13] apply pretrained deep learning models for automatic white blood cell classification, and these models perform very well. The authors in Refs. [14, 15] present a review of image enhancement techniques and work with fewer data. Fayaz et al. [16] incorporated feature extraction methods and converted images to three-channel mode; the final accuracy achieved was 92.5% with the KNN classifier. Tajik et al. [17] proposed a texture overcome matrix model. They used the different feature extraction methods such as PCA, index approach, and Gabor filtering to obtain an accuracy of 96.67%. Togacar et al. [18] employed the CNN and SVM to achieve 96.77% accuracy with the masks produced by the hypercolumn approach in the suggested method.

In Ref. [19], the author used MobileNetv2 to select the features. They generated the 1000 features using Mobile-Netv2 and used iterative neighborhood component analysis to find the most important features. These features are trained using SVM and obtained an accuracy of 99.10%. The dataset used 444 images related to three types of tumor diseases and the rest with no tumor. The authors of Ref. [20] obtained a 95.75 percent accuracy using the CNN model with 22 layers with the transfer learning approach. They used the 22 layers in their proposed model, and the dataset used in this study comprised three main kinds of tumors. The authors in [21] developed the model for segmentation and detecting brain tumors. They used Berkeley's wavelet transformation and the deep learning model for image segmentation in this research. They also used the GLCM method for extracting the features from segmented images. In this research, the author obtained an accuracy of 98.5%. The author of Ref. [22] tested many approaches for classification, including support vector machines (SVMs), K-nearest neighbors (KNNs), binary decision trees (BDTs), random forest (RF), and ensemble methods, and obtained a high accuracy of 97% utilizing the SVM classifier. The authors of Ref. [23] employed the ResNet architecture with data suspension encoding and fusion layer to achieve a 98.02 percent accuracy. The dataset used in this study included three types of tumor images. The author of Ref. [24] compared the findings using three distinct datasets. The author proposed a scalable range-based adaptive bilateral filter to reduce noise from pictures in this article. The fully convolutional network was then used for segmentation. They were able to achieve a 97% accuracy rate. Several other researchers have also applied various transformation techniques to improve the quality of images used for classification [25, 26]. The authors in Ref. [27] propose a convolutional neural network model with dilations to classify different types of brain tumors and achieve an accuracy of 97%. The authors in Ref. [28] have applied several optimizers on CNN models for brain tumor segmentation and found that the Adam optimizer had the best accuracy of 99.2% in enhancing the CNN ability in classification and segmentation. In contrast, others in Ref. [29] aim at improving feature extraction by using a texture amortization

map (TAM). Several researchers have also extensively studied the application of deep neural networks such as hybrid networks [30], capsule networks [31], and convolutional networks [32–35].

3. Transfer Learning—Overview and Notations

Transfer learning is a very powerful idea in deep learning since we can take the knowledge a neural network has learned from one task and apply that knowledge to a separate task [36–39]. The transformation of each original feature into a learned representation for knowledge transfer preserves the properties or the potential structures of the data and finds the correspondence between features [40]. We first present the notations used for a better explanation of concepts. In this work, we assume that the labels for all data instances are available. Hence, supervised classification is performed. However, our notations are broadly adapted from Ref. [40].

Definition 1. A Domain $\mathscr{D}$ comprises a feature space $\mathscr{FS}$ and a label space $\mathscr{FS}$, so $\mathscr{D} = \{\mathscr{FS}, \mathscr{FS}\}$. The set of data instances comprises elements belonging to the feature space $\mathscr{FS}$.

Definition 2. A Task $\mathscr{T}$ is composed of a label space $\mathscr{Y}$ and a function f to be learned such that $\mathscr{T} = \{\mathscr{Y}, f\}$. The function f is often also called a model or predictor since it predicts the values of y for unknown x.

Definition 3. A learned model (function) lf is defined as $lf = \mathscr{L}(\mathscr{D}, \mathscr{T})$. It may be noted that lf is the model we are interested in improving upon since f is not known to us.

Consider a scenario where a network learns to recognize objects like cars and trucks, and then uses that knowledge or uses part of that knowledge to improve the task of reading X-ray scans. Here, we have a source domain $\mathscr{D}_s$ of images and task $\mathscr{T}_s$ of identifying the labels of cars and trucks in the source domain. Thus, the source domain $\mathscr{D}_s$ can be presented as follows: $\mathscr{D}_s = \{(x, y)|x_i \in \mathscr{X}^s, y_i \in \mathscr{Y}^s, \quad i = 1, 2, \ldots, n^s\}$, and let the learned model be lf^s.

Definition 4. Given $\mathscr{D}_s, \mathscr{T}_s, \mathscr{D}_t, \mathscr{T}_t$ Transfer Learning utilizes the knowledge from the source domain to improve the performance of lf^t on the target domain $\mathscr{D}_t$.

Definition 5. We thus state that $lf^{t*} = \{lf^s_{\text{trans}}\}(\mathscr{D}_t, \mathscr{T}_t)$, where lf^{t*} is the model learned by applying for knowledge transfer from lf^s, on $\mathscr{D}_t, \mathscr{T}_t$; while $lf^t = \mathscr{L}(\mathscr{D}_t, \mathscr{T}_t)$.

In other words, by transfer learning, one can use the knowledge from the task $\mathscr{T}_s$ of identifying the labels cars and trucks in the source domain $\mathscr{D}_s$ of images to improve upon the target task $\mathscr{T}_t$ of reading X-ray images for the presence of a disease, from the target domain of X-ray images.

The scenario addressed in this research work is homogeneous, inductive transfer learning, and we apply the feature-based and parameter-based approach [41]. We aim at investigating whether lf^{t*} is an improvement on lf^t (please refer to Definition 5).

4. Materials and Methods

Three datasets have been used in this study. Two datasets are for brain MRI classification, namely, 2-class and 4-class data, and one 3-class dataset for COVID X-ray images.

4.1. Dataset Description

4.1.1. COVID-19 X-Ray 3-Class Dataset.
This dataset is available on Kaggle [42]. It contains 111 samples of COVID-19 class and 70 samples of a normal class, and 70 samples of viral pneumonia class with 512×512 size. We generated 500 samples of each class using the data augmentation technique; median filters and histogram equalization were applied and resized the image into 224×224 size. Hence, there were 1500 samples (with an equal number of classes) for training. There are 26 COVID-19 samples, 20 normal samples, and 20 viral pneumonia samples for testing in the original dataset. In this dataset, the skull boundary technique was not applied.

4.1.2. Brain MRI Tumor 2-Class Dataset.
The dataset used in this study is available on the Kaggle repository [12]. The dataset includes normal and tumor samples, with 98 samples belonging to the tumor and 155 samples to the normal. A total of 253 images were associated with the patients' MRI brain scans. The image quality is not great because of the multiple resolutions, and each image has a different resolution. The format of images is JPEG. In this article, the images were converted to grayscale, and the skull's border was found by removing the image's background color. This is to remove any extra color that may be present outside of the skull. As a result, it provided the original image's contour. The data were split into 195 samples for training and 60 samples (30 of each class) for testing. Data augmentation techniques [18–20] were used to enhance the training samples to 6412. Median filters and histogram equalization were applied. All images of the original dataset have different resolutions with sizes 512×512. We resized the size of all images to 224×224 for further processing. All the images are normalized between 0 and 1.

4.1.3. Brain MRI Tumor 4-Class Dataset.
This dataset is available on the Kaggle website [43]. This dataset contains 300 samples of glioma tumor, 306 samples of meningioma tumor, 405 samples of no tumor, and 300 samples of pituitary tumor class. Apart from this, the dataset also contains 100 samples of glioma tumors, 115 samples of meningioma tumors, 105 samples of no tumors, and 74 samples of pituitary tumors for testing purposes. No data augmentation technique was applied. Median filters and histogram equalization were applied. The original size of the dataset was 512×512, which we resized into 224×224.

Table 1 gives the train test distributions and the class imbalance, if any, in all the datasets.

4.2. Preprocessing.
The preprocessing steps applied are enumerated as follows:

(1). Capturing skull boundary of brain MRI images

(2). Data augmentation: new data samples have been generated using the data augmentation process to balance the classes using the Keras module with the following augmentations: shear range of 20%, zoom range of 20%, rotation range angle of 30%, and fill mode set to nearest.

(3). Filters and histogram equalization: the median filter is the most effective and extensively used filter for eliminating noise from pictures. A median filter is applied to the images after data augmentation. Finally, histogram equalization was applied to increase the images' contrast and improve the image quality.

(4). Resizing: all images of the original dataset have different resolutions with sizes 512×512. We resized the size of all images to 224×224 for further processing. All the images are normalized between 0 and 1.

Figure 2 presents brain MRI images and their preprocessed images.

4.3. Evaluation Parameters.
The evaluation parameters used in this research work are sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), $F1$-score, and accuracy, as presented in Table 2.

4.4. Method Implementation Details.
All our experiments take the ResNet50 trained on the ImageNet dataset as the base model. The architecture of ResNet50 is presented in Figure 3. For our proposed architecture TrFEMNet, the General Feature Extraction module (GFEM) is implemented by freezing the weights of the lower layers of the ResNet50 model; the upper-level layers are retrained with the target data instances for implementing the Specific Feature Extraction module (SFEM). The output of the GFEM and SFEM gives the feature representations, which are fed into a projection head (PH) module. The output from the PH module is fed into the softmax classifier for the final output. The PH module comprises a two-layer MLP with ReLu nonlinearity. This module extracts the features at a middle level and higher level of hierarchy and gives a better representation for improved classifier performance. Few dense layers have also been added above the convolutional layers. For comparison with TrFEMNet, we have built three other models without the projection head. The variation in the number of layers with fixed/trainable weights has been done.

We develop models based on ResNet50, with softmax as an output layer activation function. The description of the models is as follows. Model 1 consists of one dense trainable layer of 1024 neurons (other than the last classification dense layer). In this model, we freeze all convolutional layers of ResNet50, thus enabling only one dense layer for training purposes. Model 2 is similar to Model 1, except it consists of two trainable dense layers of sizes 1024 and 512, respectively. In Model 3, we freeze all convolutional layers except the last one. In ResNet50, there are 48 convolutional layers, so we

TABLE 1: Dataset train test distributions.

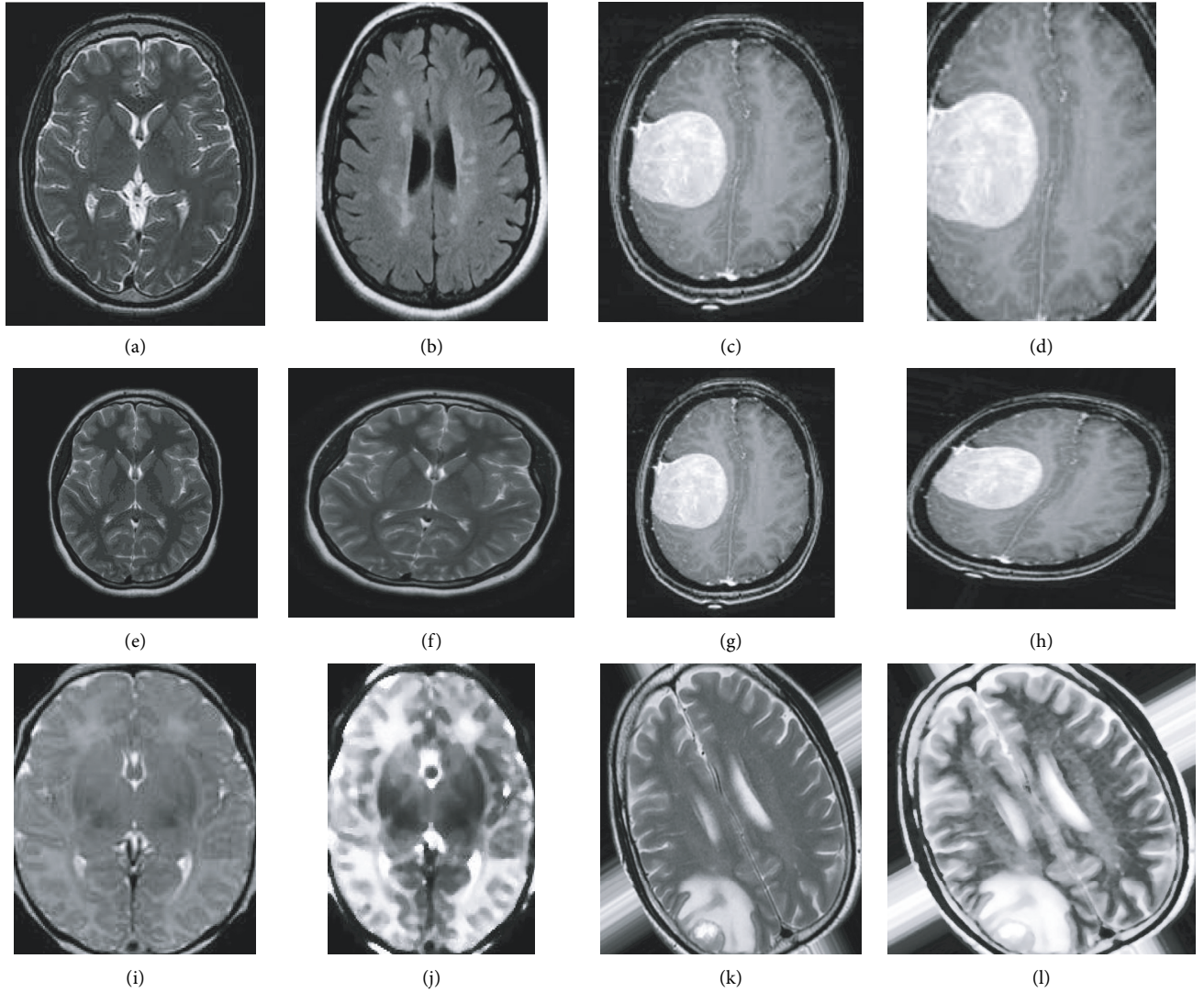| Dataset	Train samples	Test samples	Train set balanced	Test set balanced
COVID X-ray 3-class	1500	66	Yes	No
Brain MRI 2-class	6412	60	Yes	Yes
Brain MRI 4-class	1311	394	No	No

FIGURE 2: (a, b) A nontumor image and its boundary image. (c, d) A tumor image and its boundary image. (e, f) A nontumor image and its rotated image. (g, h) A tumor image and its rotated image. (i, j) A nontumor image and the resulting image after applying histogram equalization. (k, l) A tumor image and the resulting image after applying histogram equalization.

TABLE 2: Evaluation parameters for performance analysis.

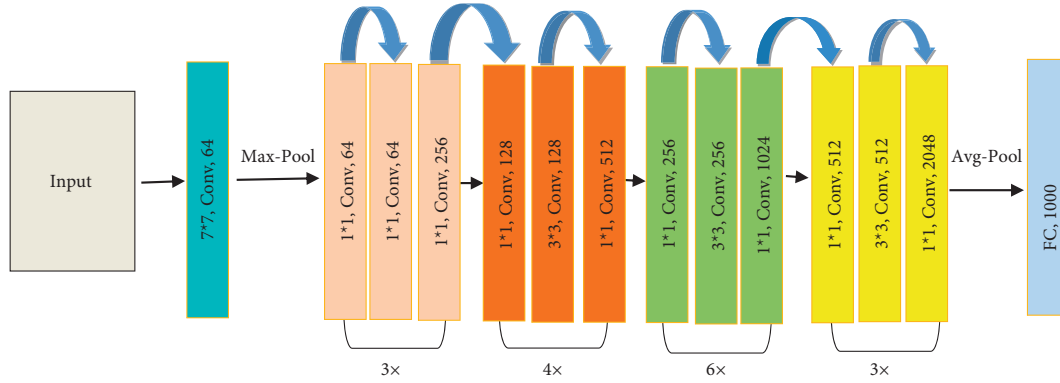| Parameter	Definition
Sensitivity (S) or Recall (R)	$\mathrm{TruePositive}\,(TP)/\mathrm{TruePositive}\,(TP) + \mathrm{FalseNegatives}\,(FN)$
Specificity (Sp)	$\mathrm{TrueNegative}\,(TN)/\mathrm{TrueNegative}\,(TN) + \mathrm{FalsePositives}\,(FP)$
$F1$-score	$2 * \mathrm{Recall}\,(R) * \mathrm{Precision}\,(P)/\mathrm{Recall}\,(R) + \mathrm{Precision}\,(P)$
Accuracy (A)	$\mathrm{TruePositive}\,(TP) + \mathrm{TrueNegatives}\,(TN)/\mathrm{TotalExamples}\,(TE)$

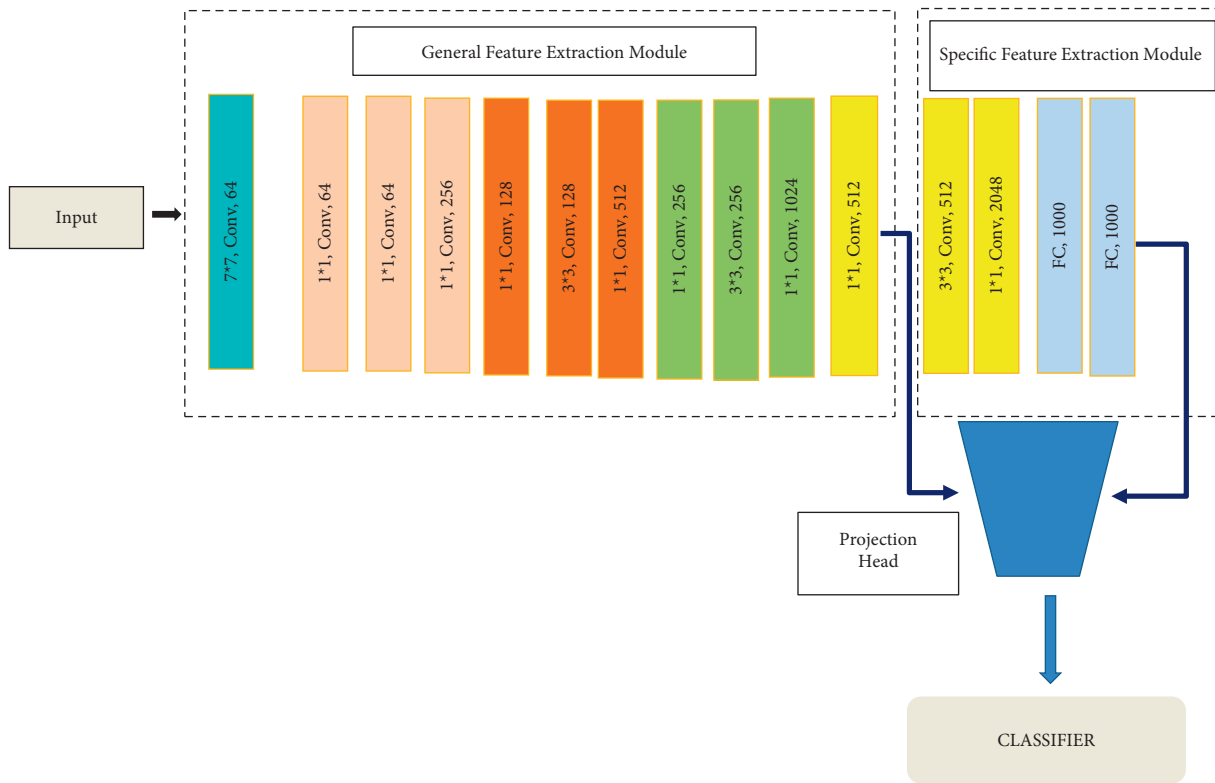FIGURE 3: The architecture of the ResNet50 model.

FIGURE 4: The architecture of the TrFEMNet model.

freeze the 47 convolutional layers and enable one convolutional and one dense layer of 1024 neurons for training purposes. For the TrFEMNet model, we keep two convolutional layers and two dense layers with 1024 and 512 neurons for the SFEM, and the frozen layers comprise the GFEM. The outputs from GFEM and SFEM are fed into the PH and then to the classifier. Figure 4 shows the schematic diagram of transfer learning for the TrFEMNet model.

Table 3 shows the parameters used in the training phase for all models.

5. Results

5.1. Experimental Results on All Datasets. The models were applied on healthcare datasets, and the results are presented in Tables 4–6.

TABLE 3: Training parameters.

Parameters	Value
Momentum	0.95
Dropout	0.3
Learning rate	0.0001
Loss	Binary cross-entropy, categorical cross-entropy
Optimizer	SGD
Epochs	30
Batch size	16

5.1.1. COVID-19 Dataset. 5.1.2. Brain MRI 2-Class Dataset. 5.1.3. Brain MRI 4-Class Dataset. 5.2. Analysis of TrFEMNet. The receiver operating curve (ROC) plots and confusion matrix of the performance of TrFEMNet for the different datasets are presented in Figures 5–7. We now

TABLE 4: Different performance measures of the deep learning models for the COVID-19 dataset.

Classifiers	Specificity (%)	Sensitivity (%)	Accuracy (%)	F1-score (%)
Model 1	95	97.82	97.13*	94.97
Model 2	95	97.65	96.20	93.65
Model 3	93.33	96.88	95.95	93.46
TrFEMNet	95*	97.84*	96.97	95.06*

TABLE 5: Different performance measures of the deep learning models for the brain MRI 2-class dataset.

Classifiers	Specificity (%)	Sensitivity (%)	Accuracy (%)	F1-score (%)
Model 1	89.47	89.47	89.47	89.46
Model 2	89.21	89.23	89.21	89.20
Model 3	100*	99.89*	99.89*	99.89*
TrFEMNet	99.35	99.37	99.39	99.37

TABLE 6: Different performance measures of the deep learning models for the brain MRI 4-class dataset.

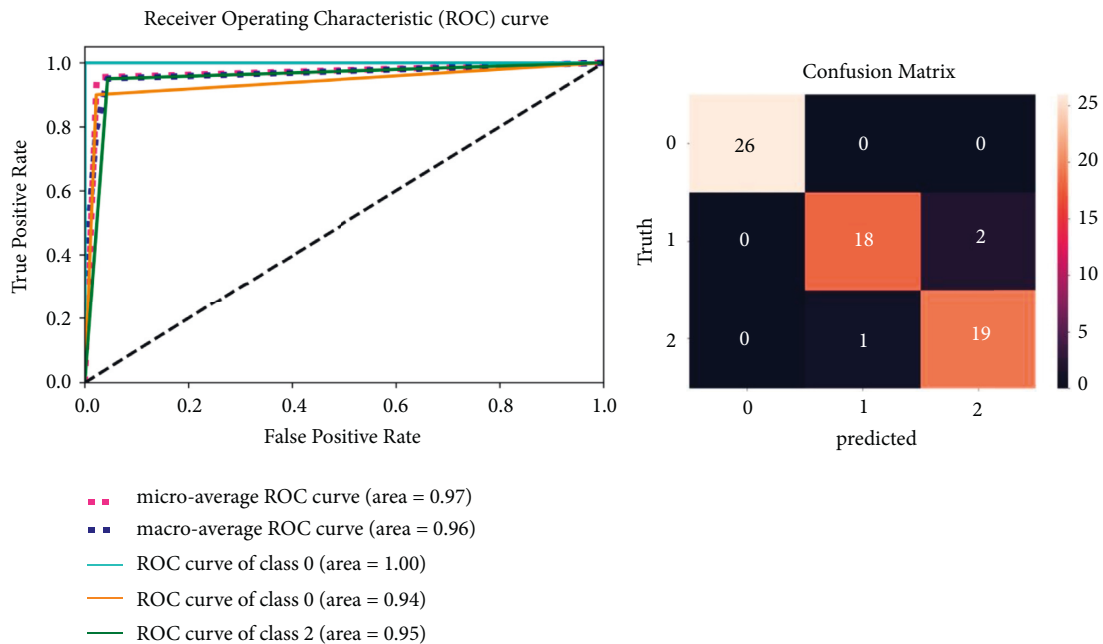| Classifiers	Specificity (%)	Sensitivity (%)	Accuracy (%)	F1-score (%)
Model 1	52.83	84.27	77.27	49.45
Model 2	51.50	83.37	75.49	50.58
Model 3	50.68	78.27	69.62	52.93
TrFEMNet	60.45*	85.01*	78.05*	70.65*

FIGURE 5: The ROC curve and confusion matrix of the TrFEMNet model for the COVID-19 dataset.

present the analysis of our proposed model with respect to the different datasets.

From the results presented in Tables 4–6, it is seen that TrFEMNet performs comparably to other models in all the cases. For the COVID-19 dataset, all the parameter values except accuracy are the highest. For the brain MRI 2-class dataset, the model gives the second highest values but within the 0.65% range. Moreover, for the brain MRI 4-class dataset, the value for specificity is 60.45%, which is 7% higher than the second best; the value for sensitivity is 85.01%, which is 0.74% higher than the second best. Similarly, values for accuracy at 78.05% and F1-score at 53.09% are also the highest.
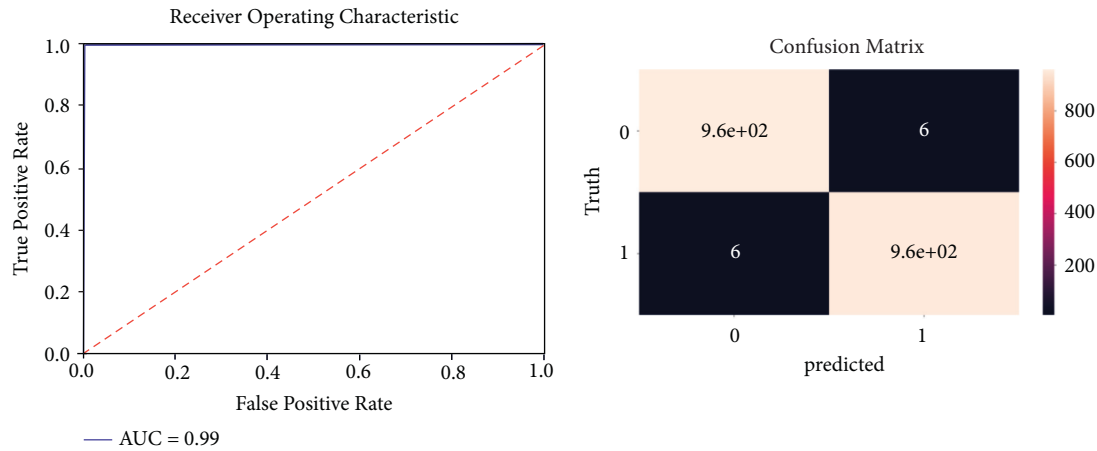
FIGURE 6: The ROC curve and confusion matrix of the TrFEMNet model for the Brain MRI 2-class dataset.
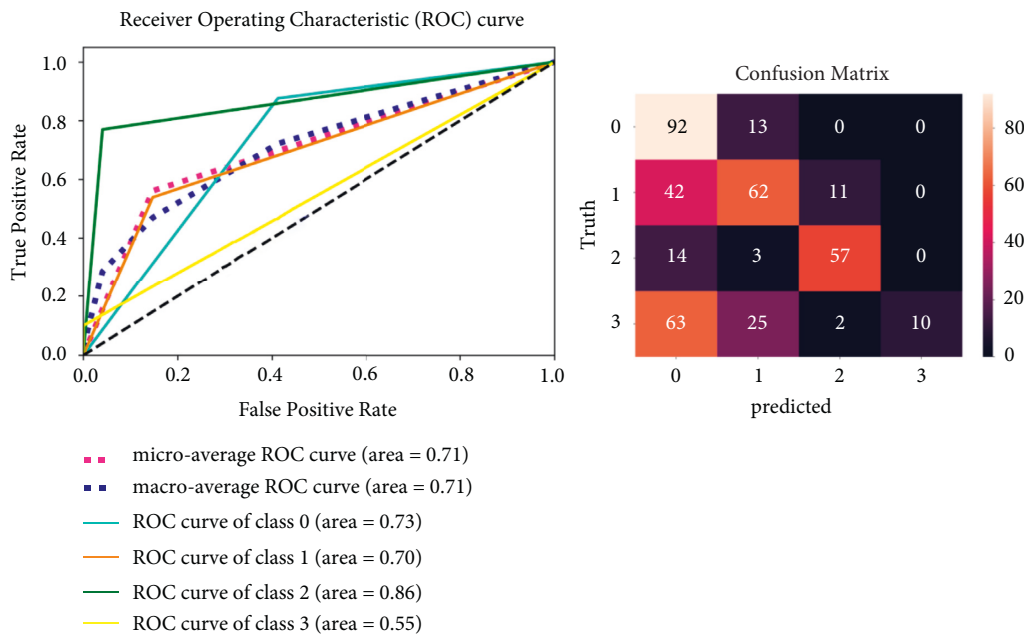

FIGURE 7: The ROC curve and confusion matrix of the TrFEMNet model for the Brain MRI 4-class dataset.

6. Discussion

The transfer learning approach applied in this research has used ResNet50 as the base model for parameter sharing from the previous training on the ImageNet dataset. The mathematical formulation of the transfer learning concept is given by means of several definitions, and Definition 5 provides the theme of the research, which has been conducted. We aim at finding out how well the model learned through transfer learning performs. For this, the concept of General Feature Extraction Module and Specific Feature Extraction Module has been introduced. The results of our experiments on the Brain MRI 2-Class dataset show that Model 1 and Model 2, which do not have any convolutional layers in their SFEM, do not perform very well. Model 3, with one convolutional layer in the SFEM, performs slightly better and achieves the best values for the brain MRI 2-class

dataset. The TrFEMNet model performs comparably to other models for most of the parameters on all datasets. For the COVID-19 dataset, the micro- and macroaverage ROC curve area is 0.97 and 0.96, respectively. For the Brain MRI 2-class dataset, the ROC curve area is 0.99, and for the Brain MRI 4-class dataset, both the micro- and macroaverage curve areas are 0.71, while for class glioma tumor, it is 0.86.

7. Conclusion

This article proposes to evaluate the transfer learning technique's effectiveness in classifying images from the medical domain. Pretrained model ResNet50 trained on a large image dataset, the ImageNet, is used as the base model. The article proposes a model TrFEMNet (Transfer Learning with Feature Extraction Modules Network), for classifying medical images. Feature representations from General

Feature Extraction Module (GFEM) and Specific Feature Extraction Module (SFEM) are input to a projection head and the classification module to learn the target data. The aim is to extract representations at different levels of hierarchy and use them for the final representation learning. In addition, for a detailed understanding of our work, we have detailed the steps for preprocessing, data augmentation, and normalization of images. A mathematical formulation for transfer learning is given in Section 3. The proposed model TrFEMNet obtains the feature representation by the GFEM and SFEM, with a projection head comprising a two-layer multilayer perceptron with nonlinearity, on the pretrained ResNet50 model. The model is evaluated by comparing it with three other models built by transfer learning, with brain MRI 2-class, brain MRI multiclass, and COVID-19 X-ray datasets. Experiments on the COVID-19 dataset, brain MRI binary classification, and brain MRI multiclass data show that TrFEMNet performs at par with other models, and for the most complex dataset, it achieves an improvement of 7% on specificity with respect to the second-best model. The outcome of this research motivates one to investigate further into the realm of transfer learning. As part of our ongoing work, we aim at investigating more architectures for the projection head and with more levels of extraction from the base model.

Data Availability

The dataset used in this study is available at https://www.kaggle.com.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] L. Torrey and J. Shavlik, *Chapter 11 transfer learning Handbook of Research on Machine Learning Applications*, E. Soria, J. Martin, R. Magdalena, M. Martinez, and A. Serrano, Eds., IGI Global, Pennsylvania, PA, U.S.A, 2009.

[2] J. Yosinski, J. Clue, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" *Advances in Neural Information Processing Systems*, vol. 27, 2014.

[3] N. Varuna Shree and T. N. R. Kumar, "Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network," *Brain Informatics*, vol. 5, no. 1, pp. 23–30, 2018.

[4] Z. N. K. Swati, Q. Zhao, M. Kabir et al., "Content-based brain tumor retrieval for MR images using transfer learning," *IEEE Access*, vol. 7, pp. 17809–17822, 2019.

[5] D. Nie, X. Cao, Y. Gao, L. Wang, and D. Shen, *Estimating CT images from MRI data using 3D fully convolutional networks Deep Learning and Data Labeling for Medical Applications*, pp. 170–178, Springer, Berlin, Germany, 2016.

[6] A. Ari and D. Hanbay, "Deep learning based brain tumor classification and detection System," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 26, no. 5, pp. 2275–2286, 2018.

[7] M. Sajjad, S. Khan, K. Muhammad, W. Wu, A. Ullah, and S. W. Baik, "Multigrade brain tumor classification using deep CNN with extensive data augmentation," *Journal of Computational Science*, vol. 30, pp. 174–182, 2019.

[8] J. Sachdeva, V. Kumar, I. Gupta, N. Khandelwal, and C. K. Ahuja, "A package-SFERCB-Segmentation, feature extraction, reduction and classification analysis by both SVM and ANN for brain tumors," *Applied Soft Computing*, vol. 47, pp. 151–167, 2016.

[9] M. Nazir, F. Wahid, and S. A. Khan, "A simple and intelligent approach for brain MRI classification," *Journal of Intelligent and Fuzzy Systems*, vol. 28, no. 3, pp. 1127–1135, 2015.

[10] M. O. Khairandish, M. Sharma, V. Jain, J. M. Chatterjee, and N. Z. Jhanjhi, "A Hybrid CNN-SVM Threshold Segmentation Approach for Tumor Detection and Classification of MRI Brain Images," *IRBM*, vol. 43, pp. 290–299, 2021.

[11] N. Ghassemi, A. Shoeibi, and M. Rouhani, "Deep neural network with generative adversarial networks-training for brain tumor classification based on MRI images," *Biomedical Signal Processing and Control*, vol. 57, Article ID 101678, 2019.

[12] N. Chakrabarty, "Brain MRI Images for Brain Tumor Detection Kaggle," 2019, https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-%20detection/metadata.

[13] M. Habibzadeh, M. Jannesari, Z. Rezaei, H. Baharvand, and M. Totonchi, "Automatic white blood cell classification using pre-trained deep learning models: Resnet and inception," in *Proceedings of the Tenth International Conference on Machine Vision*, vol. 10696, International Society for Optics and Photonics, Austria, April 2018.

[14] S. Rana, "A review of medical image enhancement techniques for image processing," *International Journal of Current Engineering and Technology*, vol. 5, no. 2, pp. 1282–1286, 2011.

[15] F. Chollet, "Building Powerful Image Classification Models Using Very Little Data," 2016, https://blog.keras.io/building-%20powerful-%20image-%20classification-%20models-%20using-%20very-%20little-%20data.html.

[16] M. Fayaz, A. S. Shah, F. Wahid, and A. Shah, "A robust technique of brain MRI classification using color features and K-Nearest neighbors algorithm," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 9, no. 10, pp. 11–20, 2016.

[17] M. N. Tajik, W. Khan, and B. Khan, "Texture feature selection using ga for classification of human brain MRI scans," in *Proceedings of the International Conference on Swarm Intelligence ICSI*, pp. 233–244, Berlin, Germany, June 2016.

[18] M. Togacar, Z. Comert, and B. Ergen, "Classification of brain MRI using hyper column technique with convolutional neural network and feature selection method," *Expert Systems with Applications*, vol. 149, Article ID 113274, 2020.

[19] A. K. Poyraz, S. Dogan, E. Akbal, and T. Tuncer, "Automated brain disease classification using exemplar deep features," *Biomedical Signal Processing and Control*, vol. 73, Article ID 103448, 2022.

[20] M. F. Alanazi, M. U. Ali, S. J. Hussain et al., "Brain tumor/mass classification framework using magnetic-resonance-imaging-based isolated and developed transfer deep-learning model," *Sensors*, vol. 21, 2022.

[21] M. Arif, F. Ajesh, S. Shamsudheen, O. Geman, D. Izdrui, and D. Vicoveanu, "Brain tumor detection and classification by MRI using biologically inspired orthogonal wavelet transform and deep learning techniques," *Journal of Healthcare Engineering*, vol. 2022, pp. 1–18, Article ID 2693621, 2022.

[22] B. Jena, G. N. Krishna, and S. Saxena, "An empirical study of different machine learning techniques for brain tumor classification and subsequent segmentation using hybrid texture feature," *Machine Vision and Applications*, vol. 33, 2022.

[23] G. Xiao, H. Wang, J. Shen, Z. Chen, Z. Zhang, and X. Ge, "Synergy factorized bilinear network with a dual suppression strategy for brain tumor classification in MRI," *Micromachines*, vol. 13, no. 1, 2022.

[24] V. V. S. Sasank and S. Venkateswarlu, "An automatic tumor growth prediction based segmentation using full resolution convolutional network," *Biomedical Signal Processing and Control*, vol. 71, Article ID 103090, 2022.

[25] Z. Ullah, M. Fayaz, and A. Iqbal, "Critical analysis of data mining techniques on medical data," *International Journal of Modern Education and Computer Science*, vol. 8, no. 2, pp. 42–48, 2016.

[26] F. Wahid, R. Ghazali, M. Fayaz, and A. S. Shah, "Using probabilistic classification technique and statistical features for brain magnetic resonance imaging (MRI) classification: an application of AI technique in bioscience," *International Journal of Bio-Science and Bio-Technology*, vol. 8, no. 6, pp. 93–106, 2016.

[27] S. S. Roy, N. Rodrigues, and Y. h. Taguchi, "Incremental dilations using CNN for brain tumor classification," *Applied Sciences*, vol. 10, no. 14, p. 4915, 2020.

[28] M. Yaqub, J. Feng, M. Zia et al., "State-of-the-Art CNN optimizer for brain tumor segmentation in magnetic resonance images," *Brain Sciences*, vol. 10, no. 7, p. 427, 2020.

[29] R. A. Gonzalez and A. M. Vazquez, "Deep learning architecture reduction for fMRI data," *Brain Sciences*, vol. 12, no. 2, p. 235, 2022.

[30] R. Nair, S. Vishwakarma, M. Soni, T. Patel, and S. Joshi, "Detection of COVID-19 cases through X-ray images using hybrid deep neural network," *World Journal of Engineering*, vol. 19, no. 1, pp. 33–39, 2021.

[31] M. Ragab, S. Alshehri, N. A. Alhakamy, R. F. Mansour, and D. Koundal, "Multiclass classification of chest X-ray images for the prediction of COVID-19 using capsule network," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–8, Article ID 6185013, 2022.

[32] L. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," *Scientific Reports*, vol. 10, no. 1, 2020.

[33] S. Shambhu, D. Koundal, P. Das, and C. Sharma, "Binary classification of covid-19 CT images using CNN: covid diagnosis using CT," *International Journal of E-Health and Medical Communications*, vol. 13, no. 2, pp. 1–13, 2021.

[34] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," *Physical and Engineering Sciences in Medicine*, vol. 43, no. 2, pp. 635–640, 2020.

[35] R. Nair, A. Alhudhaif, D. Koundal, R. I. Doewes, and P. Sharma, "Deep learning-based COVID-19 detection system using pulmonary CT scans," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 29, no. SI-1, pp. 2716–2727, 2021.

[36] Y. Bengio, *Deep learning of representations for unsupervised and transfer learning JMLR W&CP: Proc. Unsupervised and Transfer Learning*Vol. 27, PMLR, Bellevue, WA, USA, 2011.

[37] Y. Bengio, F. Bastien, A. Bergeron et al., "Deep learners benefit more from out-of-distribution examples," in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, Fort Lauderdale, FL, USA, April 2011.

[38] R. Caruana, *Learning many Related Tasks at the Same Time with Backpropagation*, pp. 657–664, MIT Press, Cambridge, MA, 1995.

[39] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei, "ImageNet: a large-scale hierarchical image database," in *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL, U.S.A, June 2009.

[40] F. Zhuang, Z. Qi, K. Duan et al., "A comprehensive survey on transfer learning," 2019, https://arxiv.org/abs/1911.02685.

[41] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.

[42] Kaggle, "Chest X-Ray Images (Pneumonia)," 2022, https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia.

[43] Kaggle, "Brain Tumor MRI Dataset," 2022, https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset.