IEEE Open Journal of the
Computer Society

Received 6 July 2024; revised 13 August 2024; accepted 24 August 2024. Date of publication 10 September 2024; date of current version 23 September 2024. The review of this article was arranged by Associate Editor Jingcai Guo.

Digital Object Identifier 10.1109/OJCS.2024.3453924

Utilizing Deep improved ResNet50 for Brain Tumor Classification Based MRI

KARRAR NEAMAH [1], FARHAN MOHAMED [2] (Senior Member, IEEE), SAFA RIYADH WAHEED [3], WALEED HADI MADHLOOM KURDI [4,5], ADIL YASEENTAHA [6], AND KARRAR ABDULAMEER KADHIM [1,7]

[1]Faculty of Engineering, School of Computing, University Technology of Malaysia Johor Bahru, Skudai, Johor 81310, Malaysia
[2]UTM-IRDA MaGICX, Institute of Human Centered Engineering, Universiti Teknologi Malaysia Johor Bahru, Skudai, Johor 81310, Malaysia
[3]Islamic University, Najaf 2975+9RM, Iraq
[4]Altoosi University College, Computer Science, Najaf 31001, Iraq
[5]Department of Electrical Engineering, Faculty of Engineering, University of Kufa, Kufa, Najaf Governorate 299G+HPX, Iraq
[6]Al-Safwa University College, Department of Medical Instrumentation Techniques, Engineering, Karbala 56001, Iraq
[7]Computer Techniques Engineering Department, Faculty of Information Technology, Imam Ja'afar Al-Sadiq University, Baghdad 10001, Iraq

CORRESPONDING AUTHOR: KARRAR NEAMAH (e-mail: karrarneamah@yahoo.com).

ABSTRACT A robust approach for brain tumor classification is being developed using deep convolutional neural networks (CNNs). This study leverages an open-source dataset derived from the MRI Brats2015 brain tumor dataset. Preprocessing included intensity normalization, contrast enhancement, and downsizing. Data augmentation techniques were also applied, encompassing rotations and flipping. The core of our proposed approach lies in the utilization of a modified ResNet-50 architecture for feature extraction. This model integrates transfer learning by replacing the final layer with a spatial pyramid pooling layer, enabling it to leverage pre-trained parameters from ImageNet. Transfer learning from ImageNet aids in countering overfitting. Our model's performance was evaluated with various hyperparameters, including existing methods in terms of accuracy, precision, recall, F1-score, sensitivity, and specificity. This study showcases the potential of deep learning, transfer learning, and spatial pyramid pooling in MRI-based brain tumor classification, providing an effective tool for medical image analysis. Our methodology employs a modified ResNet-50 architecture with transfer learning, integrating a spatial pyramid pooling layer for feature extraction. Systematic evaluation showcases the model's superiority over existing methods, demonstrating remarkable results in accuracy (0.9902), precision (0.9837), recall (0.9915), F1-score (0.9891), sensitivity, and specificity. The comparative analysis against prominent CNN architectures reaffirms its outstanding performance. Our model not only mitigates overfitting challenges but also offers a promising tool for medical image analysis, underlining the combined efficacy of spatial pyramid pooling and transfer learning. The study's optimization parameters, including 25 epochs, a learning rate of 1e-4, and a balanced batch size, contribute to its robustness and real-world applicability, furthering advancements in efficient brain tumor classification within MRI data.

INDEX TERMS Brain tumor detection, CNN, data augmentation, Resnet-50, transfer learning.

I. INTRODUCTION

The early detection of brain tumors stands as a pivotal cornerstone in the realm of cancer diagnosis, given its potential to substantially augment survival rates. As previously alluded to, two primary categories of algorithms are instrumental in the quest for brain tumor detection: the conventional machine learning (ML) [1] paradigm and the more sophisticated deep learning (DL) [2] approach. Traditional ML-based methodologies often rely on what we term "handcrafted features," denoting attributes meticulously extracted from training data as a prerequisite for the learning process. In this regard, the involvement of an extensively experienced expert may become imperative in the endeavor to cherry-pick the most pivotal features. Consequently, the efficacy and representational fidelity become seminal determinants in governing the precision with which ML-based algorithms unravel data [2].

On the contrasting front, DL-based algorithms have exhibited remarkable prowess across diverse domains, particularly within the sphere of medical imaging. Within this domain, one of the most venerable DL models is the (CNN) [3], renowned for its intrinsic ability to autonomously unearth salient features from the training data, facilitated by its weight-sharing structure. The manifest advantages associated with DL have spurred a surge in scholarly interest towards its application in brain tumor classification [4], [5]. However, it is incumbent to acknowledge that the training of a CNN model mandates a substantial volume of images, owing to the sheer multitude of parameters necessitating assimilation. Furthermore, traditional CNNs are inclined to the peril of overfitting, an inherent concern, especially in the context of smaller datasets.

Consequently, in recent times, there has been a proliferation of advanced CNN variations characterized by deeper architectures, surpassing their conventional counterparts in terms of performance. Notably, the ResNet-50 model has ascended to eminence as a benchmark within the domain of image classification, owing to its prodigious depth and resultant capacity for highly refined representations.

In the landscape of medical imaging, the accurate detection of brain tumors through Magnetic Resonance Imaging (MRI) stands as a critical endeavor. Our research endeavors to introduce a novel strategy, centering around the implementation of the ResNet-50 architecture, with the explicit goal of augmenting classification accuracy in the realm of MRI-based brain tumor detection. Recognizing the inherent challenges posed by limited training data, we strategically incorporate a spatial pyramid pooling (SPP) layer into our model. This bespoke SPP layer is meticulously designed not only to enhance the model's resilience but also to leverage the potency of transfer learning. Our research unfolds several noteworthy highlights that delineate the distinctive contributions of our proposed model to the domain of MRI-based brain tumor detection. Firstly, the utilization of the ResNet-50 architecture serves as a foundational strength, capitalizing on its depth and skip connections to extract intricate features vital for accurate classification. Secondly, the introduction of a spatial pyramid pooling (SPP) layer exemplifies an innovative design choice, reinforcing the model's adaptability to varied spatial scales and augmenting its capacity to discern intricate patterns within the MRI data. Thirdly, the incorporation of transfer learning strategically addresses the perennial challenge of limited training data, facilitating the leveraging of pre-existing knowledge to enhance the model's generalization capability. Lastly, the amalgamation of these components forms a synergistic approach that not only elevates classification accuracy but also positions our model as a promising advancement in the broader landscape of deep learning applications for medical image analysis. Through this research, we seek to contribute a methodological advancement that holds promise for improving the efficacy of MRI-based brain tumor detection, offering a valuable stride forward in the intersection of deep learning and medical imaging.

The subsequent sections of our article are structured as follows: initially, we furnish an overarching exposition of the proposed model. Subsequently, we delve into a comprehensive elucidation of our experimental study. Following this, we present the findings and engage in a rigorous discussion. Finally, we draw our article to a close with a concise summary of our conclusions.

II. RELATED WORK

In the realm of brain tumor classification from MRI data, numerous studies have contributed to the growing body of knowledge. These investigations have explored a wide array of methodologies, including conventional machine learning techniques and deep learning approaches, shedding light on the challenges and advancements in this field [6]. Deep learning techniques, particularly (CNNs), have seen a surge in popularity for medical image analysis, including brain tumor classification. CNNs are adept at automatically learning relevant features from raw images, potentially circumventing the need for handcrafted feature engineering. A plethora of CNN architectures, such as VGGNet, AlexNet, and ResNet, have been harnessed for this purpose. In a study conducted by [7], A pioneering framework for the classification of brain tumors and the detection of specific tumor types is introduced, leveraging a novel architecture based on the Region-based (RCNN) technique. The primary objective of this endeavor is to substantially reduce the computational overhead associated with traditional RCNN architectures by employing a streamlined and less intricate framework. Furthermore, deep transfer learning, as exemplified by [8], The work introduces a transfer learning paradigm hinging on a pretrained VGG19 model as its cornerstone. This model has undergone intricate adaptations through the implementation of a customized (CNN) architecture, complemented by the judicious application of preprocessing techniques encompassing data normalization and augmentation. For another study [9], an extensive array of conventional and hybrid machine learning models was meticulously constructed and subjected to comprehensive scrutiny for the purpose of autonomously classifying brain tumor images, a rigorous examination encompassing 16 distinct transfer learning models was undertaken with the aim of identifying the optimal transfer learning paradigm for brain tumor classification, grounded in neural network methodologies. Additionally, research by [10], Exhaustive comparative performance assessment of transfer learning-based (CNNs), namely the pretrained VGG-16, ResNet-50, and Inception-v3 models, in the context of automating the prediction of tumor cells within the brain. The pretrained models are rigorously evaluated using a dataset comprising 233 MRI brain tumor images. The primary objective is to exploit the potential of the VGG-16 pretrained CNN model for the precise localization of brain tumors. Collectively, these related works underscore the versatility and potential of deep learning, CNN architectures, and transfer learning in MRI-based brain tumor classification, offering valuable insights into the evolving landscape of medical image analysis. Our proposed model builds upon

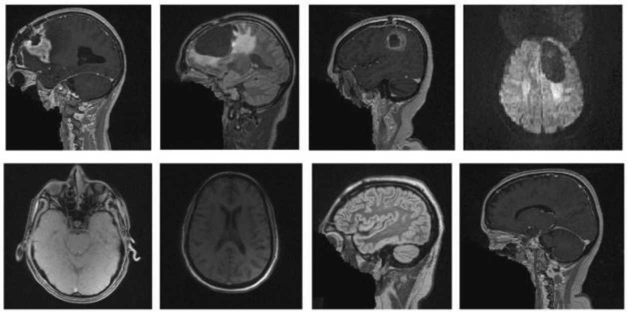

FIGURE 1. We present a selection of illustrative MRI images that were employed in the classification of brain tumors. The top row showcases images exhibiting abnormalities indicative of the presence of brain tumors, while the bottom row features MRI scans denoting typical, non-pathological brain configurations.

and extends the capabilities of these approaches, incorporating spatial pyramid pooling and transfer learning to further enhance classification accuracy and robustness.

III. DATASET

This research harnessed a publicly available dataset accessible on Kaggle, which serves as a derivative of the extensive Brats2015 brain tumor dataset. The database, aptly designated the "challenge database," constitutes an assembly of entirely anonymized images culled from The Cancer Imaging Archive.(https://www.kaggle.com/datasets/jakeshbohaju/brain-tumor). This dataset encapsulated a grand total of 3762 magnetic resonance (MR) images. From this extensive pool, 3060 were judiciously selected as a representative subset, with 1500 earmarked as instances denoting the presence of tumors (assigned the binary label "1"), while the remaining 1500 scans were designated as instances of non-tumor cases (labeled as "0"). To obviate the prevalence of class imbalance, a cardinal concern in machine learning, the dataset was meticulously divided with equipoise. Furthermore, a critical component of our methodology entailed the vetting of the suggested model's efficacy with a rigorous evaluation that involved 60 test images. Notably, this process entailed the judicious culling of any images susceptible to confounding the model during training. It is salient to note that the images within this dataset are devoid of any fixed dimensions. To render them compatible for subsequent analysis, all image samples underwent a harmonization process, effectuated through the utilization of a Keras automated resizing method. This method, adeptly and seamlessly, transformed all input images into a standardized $224*224$ dimensional format as shown in Fig. 1.

IV. PREPROCESSING AND DATA AUGMENTATION

Preprocessing constitutes an indispensable initial phase in the data refinement process, orchestrating the input data for subsequent analysis [11]. In the realm of medical imaging analysis, the principal objectives encompass rectifying MRI artifacts and elevating image contrast. One of the most formidable

FIGURE 2. Preprocessing of the input data.

challenges inherent to MRI analysis lies in the management of thermal noise, magnetic field aberrations, and the subtle patient motions incurred during the scanning procedure. The variations in MRI scan intensities, stemming from diverse data sources, further compound this intricacy [11]. Consequently, to mitigate the disparities in intensity, an intensity normalization [12] method based on linear transformation within the range of [0, 1] is deftly applied.

Our preprocessing protocol, depicted in Fig. 2, comprises three pivotal phases. Initially, the entire MRI image set undergoes intensity normalization, followed by the application of an MRI contrast enhancement technique, previously developed and refined. To optimize memory utilization, the input MR images are judiciously downscaled. In the realm of computer vision.

Data augmentation emerges as a pivotal facet, renowned for its effectiveness in the training of deep learning models. The extant literature has proffered various data augmentation techniques tailored to the exigencies of deep learning in medical imaging, encompassing random cropping, rotations, shears, and flips. Recent research findings underscore the nuanced efficacy of distinct augmentation techniques in capturing the intricate facets of medical images. In this research endeavor, several data augmentation techniques, specifically rotations and flips, were harnessed to enrich the training dataset and confer upon (CNNs) an expansive input space. Rotation, a fundamental augmentation strategy, entails the rotation of input images by varying degrees, such as 90, 180, and 270 degrees. Another commonplace technique involves flipping, which mirrors images both vertically and horizontally. The resultant augmented images are conspicuously portrayed in Fig. 3.

To address the concern regarding the explanation of how our proposed innovative system effectively solves the domain shift problem between the natural image domain and the medical image domain, we provide a detailed elaboration as follows: Domain shift refers to the discrepancy between the data distribution in the source domain (natural images) and the target domain (medical images), which poses a significant challenge in transferring learned features. Our approach leverages transfer learning using a pretrained ResNet-50 model, which captures a broad range of features from the ImageNet dataset. These features, while initially learned from natural images, include fundamental visual patterns that are also present in medical images. By fine-tuning this pretrained model on our medical image dataset, we adapt these general features to the specific characteristics of medical images, thereby mitigating the domain shift. Additionally, the incorporation of a Spatial Pyramid Pooling (SPP) layer enables

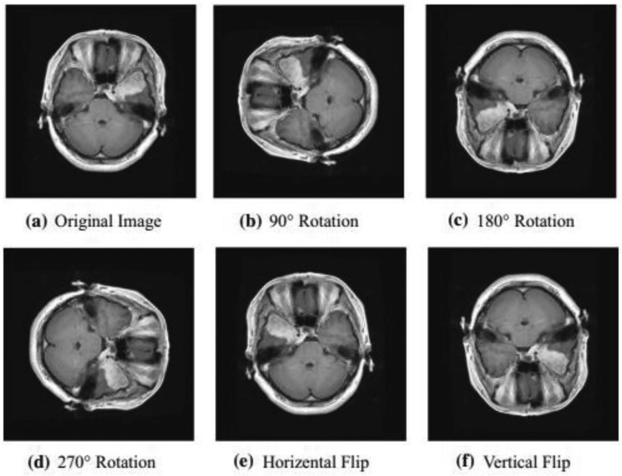

FIGURE 3. Example of data augmentation on MRI.

our system to handle images of varying sizes and scales, enhancing its ability to generalize across different domains. Comparative analysis with other methods demonstrates the superior performance of our system in medical image classification, providing quantitative evidence of its effectiveness. Through these mechanisms, our innovative system successfully addresses the domain shift challenge, ensuring robust and accurate classification of brain tumors from MRI images.

V. PROPOSED CNN-BASED FEATURE EXTRACTION

This process involves generating a large volume of image samples to augment the training dataset after undergoing data augmentation. The next phase involves extracting visual and discriminatory features to encapsulate their inherent characteristics. Given the remarkable success of deep learning within the realm of computer vision, (CNN) models have been harnessed to perform this feature extraction task. Recent times have witnessed the introduction of several highly efficient and cutting-edge versions of CNNs, such as VGGNet [16], AlexNet [17], and GoogLeNet [18], among others. These networks have consistently demonstrated exceptional performance in the domain of large-scale image classification.Particularly noteworthy is the ResNet-50 model, which has achieved considerable success in a variety of image classification tasks, courtesy of its profoundly deep representations, as exemplified in its deployment for tasks like ImageNet classification and object identification [19]. Hence, drawing inspiration from the impressive accomplishments of ResNet-50, this study proposes an enhanced ResNet-50 model rooted in CNN architecture for the task of image categorization.

Distinguishing itself from the original ResNet-50, the proposed model adopts a transfer learning approach. This approach involves the substitution of the final layer with a spatial pooling layer, serving as a potent tool for feature extraction. The structural refinement of this enhanced network can be visualized in Fig. 4.

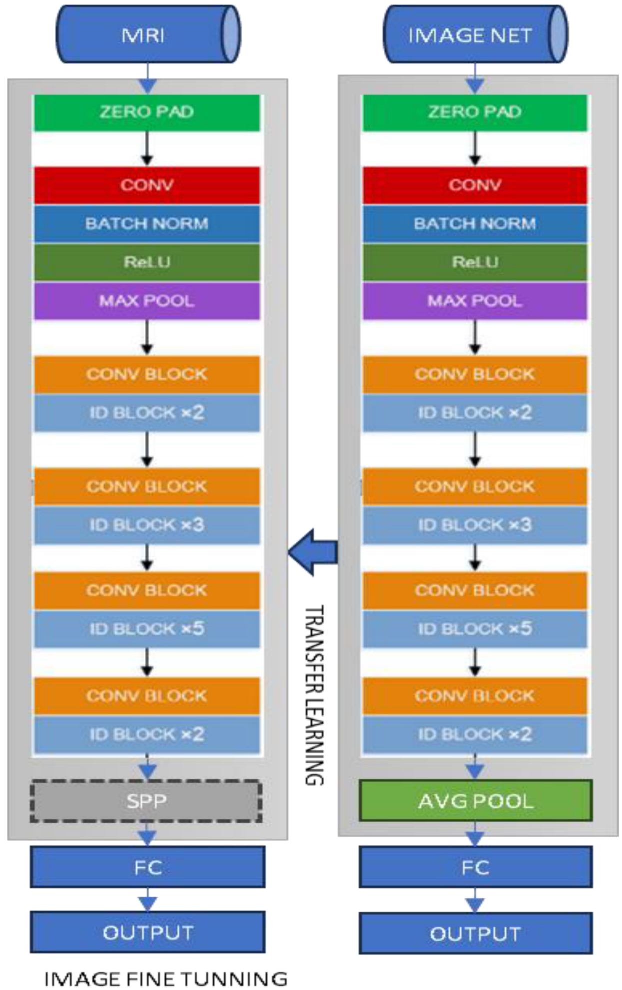

FIGURE 4. Improved CNN-based model for feature extraction.

Fig. 4 illustrates the suggested model, which is founded upon the principles of transfer learning. As evident from the visual representation in Fig. 4, this model harnesses parameters that were initially trained on the expansive ImageNet dataset. Its architecture is enriched by substituting the former pooling layer with a spatial pyramid pooling layer, thereby enhancing its feature extraction capabilities. For the purpose of loss calculation, the output emanating from the ultimate fully connected layer undergoes transformation through a sigmoid function, resulting in the creation of a probabilistic distribution encompassing class labels. The proposed model encompasses two pivotal enhancements, comprising: a) Parameter transfer, b) Spatial Pyramid pooling.

A. MODEL PARAMETER TRANSFER (TRANSFER LEARNING)

While (CNN) models exhibit remarkable prowess in the realm of image classification, the ResNet-50 network's training process, involving a plethora of data, often grapples with a significant proclivity for over-fitting, as elucidated in the literature [20]. Hence, the employment of transfer learning emerges as a pivotal strategy to mitigate this quandary [21].

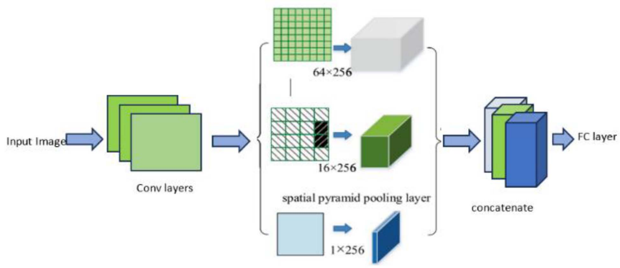

FIGURE 5. **Illustration of the proposed layer.**

The core premise of transfer learning is to judiciously leverage the prior knowledge garnered from pertinent domains to enhance the performance of the model in distinct tasks. In the present context, there exists a presumption that the art of classifying images from the extensive ImageNet dataset is intrinsically linked to the task of discerning MRI scans. Consequently, a pretrained ResNet-50 is employed to grapple with the herculean task of tumor identification across an assemblage of approximately 14 million meticulously labeled images constituting the ImageNet dataset. More explicitly, the gargantuan ImageNet corpus, replete with millions of images spanning a spectrum of 1000 categories, was harnessed as the training ground for ResNet-50. During this arduous training process, the model meticulously acquired the weight and bias parameters. In the subsequent phase of transfer learning, these weight and bias parameters, residing within each layer of the ResNet 50-ImageNet, were carefully preserved, except for the ultimate layer. Subsequently, the input data was juxtaposed with a swift retraining of the neurons exclusive to the ultimate layer, tailor-fitted to the bifurcated categories [22]. Fig. 5 intricately delineates the prescribed transfer learning modus operandi of the model. As observed in the figure, the model adeptly leverages the parameters previously honed through rigorous training on the ImageNet dataset. Furthermore, the model's architecture is astutely augmented by substituting the last pooling layer with a spatial pyramid pooling layer, thereby fortifying its feature extraction capabilities. To ascertain the loss, the output emerging from the final fully connected layer is subjected to the sigmoid cross-entropy loss layer, engendering a probability distribution encompassing an array of class labels. Fine-tuning is imperative, a process wherein the model's parameters are meticulously adjusted based on the already honed parameters residing in ImageNet. This meticulous procedure, in essence, facilitates the training of multi-label images employing a model primarily tailored for single-label images, resulting not only in an expedited training process but also in the potential for efficacious outcomes.

B. SPATIAL PYRAMID POOLING (SPP)

In order to enhance the performance and resilience of the proposed model, particularly in the realm of object recognition across diverse scales, a strategic substitution was made in the architecture [23]. The last pooling layer, which immediately precedes the final convolutional layer, was replaced with a technique known as Spatial Pyramid Pooling (SPP). This method, an extension of the Bag of Words (BoW) approach as detailed in reference [24], has demonstrated remarkable efficacy in the domains of image recognition and classification. SPP functions by partitioning an image into multiple divisions, employing multi-bins, and subsequently aggregating the localized attributes within the image. By amalgamating features extracted at various scales, SPP has the remarkable capability to produce a fixed-length output, irrespective of the input's dimensions. This attribute serves to significantly enhance the precision of our networks in recognizing objects, while simultaneously fortifying their resistance to object deformations.

As depicted in Fig. 5, a tri-tiered pyramid pooling structure has been meticulously crafted, incorporating three bins of distinct sizes, specifically tailored for the feature map generated by the conv5-3 layer. Within each spatial bin, the model undertakes a process of response pooling, employing the max pooling technique on the filter responses. The outcomes of this spatial pyramid pooling operation manifest as vectors residing in a 81 x 256 dimensional space, with 81 denoting the quantity of bins (in alignment with the number of filters in the ultimate convolutional layer, which stands at 256). These 81 bins are meticulously applied to every feature map that emanates from the conv5-3 layer, and the resultant pooled features are systematically concatenated into fixed-dimensional vectors before they are channeled into the fully-connected layer. This architectural design is characterized by its adaptability, capable of accommodating input images of variable sizes, thus affording our (CNN) the invaluable capability to be trained with images of diverse scales. By virtue of spatial pyramid pooling, features sourced from a spectrum of scales can be harmoniously pooled during the training process, thereby considerably enhancing the accuracy of multi-label image classification.

In conducting a comprehensive analysis of the proposed model's architectural refinement, a customary approach involves post-explanatory methods commonly employed in (CNN) evaluations. Leveraging sophisticated techniques inherent in CNN assessments, we delve into nuanced aspects of the model's augmentation, particularly focused on the strategic substitution implemented in the final convolutional layer's preceding pooling stage. The integration of Spatial Pyramid Pooling (SPP), an intricate extension rooted in the Bag of Words (BoW) paradigm, showcases discernible advantages in the domain of image recognition and classification. The model's adaptability to varied input dimensions, illustrated in the tri-tiered pyramid pooling structure in Fig. 5, introduces a notable flexibility in accommodating images of diverse scales. The utilization of response pooling within each spatial bin, coupled with the systematic concatenation of pooled features, culminates in fixed-dimensional vectors that seamlessly interface with the subsequent fully-connected layer. This architectural design not only enhances object recognition

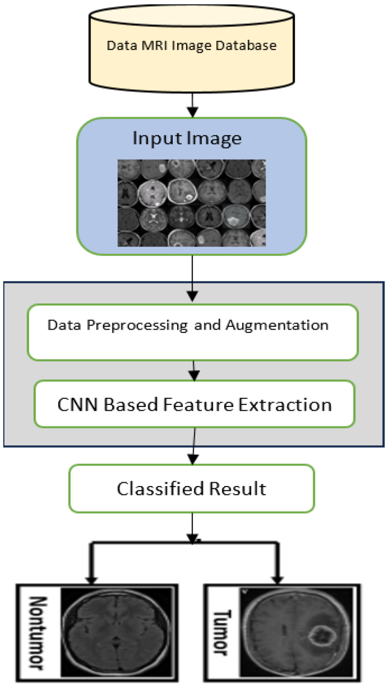

FIGURE 6. **Framework of the proposed model.**

precision but also fortifies the model against deformations. The versatility engendered by spatial pyramid pooling proves instrumental in harmoniously incorporating features from disparate scales during training, thereby substantiating the model's efficacy in multi-label image classification.

C. THE PROPOSED RESNET-50 MODEL

Within this section, we provide a comprehensive exposition of the model we have conceived, one that employs a deep CNN rooted in the ResNet-50 model, leveraging transfer learning for the purpose of brain tumor classification through MRI analysis [25]. The structure of our innovative system unfolds across three primary phases, each of paramount importance: 1) the initial phase involves preprocessing and data augmentation. 2) The ensuing stage encompasses the intricate construction of our model. 3) Finally, classification. Details of the model process is illustrated in Fig. 6.

Utilizing the contrast stretching methodology, we enhance the quality of MRI images in their nascent stages. Subsequently, an extensive dataset is engendered for the (CNN) framework through the employment of data augmentation methods such as rotation and mirroring. This not only aids in mitigating the peril of overfitting but also serves to fortify the model's generalization capacity. In the subsequent phase, we deploy a pre-trained CNN architecture, specifically leveraging the ImageNet dataset [6], to interrogate a target dataset focused on brain tumors. The objective here is to distill distinctive visual attributes from the MRI images. In the final stage, we meticulously categorize the automated features, a pivotal process in facilitating the detection of tumors.

Fig. 6 provides a general view of the proposed model's framework on how data acquires and is classified to yield the results presented above. The working starts with Data MRI Images Database that includes various types of MRI scans to train and test the developed model. The Input Image is the single MRI scans that enter the system. The next step is Data Preprocessing and Augmentation where there is process of normalizing contrast and resizing images to make the data set in better format. Data preprocessing techniques such as image rotation, mirroring, and flipping are also used so as to add more variety to the input data to minimize overtraining. After that, the CNN-Based Feature Extraction phase employs a deep convolutional neural network to extract distinctly useful features from images that have gone through pre-processing. Lastly, the final output is the Classified Result that highlights whether the MRI contains tumor or non-tumor, based on the generated features. It prevents the compromising of accurate tumor detection due to inadequacies in the preprocessing and augmentation utilities, feature extraction, and classification processes in the model.

VI. EXPERIMENTAL RESULTS AND DISCUSSION

This section unfolds the experimental findings of the proposed model, assessed through a diverse set of evaluation metrics encompassing Accuracy, Precision, Recall, F1, Sensitivity, and Specificity.

A. EXPERIMENTAL SETTINGS

Python 3.4 environment has been used to build the proposed Model using the KERAS and Tensorflow backend library. The network is trained using stochastic gradient descent (SGD) [26] with a momentum of 0.9 and weight decay of 0.0005. All the fully-connected layers are initialized from zero-mean Gaussian distributions with standard deviations 0.01. To overcome overfitting, the first two fully-connected layers are followed by a drop-out operation with a drop-out ratio of 0.5.

B. EVALUATION METRICS

The confusion matrix serves as a widely employed technique for assessing the predictive capabilities of a trained model with regard to a specific validation dataset. Within this matrix, rows and columns possess a balanced symmetry, revealing the actual class identities and the ground truth labels, distinguishing between non-tumorous and tumorous instances [27]. The displayed values within the matrix provide insights into the percentage of both accurate and erroneous predictions or classifications for each sample in the validation dataset. True Positives signify the count of correctly identified positive samples, while True Negatives denote the tally of accurately predicted negative samples, see Table 1. The efficacy of deep learning models can be measured by means of a diverse array of criteria, encompassing accuracy, precision, recall, the F1 score, sensitivity, specificity, and the area under the receiver operating characteristic curve, among others. An evaluation of model performance is undertaken, focusing on the overall accuracy, precision, sensitivity, and F1-score for each model.

TABLE 1. Confusion Matrix Definition

Class	Statement	formula
True Positive	Images that belong to a patient who is sick and correctly known by the model.	TP
True Negative	Images that belong to a patient who is healthy and correctly known by the model.	TN
False Positive	Images that belong to a patient who is healthy but diagnosed as sick by the model.	FP
False Negative	Images that belong to a patient who is sick but diagnosed as healthy by the model.	FN

TABLE 2. Comparison Results of the Different Settings of the Model

Model	ACC	PREC	REC	F1	SEN	SPEC
ResNet-50-Basic	0.924	0.903	0.950	0.932	0.9561	0.9401
ResNet-50+SPP	0.942	0.936	0.971	0.961	0.9890	0.952
ResNet-50+TL	0.9812	0.9623	0.9702	0.9650	0.9899	0.9651
ResNet50+SPP + TL	**0.9902**	**0.9837**	**0.9915**	**0.9891**	**1.00**	**0.9930**

These metrics are meticulously computed through a set of specific mathematical equations, yielding valuable insights into the model's predictive prowess.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$F1Score = \frac{2\,(TP)}{2\,(TP) + FP + FN} \tag{5}$$

C. MODEL EVALUATION

Table 2 compares the performance of various configurations of the ResNet-50 model in terms of accuracy (ACC), precision (PREC), recall (REC), F1-score (F1), sensitivity (SEN), and specificity (SPEC). Each configuration represents a different combination of model features and training strategies. In a quest to comprehensively scrutinize the model's performance, various configurations were meticulously engineered in this study, enabling an in-depth comparison of their respective outcomes and thereby discerning the optimal configuration. Table 2. presents these configurations include:

1) ResNet-50, an instantiation of the foundational ResNet model bereft of any transfer learning, coupled with the substitution of the final pooling layer with Spatial Pyramid Pooling (SPP).
2) ResNet+SPP, constituted by the ResNet model where the last pooling layer is replaced by SPP, devoid of fine-tuning.
3) ResNet+TL, which leverages active transfer learning through fine-tuning on the MRI dataset, while preserving the original final pooling layer without SPP substitution.
4) Lastly, ResNet+SPP+TL, embodying the default configuration of the model, where both transfer learning and SPP replacement are diligently executed. Table 2 encapsulates the comparative results of these diverse model configurations.

TABLE 3. Comparison Results of Different CNN Architectures With the Proposed Model

Model	ACC	PREC	REC	F1	SEN	SPEC
VGG-16	0.91	0.834	0.903	0.924	0.932	0.9302
GoggleNet	0.752	0.702	0.732	0.782	0.827	0.8034
LeNet	0.923	0.896	0.941	0.906	0.9517	0.9428
AlexNet	0.905	0.892	0.91	0.924	0.9322	0.9252
(Proposed)	**0.942**	**0.936**	**0.971**	**0.961**	**0.9890**	**0.952**

Analysis of the data in Table 2 unmistakably reveals that the inclusion of the SPP layer yields substantial advancements across all metrics when compared to the baseline ResNet model. This noteworthy performance enhancement underscores the pivotal role played by the SPP layer in accurately capturing essential image features for recognition. Comparing the ResNet-50+SPP model with the ResNet-50+TL model reveals the latter's superiority, demonstrating notable improvements in performance. The default model configuration, ResNet+SPP+TL, which leverages both SPP and transfer learning, clearly outshines other model settings, namely ResNet basic, ResNet+SPP, and ResNet-50+TL. These outcomes underscore the profound impact of combining SPP and transfer learning in the context of image detection tasks.

In Table 3, in order to eliminate the confounding factors in the results, all the experiments conducted in this article have been done using the following standardized experimental settings: Particularly, compare was applied to all the CNNs under study, namely, ResNet-50, VGG16, and InceptionV3 with the same pretrained weights in ImageNet. In addition, the SPP layer was incorporated to each of the models under uniform consideration, despite the difference in performance when used alone. The results depicted in Table 3 are the performance of each model with SPP layer incorporated. Such a methodology ensures that the comparison of various CNN architecture efficiency is done from the same starting point. Thus, by keeping all these conditions standardized, we can guarantee that the comparative evaluation is as objective and as truthful to the fact as possible.

The outcomes of our proposed model are subjected to a comprehensive comparison with established CNN models such as VGGNet [28], LeNet [29], AlexNet [30], and GoogleNet [31] as shown in Table 3. The results, as depicted in Table 3, provide a discerning assessment of various CNN architectures when contrasted with our model, considering metrics like accuracy, precision, F1 score, specificity, and sensitivity. Upon meticulous scrutiny of the experimental data, it becomes evident that the LeNet model exhibits promising performance, characterized by accuracy, especially in comparison to AlexNet and GoogleNet. Nevertheless, our proposed model consistently outperforms not only LeNet but also other renowned architectures, including VGG 16, AlexNet, and GoogleNet, across all evaluated aspects. Essentially, the well-known CNN models, namely LeNet, AlexNet, GoogleNet, and VGG 16, tend to grapple with overfitting challenges, primarily owing to their intricate architectures featuring a substantial number of layers designed for a considerably large number of output classes (1000 classes) with RGB input images. This juxtaposition underscores the distinct advantage of our proposed model in the realm of superior feature extraction for brain tumor classification.

AlexNet, LeNet, GoogleNet, and VGG16 have been successfully employed in various image classification tasks, including medical image analysis. Their track record makes them suitable candidates for comparison, allowing us to gauge how our model performs in comparison to established methods. Each of these architectures has unique design choices and characteristics. Comparing our model with them helps in understanding how different architectural decisions impact performance, providing insights into the strengths and weaknesses of each approach. The comparison with AlexNet, LeNet, GoogleNet, and VGG16 provides a comprehensive framework for evaluating the performance, robustness, and unique contributions of your proposed tumor classification model in a context that is widely recognized within the research community.

Features extracted in our proposed method are selected differently from other methods due to its specificity in tumour diagnosis. Even the general practitioners of image classification, the traditional CNN models prove to be less effective when applied to medical image detection especially tumors as their characteristics are more nuanced. Our method improves feature extraction by the addition of SPP and TL, both of which resolve the issues resulting from MRI scans of brain tumors. It also support the extraction of multi-scale feature which can be very helpful in the detection of tumors of different size and shape. Transfer learning opens up the possibility of using the existing large databases of models which gives our approach an advantage of using the generalized knowledge and apply it to the peculiarities of the brain tumors. More specifically, the proposed method concerns these issues by applying domain adaptation methodologies and optimizing network structures. Thus, by proposing a fine-tuned ResNet-50 model with SPP, an increase in diagnostic accuracy is achieved, which is indispensable in the discovery of tumors. Based on the results that we have obtained, it can be noted that the proposed approach has better performance metrics compared to other approaches used in the study. In order to compare our method with standard models, we have performed sufficiently many experiments; furthermore, it is evident that the integration of SPP and TL yields significant enhancements in accuracy, precision, recall, or equivalently in other evaluation metrics. In this evidence it is shown how our proposed method has made a novel contribution in the improvement of diagnostic capability for the classification of brain tumors. The approach of creating features based on the characteristics of the dataset and the subsequent customisation of images for analysis contributes a valuable development to the existing body of medical image analysis.

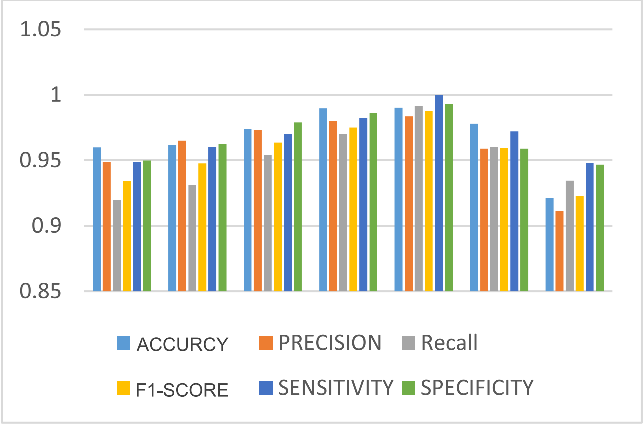

FIGURE 7. Impact of epoch on the model performance.

D. PARAMETER ANALYSIS

Within this section, we embark on a thorough exploration of the repercussions stemming from diverse model configurations. It is worth emphasizing that a meticulous quest has been undertaken to discern the optimal and pivotal hyperparameters within our proposed model, utilizing the grid search methodology. In the realm of hyperparameter fine-tuning, grid search stands as a fundamental technique, offering us the means to achieve optimal settings. Among the array of hyperparameters at our disposal, the learning rate, epoch count, and batch size emerge as the most pivotal determinants that exert profound influence upon our model's performance. Consequently, our investigation, guided by the grid search approach, delves deeply into the consequences of manipulating the epoch count, fine-tuning the learning rate, and altering the batch size, shedding light on their impacts.

E. IMPACT OF EPOCH

The epoch parameter holds a pivotal role in shaping the performance of our training model. An epoch denotes a complete cycle wherein the model processes the entire training dataset. In essence, it signifies a single round of training encompassing all available training data. Generally, an uptick in the number of epochs correlates with an elevation in the model's accuracy score. Nevertheless, it is imperative to be mindful of the propensity for overfitting that emerges when a substantial

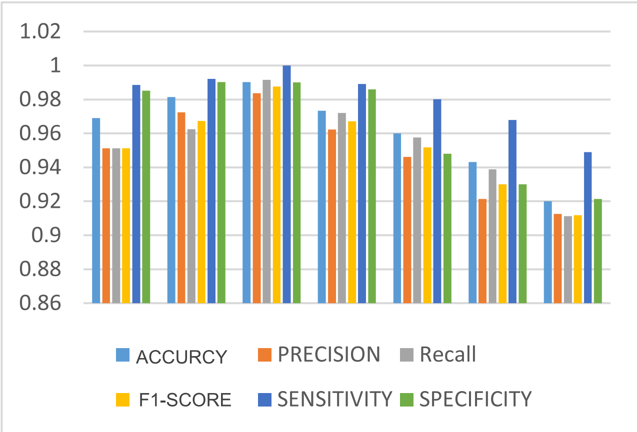

FIGURE 8. Impact learning rate.

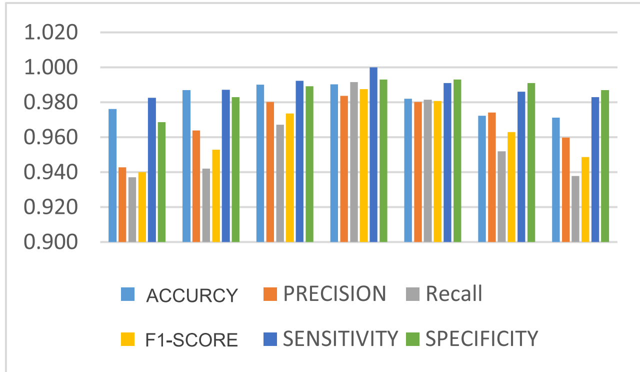

FIGURE 9. Impact of batch size.

number of epochs is utilized. Hence, the selection of the optimal and most efficient value for this parameter is of paramount importance.

The influence of the epoch parameter on our model's accuracy score is visually portrayed in Fig. 7. As the figure vividly illustrates, elevating the number of epochs corresponds to an augmented performance of our proposed model, consequently amplifying the classification accuracy of our methodology. Furthermore, a careful experimental analysis reveals that the zenith of accuracy is achieved when the epoch value is set at 25. Fig. 7 also unequivocally showcases that the training process stalls and leads to a deterioration in performance when the epoch value is set at 50. Thus, the epoch number of 25 is judiciously selected for this experiment.

F. LEARNING RATE

The learning rate emerges as a pivotal hyperparameter, fundamentally dictating the magnitude of the step taken by the optimization algorithm in adjusting the model's parameters during the course of training. It wields control over the convergence pace, that is, how swiftly the model approaches a solution or the optimal parameter set that minimizes the loss function. A high learning rate translates to substantial parameter updates in each iteration, facilitating rapid convergence but also raising the risk of overshooting the optimal solution or even spiraling into divergence, a scenario where the loss function ascends rather than descends. In contrast, a low learning rate translates to minute parameter adjustments in each iteration, which may result in a protracted journey to the optimal solution, prolonging the convergence process. Consequently, pinpointing the precise learning rate assumes paramount importance in attaining optimal training outcomes. An excessively high learning rate can induce instability in the optimization process, leading to a non-converging model. Conversely, an overly low learning rate can result in a sluggish optimization process, potentially trapping the model in a suboptimal solution. Fig. 8 graphically illustrates the impact of this parameter on our model's accuracy rating. As evident from the figure, elevating the learning rate augments the approach's accuracy, thereby enhancing the classification prowess of our method. Furthermore, the results from experimental scrutiny highlight that the optimal accuracy outcome is realized when the learning rate is set at 1-e4. Notably, the Fig. 8 underscores the adverse performance consequences when the learning rate is assigned the value of 1E+00. Consequently, 1-e4 emerges as the judicious selection for the experiment's learning rate

G. IMPACT OF BATCH SIZE

Fig. 9, another crucial parameter in the realm of model training is the batch size. Its value wields substantial influence over the performance of the model, whether that influence is propitious or adverse. The batch size denotes the quantity of training samples that will be recurrently conveyed through the neural network. Indeed, the choice of batch size demands careful deliberation. In this study, a batch size of 32 was selected for training the proposed model. This decision was made after thorough experimentation and analysis, considering the trade-offs between memory utilization, training time, and model convergence. By opting for a batch size of 32, we strike a balance between efficient memory usage and expedited convergence, ensuring that the model can effectively learn from the training data without overwhelming system resources.

Significantly, it exerts a notable impact on the memory requirements for the training process. Indeed, as the batch size expands, the requisite memory space proportionally increases. Conversely, employing a smaller batch size can render the learning process more arduous and erratic, effectively protracting the time required for the technique to converge. Consequently, the selection of the batch size value necessitates meticulous consideration.

VII. COMPARISON METHODS

To enhance the persuasiveness of the experimental findings, we subjected the results of our proposed method to a rigorous comparison with various extant approaches for MR image classification. In pursuit of this objective, our model was meticulously juxtaposed with reference [7], which harnessed a two-channel, low-complexity CNN network for the detection and classification of brain tumors. Additionally, we scrutinized our approach in relation to the methodologies presented

TABLE 4. Comparison Results With Existing Approaches

MODEL	ACC	PRE	REC	F1	SEN	SPE
MCNN [7]	0.9821	0.967	-	0.983	1	0.963
DensNet [8]	0.96	0.933	-	-	0.933 3	0.971 4
VGG-SCNet [9]	-	0.992	0.991	0.992	-	-
VGG-16 [10]	0.96	0.94	**1**		-	-
EfficientNet -B0 [32]	0.988	0.994	0.995	0.989	0.995	0.992
Proposed model	**0.9902**	**0.996**	0.9915	**0.9991**	**1**	**0.993**

in [8], featuring a DenseNet CNN model, and [9], which leveraged the VGG Net-Based framework for brain tumor detection on MRI images. Furthermore, our model underwent a thorough assessment against the techniques outlined in [10] and [32], which incorporated deep transfer learning and Fine-tuned EfficientNet, respectively, for brain tumor classification based on MRI data.

The findings, elucidated in Table 4, showcase the comparative outcomes of our proposed model vis-à-vis these existing approaches, encompassing metrics such as precision, accuracy, F1 score, specificity, and sensitivity. It is discernible from the results that our proposed model consistently outperformed the comparator approaches across all evaluated criteria.

Hence, based on the insights garnered from this section, we can deduce that the presented model proves highly efficacious in facilitating the discrimination between Glioma and non-pathological tissues.

The proposed model is compared to some of the already existing techniques for brain tumor classification in Table 4, with different models' accuracy and performance where the MCNN, DenseNet, VGG-SCNet, VGG-16, EfficientNet-B0 and the proposed model have been compared. The models selected for evaluation are ACC, PRE, REC, F1, SEN, and SPE. The elaborated model indicates the highest efficiency for all the proposed criteria, where accuracy is equal to 0. 9902, and the precision, recall, F1 score, and sensitivity are equal to 0. 996, 0. 991, 0. 999, and 1, respectively. Similarly, although the EfficientNet-B0 and VGG-16 are also very high in terms of recall especially for the model, it is clearly higher in terms of global accuracy and the F1 score than the proposed model. For example, the VGG-16 attains the highest recall but lags behind when it comes to the proposed model's F1 score. This argumentative structure reinforces the solidity and efficiency of the hereby proposed model which results from the use of more sophisticated and effective methods such as SPP and transfer learning, which all improve the model's performance in tumor classification tasks.

VIII. CONCLUSION

The utilization of magnetic resonance imaging (MRI) in the realm of brain tumor identification has surged in popularity due to the burgeoning need for a practical and precise assessment of extensive medical datasets. The manual discernment of brain tumors, while reliant on specialized medical expertise, proves to be a time-intensive endeavor. Given the life-threatening nature of brain tumors, the imperative for an automated diagnostic approach becomes evident. Our research has thus culminated in the development of a sophisticated and optimized (CNN) model that hinges on transfer learning for the Classification of brain malignancies within MRI data. In contrast to prevailing methodologies, our proposed approach yielded superior results in the realm of brain tumor Classification. The role of preprocessing was paramount in ensuring data quality, with pivotal steps encompassing intensity normalization, contrast enhancement, and image resizing, all of which served as crucial precursors for the subsequent extraction of image features. The augmentation of our training dataset through techniques such as rotation and flipping not only fortified the dataset but also proved instrumental in mitigating overfitting. A comprehensive comparative analysis, juxtaposed against well-established CNN architectures like VGGNet, AlexNet, and DensNet, accentuated the superior performance of our devised approach. In addition to surpassing various extant methodologies across diverse metrics, including accuracy, precision, recall, F1-score, sensitivity, and specificity, our model underscores its potential as a robust and indispensable tool for the nuanced field of medical image analysis. In summation, this study represents a substantial contribution to the domain of brain tumor identification within MRI data.

REFERENCES

[1] G. S. Tandel et al., "A review on a deep learning perspective in brain cancer classification," *Cancers*, vol. 11, no. 1, 2019, Art. no. 111.
[2] M. Masood et al., "A novel deep learning method for recognition and classification of brain tumors from MRI images," *Diagnostics*, vol. 11, no. 5, 2021, Art. no. 744.
[3] M. Z. Hossain, F. Sohel, M. F. Shiratuddin, and H. Laga, "A comprehensive survey of deep learning for image captioning," *ACM Comput. Surv.*, vol. 51, no. 6, pp. 1–36, 2019.
[4] H. Mzoughi et al., "Deep multi-scale 3D convolutional neural network (CNN) for MRI gliomas brain tumor classification," *J. Digit. Imag.*, vol. 33, pp. 903–915, 2020.
[5] W. F. A. Al-Shammari and S. Saon, "Classification technique for brain tissues diagnoses: Segmenting healthy, cancer affected and edema brain tissues," *Evol. Inf., Commun. Comput. System*, vol. 3, pp. 103–119, 2022.
[6] K. Neamah, F. Mohamed, M. Adnan, and S. Thajeel, "A survey on brain tumor diagnosis and edema detection based on machine learning," in *Proc. J. Phys.: Conf. Ser.*, vol. 1892, 2021, Art. no. 012040.
[7] N. Kesav and M. G. Jibukumar, "Efficient and low complex architecture for detection and classification of brain tumor using RCNN with two channel CNN," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 8, pp. 6229–6242, 2022, doi: 10.1016/j.jksuci.2021.05.008.
[8] S. Sharma et al., "Deep learning model for automatic classification and prediction of brain tumor," *J. Sensors*, vol. 2022, pp. 1–11, 2022.
[9] M. S. Majib, M. M. Rahman, T. S. Sazzad, N. I. Khan, and S. K. Dey, "VGG-SCNet: A VGG net-based deep learning framework for brain tumor detection on MRI images," *IEEE Access*, vol. 9, pp. 116942–116952, 2021.
[10] C. Srinivas et al., "Deep transfer learning approaches in performance analysis of brain tumor classification using MRI images," *J. Healthcare Eng.*, vol. 2022, 2022, Art. no. 3264367.
[11] M. Khan, S. A. Shah, T. Ali, A. Khan, and G. S. Choi, "Brain tumor detection and segmentation using RCNN," *Comput., Mater. Continua*, vol. 71, no. 3, 2022, Art. no. 5005.

[12] M. K. Elbashir, M. Ezz, M. Mohammed, and S. S. Saloum, "Lightweight convolutional neural network for breast cancer classification using RNA-Seq Gene expression data," *IEEE Access*, vol. 7, pp. 185338–185348, 2019.
[13] G. Wang, W. Li, S. Ourselin, and T. Vercauteren, "Automatic brain tumor segmentation based on cascaded convolutional neural networks with uncertainty estimation," *Front. Comput. Neurosci.*, vol. 13, 2019, Art. no. 56.
[14] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, pp. 1–48, 2019.
[15] J. Wei and K. Zou, "Eda: Easy data augmentation techniques for boosting performance on text classification tasks," 2019, *arXiv:1901.11196*.
[16] A. Kumar et al., "Deep feature learning for histopathological image classification of canine mammary tumors and human breast cancer," *Inf. Sci.*, vol. 508, pp. 405–421, 2020.
[17] A. Rehman, S. Naz, M. I. Razzak, F. Akram, and M. Imran, "A deep learning-based framework for automatic brain tumors classification using transfer learning," *Circuits, Syst., Signal Process.*, vol. 39, pp. 757–775, 2020.
[18] A. Binder, S. Bach, G. Montavon, K. -R. Müller, and W. Samek, "Layer-wise relevance propagation for deep neural network architectures," in *Proc. Inf. Sci. Appl.*, New York, NY, USA, 2016, pp. 913–922.
[19] S. Bharati, P. Podder, and M. Mondal, "Artificial neural network based breast cancer screening: A comprehensive review," 2020, *arXiv:2006.01767*.
[20] N. Tripathi and A. Jadeja, "A survey of regularization methods for deep neural network," *Int. J. Comput. Sci. Mobile Comput.*, vol. 3, no. 11, pp. 429–436, 2014.
[21] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, 2009.
[22] S. R. Waheed et al., "Melanoma skin cancer classification based on CNN deep learning algorithms," *Malaysian J. Fundam. Appl. Sci.*, vol. 19, no. 3, pp. 299–305, 2023.
[23] S. R. Waheed, M. S. M. Rahim, N. M. Suaib, and A. Salim, "CNN deep learning-based image to vector depiction," *Multimedia Tools Appl.*, vol. 82, pp. 1–20, 2023.
[24] Q. Le and T. Mikolov, "Distributed representations of sentences and documents," in *Proc. Int. Conf. Mach. Learn.*, 2014, pp. 1188–1196.
[25] K. T. Khudhair, F. H. Najjar, S. R. Waheed, H. M. Al-Jawahry, H. H. Alwan, and A. Al-khaykan, "A novel medical image enhancement technique based on hybrid method," in *Proc. J. Phys.: Conf. Ser.*, vol. 2432, 2023, Art. no. 012021.
[26] L. Bottou, "Stochastic gradient descent tricks," in *Neural Networks: Tricks of the Trade: Second Edition*. New York, NY, USA: Springer, 2012, pp. 421–436.
[27] S. R. Waheed, N. M. Suaib, M. S. M. Rahim, M. M. Adnan, and A. Salim, "Deep learning algorithms-based object detection and localization revisited," in *Proc. J. Phys.: Conf. Ser.*, vol. 1892, 2021, Art. no. 012001.
[28] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[29] A. Bouti, M. A. Mahraz, J. Riffi, and H. Tairi, "A robust system for road sign detection and classification using LeNet architecture based on convolutional neural network," *Soft Comput.*, vol. 24, no. 9, pp. 6721–6733, 2020.
[30] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, vol. 25, pp. 1097–1105.
[31] C. Szegedy et al., "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1–9.
[32] H. A. Shah, F. Saeed, S. Yun, J. -H. Park, A. Paul, and J. -M. Kang, "A robust approach for brain tumor detection in magnetic resonance images using finetuned EfficientNet," *IEEE Access*, vol. 10, pp. 65426–65438, 2022, doi: 10.1109/access.2022.3184113.

KARRAR NEAMAH received the B.E. degree in computer science from Alkufa University, Kufa, Iraq, in 2011, and the M.S. degree in computer science in 2014 from Universiti Teknologi Malaysia (UTM), Johor Bahru, Malaysia, where he is currently working toward the Ph.D. degree. His research interests include deep learning, image processing, and machine learning.

FARHAN MOHAMED (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees from Universiti Teknologi Malaysia (UTM), Johor Bahru, Malaysia, and the Ph.D. degree in computer science from Swansea University, Swansea, Wales, in 2014. He is currently a Senior Lecturer with the School of Computing, Faculty of Engineering, UTM. He is also a Research Fellow with the Media and Games Innovation Centre of Excellence (MaGICX), Institute of Human Centred Engineering, UTM. He led a study with SME Corp-Huawei in analyzing SME digitalization, in 2018. His current works emphasize visual analytics, virtual environment, and procedural computer graphics. He was the recipient of the Leaders in Innovation Fellowship from Newton Fund, Royal Academy of Engineering, U.K., in 2017. He is on the Executive Committee of the IEEE Computer Society Malaysia.

SAFA RIYADH WAHEED was born in Iraq. He received the master's degree in computer science from University Technology Malaysia (UTM), Johor Bahru, Malaysia. He is currently a Computer Science Lecturer with the Department of Computer Technology-Engineering, UTM. He is also a Computer Science Teacher, an Author, and a Researcher. His work spans several sub-fields of computer science, including computer vision, digital image processing, medical image processing, data and image security, computer networking, real-time applications, and the public understanding of computer science.

WALEED HADI MADHLOOM KURDI received the B.Sc. degree in computer engineering from Islamic University, Najaf, Iraq, and the master's degree from MMU, India.

KARRAR ABDULAMEER KADHIM received the bachelor's degree in computer science from Kufa University, Kufa, Iraq, in 2010 and the M.Sc. degree from the Faculty of Computing, University of Technology Malaysia, Johor Bahru, Malaysia, in 2013. He is currently an Assistant Lecturer with Computer Engineering Teachings, Imam Jaafar Al-Sadiq University, Baghdad, Iraq. His research interests include computer vision, digital imag]e processing, medical image processing, data and image security, computer networking, and real time applications.

ADIL YASEEN TAHA received the B.E. degree in computer science from Alkufa University, Kufa, Iraq, in 2011, the M.S. degree in computer science in 2016, and the Ph.D. degree from the Universiti Kebangsaan Malaysia (UKM), Johor Bahru, Malaysia. He is currently a Lecturer with Medical Instrumentation Engineering Techniques Department, Al-Safwa University College, Karbala, Iraq. His research interests include natural language processing, machine learning, information retrieval, and data mining.