



Classification of orbital tumors using convolutional neural networks

Esraa Allam¹ · Abdel-Badeeh M. Salem¹ · Marco Alfonse^{1,2}

Received: 20 March 2023 / Accepted: 13 December 2023 / Published online: 17 January 2024
© The Author(s) 2024

Abstract

Orbital tumors are the most common eye tumors that affect people all over the world. Early detection prevents the progression to other regions of the eye and the body. Also, early identification and treatment could reduce mortality. A computer-assisted diagnosis (CAD) system to help physicians diagnose tumors is in great demand in ophthalmology. In recent years, deep learning has demonstrated promising outcomes in computer vision systems. This work proposes a CAD system for detecting various forms of orbital tumors using convolutional neural networks. The system has three stages: preprocessing, data augmentation and classification. The proposed system was evaluated on two datasets of magnetic resonance imaging (MRI) images containing 1404 MRI T1-weighted images and 1560 MRI T2-weighted images. The results have shown that the system is capable of detecting and classifying the tumor in each image type, and the recognition rate for the T1-weighted image is 98% and for the T2-weighted image is 97%.

Keywords Deep learning · Convolutional neural networks · Orbital tumor · Image classification · Eye tumors

1 Introduction

The bony cavity known as the orbit contains the eye's controlling muscles, nerves, and blood vessels in addition to enclosing and safeguarding the eyeball. Orbital tumors are irregular tissue growths in the tissues that surround the eye [1]. These lesions may be either benign or malignant. Tumors involving the orbit are divided into primary orbital tumors and secondary orbital tumors which expand from other regions into the orbit. Primary intraocular tumors begin on the inside of the globe. Melanoma, followed by primary intraocular lymphoma, is the most frequent primary eye malignancy in adults. Retinoblastoma (cancer

that begins in the retinal cells) is the most frequent primary intraocular cancer in children, followed by medulloepithelioma (but is still extremely rare). Secondary intraocular tumors begin outside the globe and spread inside it. These are not really “eye cancers,” but they are more common than primary intraocular cancers. Breast and lung cancers are the most common primaries that send secondaries to the eye. The orbital tumor has many symptoms and various signs; these are depending on the type and the site of the tumor. Most patients notice bulging of the eyeball, double or loss of vision, and droopy or swollen eyelid. In some cases, the patient complains of pain because of infections and inflammation [2].

Imaging investigations are vital for diagnosing orbital tumors. Magnetic resonance imaging (MRI) and computed tomography (CT) scans are used as detection methods. MRI images are better than CT in detecting soft tissues and demonstrating the proximity of orbital lesions to the optic nerve [3]. It can diagnose lesions by providing a clear image of structures within the orbit. MRI has two image types: T1-weighted and T2-weighted images to produce multidirectional images demonstrating structures, tissues and their anatomical interactions.

The lesion shows a hypointense signal to fat and an isointense signal to extraocular muscle on T1-weighted MRI images, while a hyperintense signal to fat on T2-

✉ Esraa Allam
esraa.Nabeeh@cis.asu.edu.eg

Abdel-Badeeh M. Salem
abSalem@cis.asu.edu.eg

Marco Alfonse
marco_alfonse@cis.asu.edu.eg

¹ Computer Science Department, Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt

² Laboratoire Interdisciplinaire de l'Université Française d'Égypte (UFEID LAB), Université Française d'Égypte, Cairo, Egypt

weighted MRI images. This feature facilitates the identification of the type of lesions and increases diagnostic accuracy. A biopsy could be performed so that the tumor tissue could be microscopically studied for a conclusive diagnosis [4].

A vast area of computer science called artificial intelligence (AI) is focused on developing smart machines that can carry out tasks that would ordinarily need human intelligence. In the medical industry, machine-learning models are utilized to explore medical records and provide insights to enhance patient care and health outcomes. In both research and clinical contexts, physicians are supported by AI algorithms and other AI-powered applications. The most widely used applications of AI in medical contexts today are clinical decision support and image analysis. Clinical decision support methods assist doctors in making decisions regarding treatments, drugs and other patient requirements by giving rapid access to pertinent information or research. In medical sciences, AI applications include patient diagnosis, prognosis, and medication development. Also, AI techniques are used in image analysis to examine CT scans, MRIs, X-rays and other medical images for lesions or other diseases which radiologists could overlook [5].

Computer-aided diagnosis (CAD) is a technology or application that assists doctors in understanding and interpreting medical images. CAD has gained popularity as a useful tool for supporting clinical decisions for several diseases [6]. Diagnostic imaging modalities such as X-ray, MRI, endoscopy, and ultrasound generate a large amount of data that a radiologist or other medical practitioner must extensively examine, evaluate and interpret in a short amount of time. CAD systems analyze digital images or videos for common features and critical aspects, like possible diseases, and offer feedback to the physician. A frequent sort of application is tumor detection or classification [7].

In this article, a CAD system has been proposed for the detection of orbital tumors. Our contribution is developing a new system using convolutional neural networks (CNN) to classify orbital tumors from MRI images. Preprocessing and data augmentation are applied to the images to improve the performance of the system. Following is how the paper is organized: The related work is described in Sect. 2. The background is shown in Sect. 3. The proposed system is illustrated in detail in Sect. 4. The study's findings are displayed in Sect. 5. In Sect. 6, the interpretation

of results and discussion is presented. The conclusion and the scope for further research are presented in Sect. 7.

2 Related works

There are many studies on the classification of eye tumors. Biswarup et al. [8] demonstrated a technique for detecting ocular melanoma in medical images. The system is separated into three steps: image pre-processing, image labeling, and classification. In the first stage, images were rescaled, cropped showing only the eye image's middle section, and enlarged appropriately throughout the preparation processes. Then, the images were labeled as melanoma or non-melanoma by medical experts. CNN is employed as the classifier in the classification stage. The CNN model is made up of two convolutional layers, two sub-sampling layers, and three fully connected layers. The dataset was gathered from the New York Eye Cancer Centre database. It has 170 images, 110 of which are ocular melanoma and 60 are normal. Ophthalmologists validated the suggested method, which has a 91.76% accuracy rate.

Sheshang et al. [9] developed a system for detecting ocular melanoma in medical images. Preprocessing, segmentation, and identification are the system's three stages. Image preparation techniques such as grayscale transformation and the median filter were used. Otsu segmentation was applied as a thresholding method to images during the segmentation step to convert a gray image to a monochrome which is a standard activity for image cessation. Otsu process is one of the binarization algorithms. During the classification stage, CNN is used as a classifier. The CNN model is made up of four convolutional layers, four maximum pooling layers, and two fully connected layers. The dataset was gathered from two publicly accessible websites, "Miles Research" [10] and "Eye Cancer" [11]. There are 200 images in the collection. The system was examined on its capacity to correctly detect eye melanoma which results in a high rate of 92.5%.

Parmod et al. [12] created a classification system for detecting retinoblastoma in fundus images. The system's two stages are segmentation and classification. The Otsu multi-thresholding approach was applied to segment the tumors from the fundus images. The next stage was to categorize the retinoblastoma in fundus images using the AlexNet and ResNet50 deep learning models. The dataset was obtained from the MathWorks [13] website which

contains 278 fundus images. The system recognized retinoblastoma with an accuracy rate of 93.16% for the ResNet50 and 88.12% for the AlexNet.

3 Background

Deep learning (DL) is a new methodology that has huge growth and evolution in medical fields. This opened a new door for medical image analysis. DL applications in healthcare tackle a wide range of concerns, including cancer classification, segmentation, detection, screening and infection monitoring to individualized therapy recommendations. Nowadays, a huge amount of raw data is given at the disposal of physicians in various data forms such as radiological imaging, genetic sequencing, and pathological imaging. It offers the possibility of transforming all of this knowledge into usable information. DL has recently achieved high performance over human performance on tasks such as image classification [14]. In general, feature extraction is accomplished with better performance by DL techniques. For that purpose, researchers focus on using DL approaches to extract discriminative features using the least amount of human effort and field experience.

A CNN is a deep learning model used to identify essential features in data with a grid pattern, such as images. It has become essential in various computer vision tasks and is gaining interest in a wide range of fields. Nowadays, DL is quite popular because of CNN [14]. A common CNN architecture composes convolution and pooling layers, followed by single or several fully connected layers. To enhance CNN performance, numerous regulatory units such as batch normalization and dropout are added in addition to different mapping functions. A convolution process with no padding, a size of kernel 3×3 , and a stride of 1 is illustrated in Fig. 1.

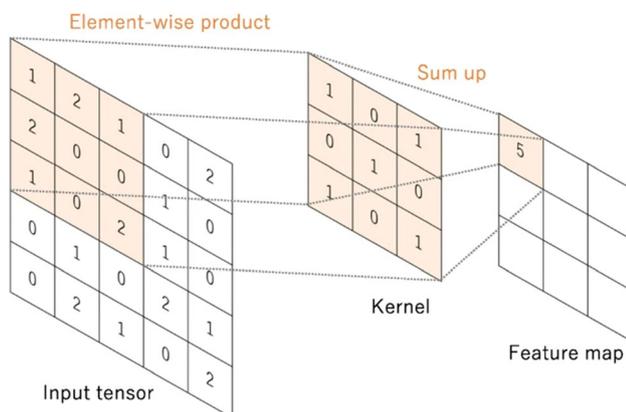


Fig. 1 A convolution process with a kernel size of 3×3 , no padding, and a stride of 1 [15]

The arrangement of CNN components is critical in building new architectures, thereby obtaining improved performance [16]. The initial layer utilized to extract the various features from the input images is the convolutional layer. The mathematical process of convolution is executed between the input and a filter in this layer. Swiping the filter across the input computes the dot product between the filter and the regions of the input based on the filter's size. This layer produces a feature map, which provides important image information, such as the edges and corners. The following layers learn distinctive features from the input image using this feature map. Convolutional operations have kernels as a parameter where kernel size, number of kernels, stride, padding, and activation functions are hyperparameters [17].

A pooling layer downsamples the feature maps in-plane to introduce translation invariance to minor shifts and distortions and to reduce the number of learnable parameters. Despite the fact that filter size, padding, and stride are hyperparameters in pooling operations, similar to convolution processes, no learnable parameters exist in pooling layers. The last layer is the fully connected layer, which transforms the output feature maps into a vector or one-dimensional array [17]. The model can be connected to one or more fully connected layers, with a learnable weight connecting each input to each output. In classification tasks, this layer transfers the output into probabilities for each class.

4 The proposed system

This section illustrates the pipeline to implement our proposed system. The pipeline is divided into several stages; starting with image acquisition, image preprocessing, image augmentation, model training and ending with model evaluation. The proposed system pipeline is depicted in Fig. 2.

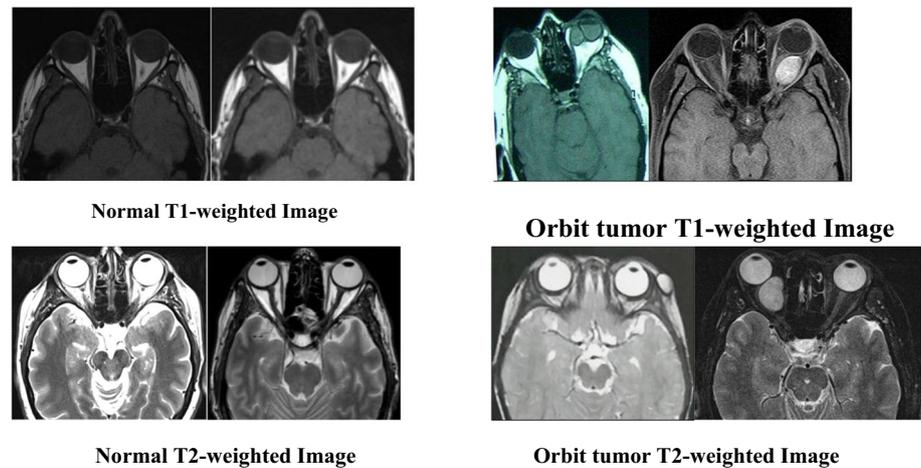
4.1 Image acquisition

We used two datasets in our experiment. Both types of images of orbital tumors were collected from Ain Shams University Hospital, Ophthalmology Department. The normal images were collected manually from different articles from a public website [18]. They consist of different MRI images: T1-weighted and T2-weighted images. The total size of T1-weighted images is 133 MRI images: 20 normal images and 113 orbit tumor images. The total size of T2-weighted images is 109 MRI images: 24 normal images and 85 orbit tumor images. The tumor datasets consist of 30 patients that have different orbital tumor types. The tumor T1-weighted image dimension is

Fig. 2 The proposed system pipeline



Fig. 3 Sample of MRI images (Normal and Orbit tumor)



(800 × 600) pixel resolution, and the normal image is (509 × 400) pixel resolution. The tumor T2-weighted image dimension is (824 × 800) pixel resolution, and the normal image is (516 × 471) pixel resolution. Figure 3 shows the different MRI image samples used in the experiment.

4.2 Image preprocessing

In our work, datasets had small samples that were not enough to be used in the training phase. In order to prevent overfitting in the experiment, one of the resampling techniques was applied. First, oversampling technique was implemented for each type of MRI image (normal and tumor) to upgrade the number of images that were used in the classification. Many image processing techniques were applied including Gaussian blur filter, Median filter, unsharp mask filter, sharpening filters, edge enhancement, brightness filter and contrast filter.

Gaussian blur filter is a low pass filter used for reducing noise (high-frequency components) and blurring regions of an image and it is calculated by Eq. 1 [19].

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

where σ is the standard deviation. In the context of filtering, the standard deviation σ is a parameter, which determines the width of the filter.

Median filtering is a nonlinear method used to remove noise from images. It is widely used as it is very effective at removing noise while preserving edges. It is particularly effective at removing ‘salt and pepper’ type noise. The median filter works by moving through the image pixel by

pixel, replacing each value with the median value of neighboring pixels. The pattern of neighbors is called the “window,” which slides pixel by pixel, over the entire image. [19]. The unsharp mask filter is an extremely versatile sharpening tool that improves the definition of fine detail by removing low-frequency spatial information from the original image. It involves the subtraction of an unsharp mask from the specimen image. An unsharp mask is simply a blurred image produced by spatially filtering the specimen image with a Gaussian low-pass filter [20]. The sharpening filters are a type of image processing that enhances the contrast between neighboring pixels, making the edges and details more visible and defined. Sharpening filters can also reduce the effects of noise, blur, or compression artifacts that degrade the quality of digital images. Edge enhancement is an image processing filter that enhances the edge contrast of an image or video in an attempt to improve its acutance (apparent sharpness). The filter works by identifying sharp edge boundaries in the image, such as the edge between a subject and a background of a contrasting color and increasing the image contrast in the area immediately around the edge. This has the effect of creating subtle bright and dark highlights on either side of any edges in the image, called overshoot and undershoot, leading the edge to look more defined when viewed from a typical viewing distance [21]. Brightness is a relative term defined as the intensity of a pixel relative to another pixel. To increase the brightness, we need to increase the intensity of each pixel by a constant and similarly to darken the image we need to decrease the intensity of every pixel of the image [22]. Contrast can simply be explained as the difference between maximum

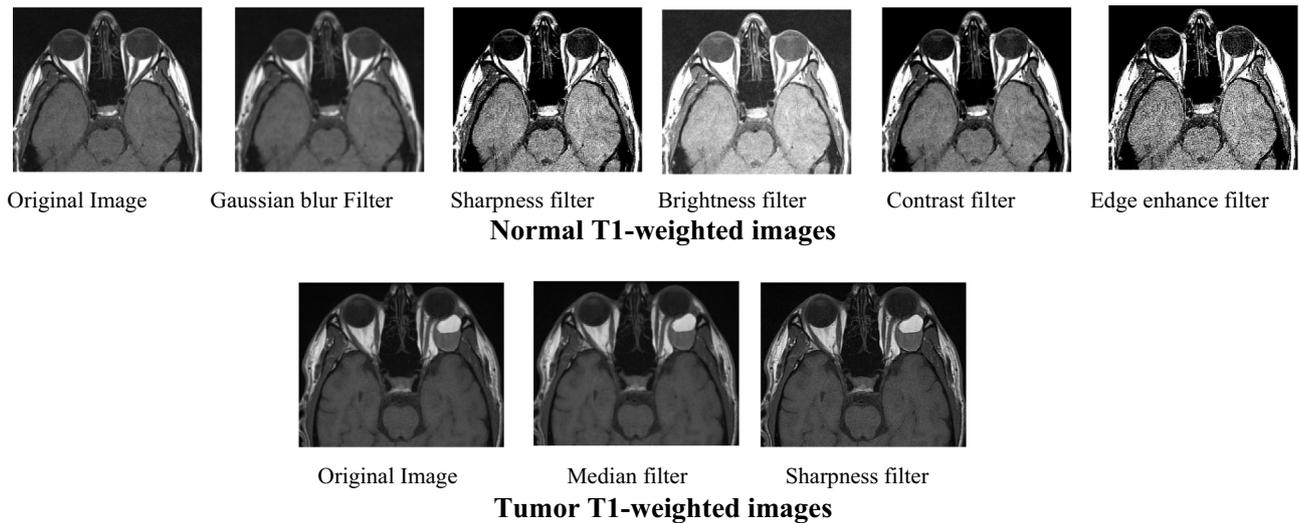


Fig. 4 Sample of MRI T1-weighted images with filter (Normal and Orbit tumor)

and minimum pixel intensity in an image. To change the contrast of an image we just need to change the value of the max and min intensity pixels [23].

For normal type, Gaussian blur filter, Median filter, Unsharp mask filter with radius size (0.5, 1, 2 and 3), sharpness filter with factor (0.5, 1, 1.5, 2, 3 and 4), EDGE Enhance Filter, EDGE Enhance More Filter, Brightness Filter with factor (0.5) and Contrast Filter with factor (0.5) were applied to images. For tumor type, median filter, sharpness filter with factors (0.5, 1, 3 and 4) and Unsharp Mask filter with radius size (1.5) were executed on images. Figure 4 shows samples for MRI T1-weighted images after applying the image filter. Figure 5 shows samples for MRI T2-weighted images after applying the image filter.

The total number of T1-weighted images is 1404 MRI images: 702 are normal and 702 are tumor images, and the total number of T2-weighted images is 1560 MRI images: 780 are normal and 780 are tumor images. Tables 1 and 2 show the total number of MRI images after applying the image processing techniques.

Second, the datasets in our experiment were split into training, validation and testing sets with a ratio of 80% and 20% rule; 10% is validation and 10% is testing. The 80–20 split ratio is prevalently used in deep learning and is used frequently in medical images. This 80–20 split technique is known as “image-level approach”. In this split, each patient image is included in either the training or validation set, but not in both and this is known as “patient-level method” [24].

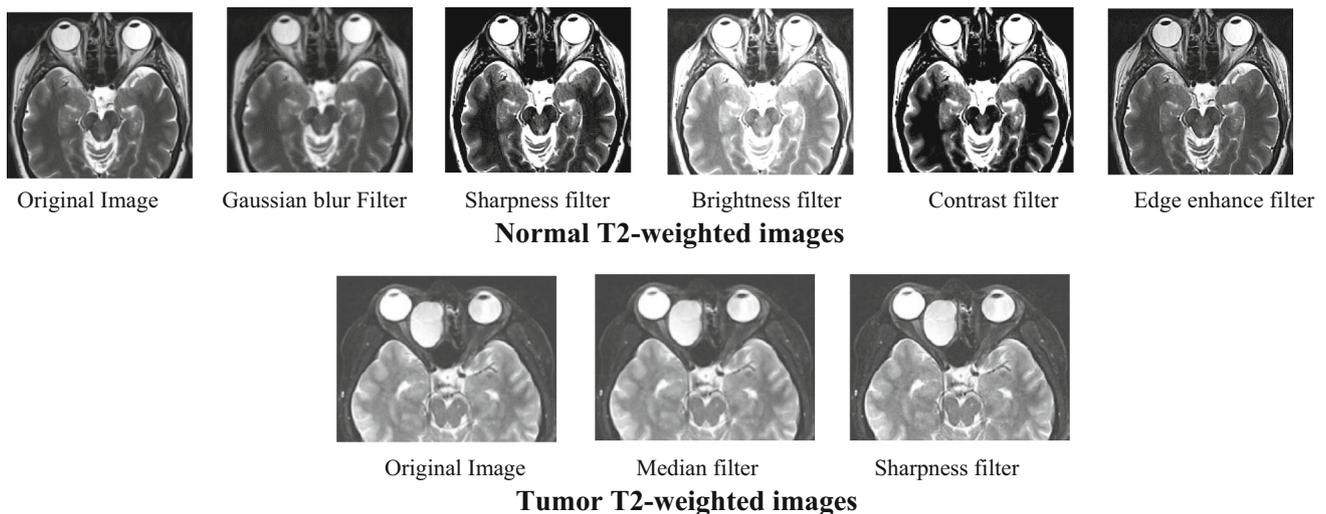


Fig. 5 Sample of MRI T2-weighted images with filter (Normal and Orbit tumor)

Table 1 The total number of T1-weighted MRI images after applying image processing techniques

	Normal T1-weighted images	Tumor T1-weighted images
Original images	20	133
Gaussian blur filter	40	–
Median filter	60	266
Unsharp mask filter with radius size (0.5, 1, 2 and 3)	140	399
Sharpness filter with factor (0.5, 1, 1.5, 2, 3 and 4)	462	702
EDGE Enhance Filter	522	–
EDGE Enhance More Filter	582	–
Brightness Filter	642	–
Contrast Filter	702	–

Table 2 The total number of T2-weighted MRI images after applying image processing techniques

	Normal T2-weighted images	Tumor T2-weighted images
Original images	24	85
Gaussian blur filter	48	–
Median filter	72	196
Unsharp mask filter with radius size (0.5, 1, 2 and 3)	167	337
Sharpness filter with factor (0.5, 1, 1.5, 2, 3 and 4)	386	780
EDGE Enhance Filter	513	–
EDGE Enhance More Filter	640	–
Brightness Filter	710	–
Contrast Filter	780	–

4.3 Image augmentation

Many medical imaging collections in clinical settings suffer from imbalanced data, and it is rare to find a large amount of data for a specific clinical case. The skewed data distribution is an essential issue that occurs frequently in medical image classification problems. One of the methods to overcome the insufficient amount of data utilized in the training phase is data augmentation [25]. It is a method of increasing the dataset size by implementing various

modifications such as rescaling, resizing and rotation of current data in the runtime to generate new samples while keeping the same label. Augmenting images increases the overall amount of images available to the model, permitting it to learn more effectively [26]. Data augmentation has been deemed as a sort of dataset regularization since it reduces over-fitting and improves overall performance by enriching the training dataset itself [27]. Moreover, data augmentation is used to solve the issue of imbalanced classification by oversampling the minority class to enlarge the datasets to make the model perform better on the training data [28].

In our experiment, the dataset has small samples and we used data augmentation to increase the amount of data. To prevent over-fitting in the training phase, we applied some geometric transformation techniques to the MRI images to enhance the training phase. First, the images were resized which changed the width and height to 224×224 pixels. Second, images were rescaled which transformed the pixel range from $[0, 255]$ to $[0, 1]$. Third, images were rotated at 40° degrees. Fourth, the width and height of images had shifted by 20%. Fifth, images were sheared by 20%. Sixth, images were zoomed in by 20%. Finally, images were flipped horizontally which reflect the images around the central horizontal axis.

After applying image augmentation, the total number of T1-weighted images is 11,232 MRI images and the total number of T2-weighted images is 12,480 MRI images. Table 3 shows the total number of MRI images after applying image augmentation techniques.

4.4 The architecture of the CNN model

Our CNN model is split into fifteen sequence levels. It was trained with 80 epochs and 64 batch sizes. The optimizer type of CNN is ‘Adam,’ the loss function is binary cross-entropy, and the learning rate is 0.01. Figure 6 illustrates the CNN model architecture.

Layer 1 in our CNN architecture is the convolutional layer with the ReLu (Rectified Linear Unit) activation function. This layer obtains the pre-processed image as an input with a size $n*n = 224*224$. The convolutional kernel size (filter size) is $f*f = 3*3$, there is no padding (p), stride (s) is 1 and the number of filters (neurons) is 32. We get feature maps of size $32@222*222$ after this convolution process, where 32 is the number of neurons and 222 is the

Table 3 The total number of MRI images after applying image augmentation techniques

	Original images	Resizing	Rescaling	Rotation	Shifting	Shearing	Zooming	Flipping horizontally
T1-weighted images	1404	2808	4212	5616	7020	8424	9828	11,232
T2-weighted images	1560	3120	4680	6240	7800	9360	10,920	12,480

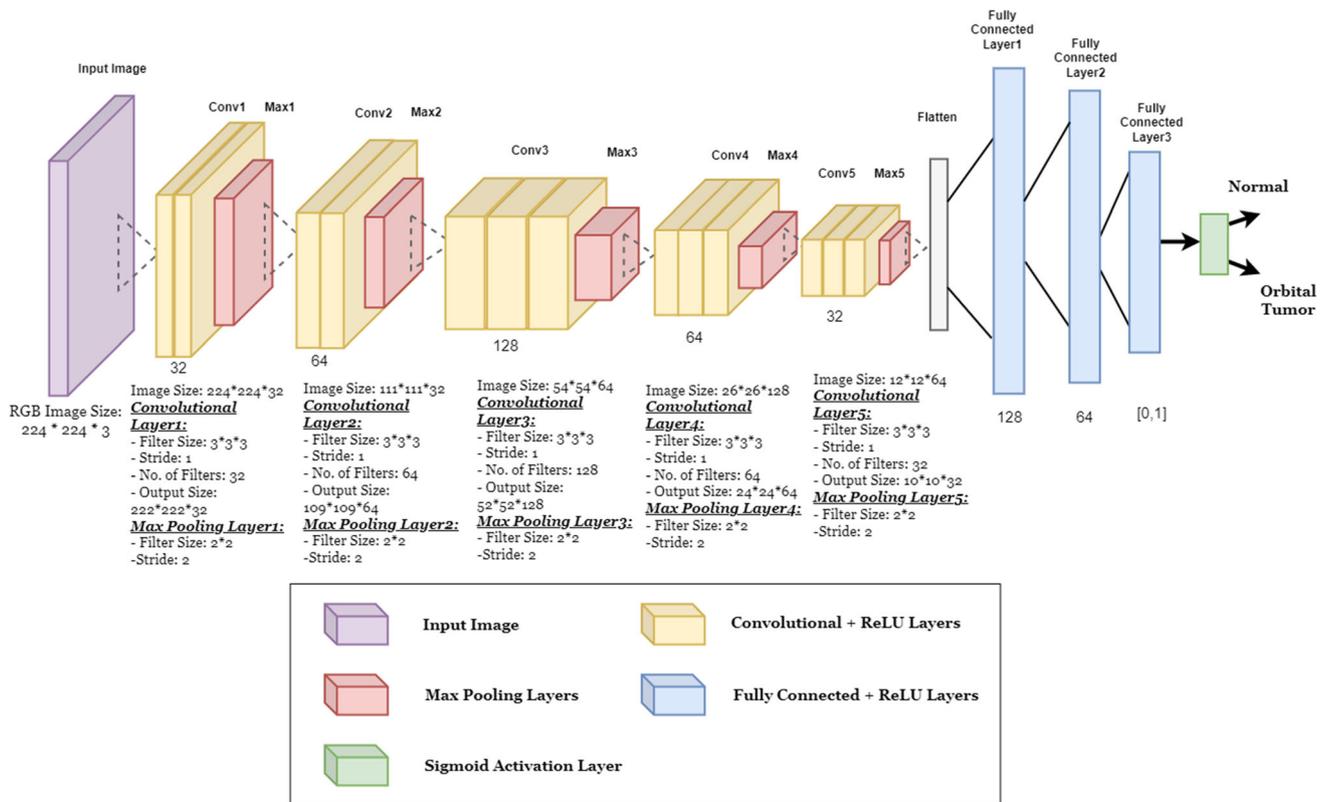


Fig. 6 The CNN Model Architecture

output of the formula $((n + 2p - f) / s) + 1 = ((224 + 2*0 - 3) / 1) + 1 = 222$. The output of this layer recognizes essential characteristics such as straight edges and corners.

Layer 2 is the subsampling layer which we use as the max pooling layer. The pooling size is 2*2, padding is 0 and stride is the default value so it is the same as pooling size 2. We get a feature map of size 32@111*111 after the max pooling operation, where 32 is the number of neurons and 111 is the output of the formula $((n + 2p - f) / s) + 1 = ((222 + 2*0 - 2) / 2) + 1 = 111$. This layer has no activation function.

Layer 3 is the convolutional layer where the ReLu activation function is applied. The convolutional filter size is 3*3, the padding is 0, the stride is 1, and the number of neurons is 64. After this convolution process, we get feature maps of size 64@109*109, where 64 is the number of neurons. The ReLu activation is applied in each feature map.

Layer 4 is the max pooling layer. The pooling size is 2*2, padding is 0, and stride is 2. We get a feature map of size 64@54*54 after the max pooling operation, where 64 is the number of neurons. This layer has no activation function.

Layer 5 is the convolutional layer where the ReLu activation function is applied. The convolutional filter size

is 3*3, the padding is 0, the stride is 1, and the number of neurons is 128. After this convolution process, we get feature maps of size 128@52*52, where 128 is the number of neurons. The ReLu activation is applied to each feature map.

Layer 6 is the max pooling layer. The pooling size is 2*2, padding is 0 and stride is 2. We get feature map size 128@26*26 after the max pooling operation, where 128 is the number of neurons. This layer has no activation function.

Layer 7 is the convolutional layer where the ReLu activation function is applied. The convolutional filter size is 3*3, the padding is 0, the stride is 1, and the number of neurons is 64. After this convolution process, we get feature maps of size 64@24*24, where 64 is the number of neurons. The ReLu activation is applied to each feature map.

Layer 8 is the max pooling layer. The pooling size is 2*2, padding is 0, and stride is 2. We get a feature map of size 64@12*12 after the max pooling operation, where 64 is the number of neurons. This layer has no activation function.

Layer 9 is the convolutional layer where the ReLu activation function is applied. The convolutional filter size is 3*3, the padding is 0, the stride is 1, and the number of neurons is 32. After this convolution process, we get

feature maps of size 32@10*10, where 32 is the number of neurons. The ReLu activation is applied to each feature map.

Layer 10 is the max pooling layer. The pooling size is 2*2, padding is 0, and stride is 2. We get a feature map of size 32@5*5 after the max pooling operation, where 32 is the number of neurons. This layer has no activation function.

Layer 11 is a flattened layer to reshape the previous layer’s output and produces a one-dimensional vector to be used as an input to a fully connected layer.

Layer 12 is a dropout layer, a mask that eliminates some neurons’ contributions to the next layer. It is one of the regularization techniques to prevent overfitting during the training phase.

Layer 13 is a fully connected layer. It receives the input from the flattened layer and outputs a one-dimensional vector of size 128. Each element in the vector receives the ReLu activation function.

Layer 14 is a fully connected layer. It receives the input from the previous layer and outputs a one-dimensional vector of size 64. Each element in the vector receives the ReLu activation function.

Layer 15 is the last layer in our CNN architecture. It is another fully connected layer. It calculates the class scores for each image to which the class belongs, resulting in a binary class [0, 1] where the normal image is represented by 0 and the tumor image is represented by 1. For the final output, the sigmoid activation function is used.

5 Results

Two MRI datasets were used in the experiment: T1-weighted and T2-weighted images. The MRI T1-weighted image has 1404 MRI images: 702 normal and 702 tumor images. The MRI T2-weighted image has 1560 MRI images: 780 normal and 780 tumor images. Each dataset was split into three groups training, validation and testing. For this purpose, 80% of all images were allocated to the training group and the remaining images were divided into the validation group (10%) and the testing group (10%). The performance of the developed system was evaluated and assessed using accuracy, recall, precision and f1-score. In order to calculate these matrices, we first calculated True Positive (TP), False Positive (FP), False Negative (FN) and True Negative (TN) for both MRI T1-weighted images and MRI T2-weighted images. Figures 7 and 8 show the confusion matrix for MRI T1-weighted images and MRI T2-weighted images, respectively.

The accuracy, recall, precision, and f1-score of our proposed system are shown in Table 4. The model training and validation accuracy and loss results across epochs for

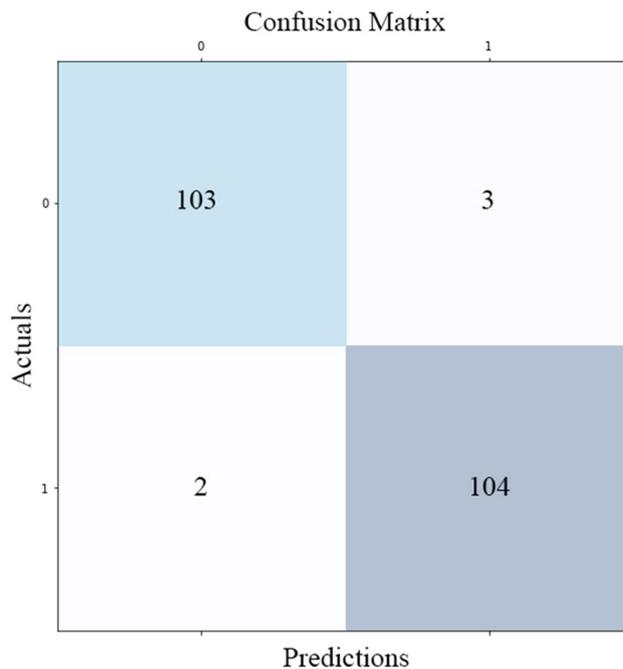


Fig. 7 Confusion Matrix for MRI T1-weighted images

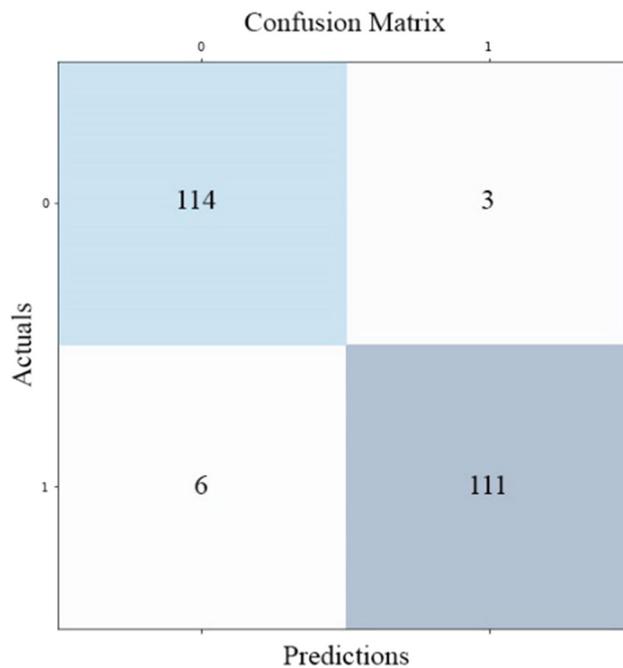


Fig. 8 Confusion Matrix for MRI T2-weighted images

Table 4 The results of our proposed CNN model

	T1-weighted images (%)	T2-weighted images (%)
Accuracy	98	97
Recall	97	94.8
Precision	99	98
f1-score	99	97

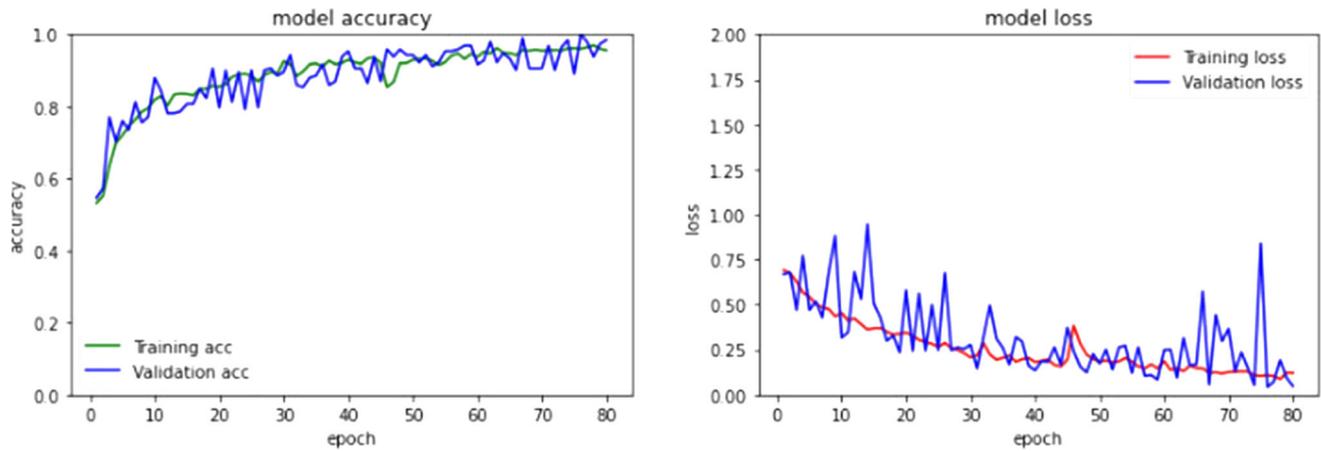


Fig. 9 The model training and validation accuracy and loss for MRI T1-weighted images

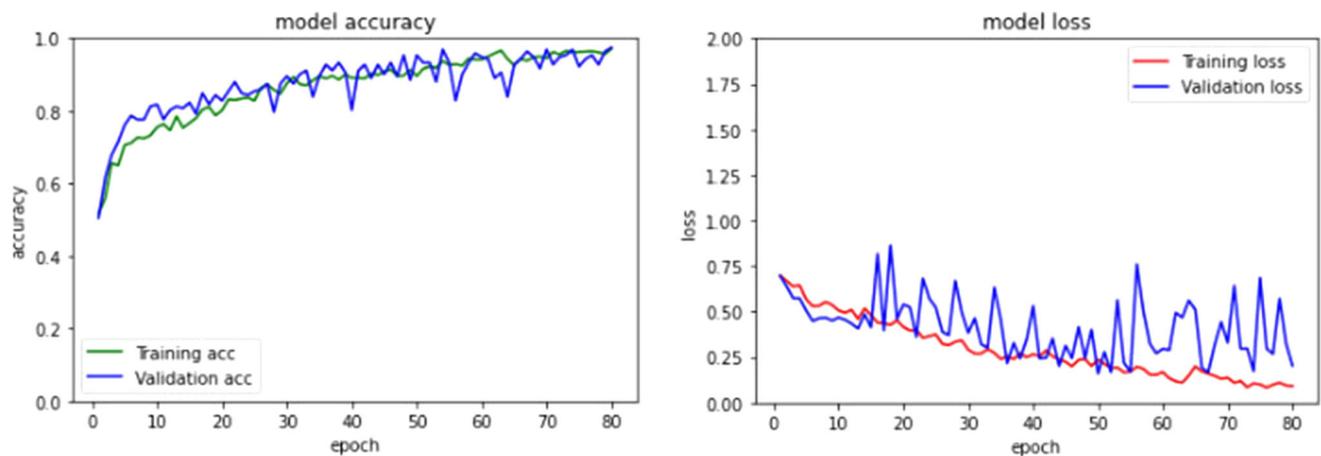


Fig. 10 The model training and validation accuracy and loss for MRI T2-weighted images

MRI T1-weighted images are demonstrated in Fig. 9, and the model training and validation accuracy and loss results across epochs for MRI T2-weighted images are demonstrated in Fig. 10.

6 Discussions

This paper proposed a system for orbital tumor identification from MRI images. It was trained on private MRI datasets for orbital tumors of 1404 MRI T1-weighted images and 1560 MRI T2-weighted images. We applied different preprocessing techniques to increase the size of the datasets. Various data augmentation techniques were applied to enhance the model accuracy in the training phase. In the experiment, horizontal flip, shifting zooming, shearing, rescaling, resizing and rotation manipulation were used as augmentation techniques. This proposed system used a CNN with five convolutional and pooling

layers. In addition, it contains three fully connected layers and the final one was modified to have one neuron.

Our classification system's outcome was evaluated and analyzed using different evaluation metrics, such as f1-score, recall, precision and accuracy. The benefit of using several different evaluation metrics is to ensure our model's performance. Furthermore, while separating the data, two approaches were utilized: shuffling and randomly splitting the images into training, validation, and testing sets, which is called the "image-level approach". Also, the "patient-level method" implies that each patient image was utilized in either the training or validation set, but not both. To ensure that the splitting data were representative of the entire distribution of the data, we applied a shuffling approach when splitting the data into training/validation/test sets. In addition, shuffling data reduced variance and ensured that the proposed model remained generic and overfitting-free. As an additional validation step, the second approach was utilized to ensure that our proposed

Table 5 Performance comparison with previous work

Publication	Objective	Methodology	Dataset	Performance
Biswarup, et al. [8]	Classification of eye melanoma	CNN	The dataset contains 170 images; affected and unaffected	A recognition rate is 91.76%
Sheshang, et al. [9]	Classification of eye melanoma	CNN	The dataset contains 200 images	Accuracy is 92.5%
Parmod et al. [12]	Identification of Retinoblastoma from the fundus images	AlexNet ResNet50	The dataset contains 278 fundus images	Accuracy is 88.12% for AlexNet and 93.16% for ResNet50 algorithms
The proposed system	Identification of orbital tumors from MRI images	CNN	The datasets contain 1404 MRI T1-weighted and 1560 MRI T2-weighted images	Accuracy is 98% for MRI T1-weighted and 97% for MRI T2-weighted images

system could categorize fresh patient images which had not been exposed during the training phase.

To the best of our knowledge, this work is novel as no previous research has been conducted to classify orbital tumors; this is the first research on this type of tumor. But there are previous studies that classified different types of eye tumors. The results are compared with the existing technique in terms of performance indices (see Table 5).

CNN architecture achieved promising accuracy in categorizing the MRI image dataset. Owing to its high resolution for diverse tissues and lack of radiation, MRI is the most often utilized non-invasive imaging technology for tumor detection. On the other hand, tumor categorization utilizing MRI scans is a difficult endeavor due to overlapping intensities and unpredictability in orientation, size and shape. A neural network with sufficient depth, such as CNN, is appropriate for controlling the variance and learning the high-level characteristics, in addition to handling noise disturbances and low image contrast [29]. It has proven to be more capable of handling the variety and complexity of MRI medical images. Moreover, medical databases are typically limited and difficult to obtain. Our system's use of data augmentation enhanced model accuracy by expanding the variety of accessible data without the requirement for further data collection.

7 Conclusion and future work

Our study demonstrates a new classification system for an orbital tumor using MRI images. Preprocessing techniques were applied to increase the size of the dataset. These techniques include median filter, Gaussian blur, brightness, contrast and edge enhance filters. Data augmentation techniques were applied to enhance the performance of the system, which includes horizontal flip, shifting, zooming, shearing, rescaling, resizing and rotation manipulation. The system was trained with two private datasets of MRI

images 1404 T1-weighted and 1560 T2-weighted images. The recognition rate for the T1-weighted image is 98% and for the T2-weighted image is 97%. For future work, we aim to apply this system to different image types, such as computed tomography, ultrasound and histological images in addition to other types of eye tumors, such as iris, conjunctiva, uvea and secondary tumors.

Acknowledgements The authors would like to thank Dr. Azza Mohamed Ahmed, Professor of Ophthalmology, Faculty of Medicine, Ain Shams University for her contribution to the collection of the dataset and validation of the results.

Funding Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

Data availability The data that support the findings of this study are provided by the Ophthalmology Department, Faculty of Medicine, Ain Shams University, Cairo, Egypt.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The authors declare the following financial interests/personal relationships which may be considered as potential competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Kannan S, Hasegawa M, Yamada Y, Kawase T, Kato Y (2019) Tumors of the orbit: case report and review of surgical corridors and current options. *Asian J Neurosurg* 14(03):678–685. https://doi.org/10.4103/ajns.AJNS_51_19
- Blandford AD, Perry JD (2019) Classification of orbital tumors. In: *Clinical ophthalmic oncology: orbital tumors*. pp 9–15. https://doi.org/10.1007/978-3-030-13558-4_2
- Zhang L, Li X, Tang F, Gan L, Wei X (2020) Diagnostic imaging methods and comparative analysis of orbital cavernous hemangioma. *Front Oncol* 10:577452. <https://doi.org/10.3389/fonc.2020.577452>
- Karcioglu ZA (2019) Overview and imaging of orbital tumors. In: *Surgical ophthalmic oncology: a collaborative open access reference*, pp 107–116
- Mirbabaie M, Stieglitz S, Frick NR (2021) Artificial intelligence in disease diagnostics: a critical review and classification on the current state of research guiding future direction. *Health Technol* 11(4):693–731. <https://doi.org/10.1007/s12553-021-00555-5>
- Yan Y, Yao XJ, Wang SH, Zhang YD (2021) A survey of computer-aided tumor diagnosis based on convolutional neural network. *Biology* 10(11):1084. <https://doi.org/10.3390/biology10111084>
- Scheetz J, He M, van Wijngaarden P (2021) Ophthalmology and the emergence of artificial intelligence. *Med J Aust* 214(4):155–157. <https://doi.org/10.5694/mja2.50932>
- Ganguly B, Biswas S, Ghosh S, Maiti S, Bodhak S (2019) A deep learning framework for eye melanoma detection employing convolutional neural network. In: 2019 International conference on computer, electrical & communication engineering (ICCECE), pp 1–4. IEEE. <https://doi.org/10.1109/ICCECE44727.2019.9001858>
- Degadwala S, Vyas D, Dave HS, Patel V, Mehta JN (2022) Eye melanoma cancer detection and classification using CNN. In: *Second international conference on image processing and capsule networks: ICIPCN 2021 2* (pp 489–497). Springer. https://doi.org/10.1007/978-3-030-84760-9_42
- Miles Research, available: <http://www.milesresearch.com/main/links.htm>, 2023. Accessed 30 June 2023
- The Eye Cancer Foundation. <http://www.eyecancer.com/research/image-gallery/>, 2023. Accessed 30 June 2023
- Kumar P, Suganthi D, Valarmathi K, Swain MP, Vashistha P, Buddhi D, Sey E (2023) A multi-thresholding-based discriminative neural classifier for detection of retinoblastoma using CNN models. *BioMed Res Int* 2023. <https://doi.org/10.1155/2023/5803661>
- Jeba J (2023) Retinoblastoma dataset, MATLAB Central File Exchange. (<https://www.mathworks.com/matlabcentral/fileexchange/99559-retinoblastoma-dataset>)
- Tsuneki M (2022) Deep learning models in medical image analysis. *J Oral Biosci* 64(3):312–320. <https://doi.org/10.1016/j.job.2022.03.003>
- Yamashita R, Nishio M, Do RKG, Togashi K (2018) Convolutional neural networks: an overview and application in radiology. *Insights Imaging* 9(4):611–629. <https://doi.org/10.1007/s13244-018-0639-9>
- Suganyadevi S, Seethalakshmi V, Balasamy K (2022) A review on deep learning in medical image analysis. *Int J Multimedia Inf Retrieval* 11(1):19–38. <https://doi.org/10.1007/s13735-021-00218-1>
- Chen L, Li S, Bai Q, Yang J, Jiang S, Miao Y (2021) Review of image classification algorithms based on convolutional neural networks. *Remote Sensing* 13(22):4712. <https://doi.org/10.3390/rs13224712>
- Radiopaedia Image Bank (2022). <https://radiopaedia.org/>. Accessed 20 Oct 2022
- Tan L, Jiang J (2019) Chapter 13: Image processing basics, digital signal processing (third edition), Academic Press, pp 649–726. <https://doi.org/10.1016/B978-0-12-815071-9.00013-0>
- Ngo D, Lee S, Kang B (2020) Nonlinear unsharp masking algorithm. In: 2020 International conference on electronics, information, and communication (ICEIC), pp 1–6. IEEE. <https://doi.org/10.1109/ICEIC49074.2020.9051376>
- Suetens P (2017) *Fundamentals of medical imaging*. Cambridge University Press
- Mohammed FG, Rada HM, Mohammed SG (2013) Contrast and brightness enhancement for low medical X-ray images. *Int J Sci Eng Res* 4(5):1519
- Perumal S, Velmurugan T (2018) Preprocessing by contrast enhancement techniques for medical images. *Int J Pure Appl Math* 118(18):3681–3688
- Rezaei M, Yang H, Meinel C (2017) Deep neural network with l2-norm unit for brain lesions detection. In: *Neural information processing: 24th international conference, ICONIP 2017, Guangzhou, China, November 14–18, 2017, Proceedings, Part IV 24*, (pp. 10637:798–807). Springer. https://doi.org/10.1007/978-3-319-70093-9_85
- Gao L, Zhang L, Liu C, Wu S (2020) Handling imbalanced medical image data: a deep-learning-based one-class classification approach. *Artificial intelligence in medicine*, pp 108:101935. <https://doi.org/10.1016/j.artmed.2020.101935>
- Ibrahim M, Torki M, El-Makky N (2018) Imbalanced toxic comments classification using data augmentation and deep learning. In: 2018 17th IEEE international conference on machine learning and applications (ICMLA), pp 875–878. IEEE. <https://doi.org/10.1109/ICMLA.2018.00141>
- Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. *J Big Data* 6(1):1–48. <https://doi.org/10.1186/s40537-019-0197-0>
- Ismael SAA, Mohammed A, Hefny H (2020) An enhanced deep learning approach for brain cancer MRI images classification using residual networks. *Artif Intell Med* 102:101779. <https://doi.org/10.1016/j.artmed.2019.101779>
- Tandel GS, Biswas M, Kakde OG, Tiwari A, Suri HS, Turk M, Suri JS (2019) A review on a deep learning perspective in brain cancer classification. *Cancers* 11(1):111. <https://doi.org/10.3390/cancers11010111>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.