Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 19 No. 9 (2023) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v19i09.38819

PAPER

Brain Tumor Classification Deep Learning Model Using Neural Networks

Gisella Luisa Elena Maquen-Niño[1](✉), Ariana Ayelen Sandoval-Juarez[2], Robinson Andres Veliz-La Rosa[2], Gilberto Carrión-Barco[1], Ivan Adrianzén-Olano[3], Hugo Vega-Huerta[2], Percy De-La-Cruz-VdV[2]

[1]Universidad Nacional Pedro Ruíz Gallo, Lambayeque, Perú

[2]Universidad Nacional Mayor de San Marcos, Lima, Perú

[3]Universidad Nacional Toribio Rodríguez de Mendoza, Chachapoyas, Perú

gmaquenn@unprg.edu.pe

ABSTRACT

The timely diagnosis of brain tumors is currently a complicated task. The objective was to build an image classification model to detect the existence or not of brain tumors by adding a classification header to a ResNet-50 architecture. The CRISP-DM methodology was used for data mining. A dataset of 3847 brain MRI images was used, 2770 images for training, 500 for validation, and 577 for testing. The images were resized to a 256 × 256 scale and then a data generator is created that is responsible for dividing pixels by 255. The training was performed and then the evaluation process was carried out, obtaining an accuracy percentage of 92% and a precision of 94% in the evaluation process. It is concluded that the proposed CNN model composed of a head with a ResNet50 architecture and a seven-layer convolutional network achieves adequate accuracy, becoming an efficient and complementary proposal to other models developed in previous works.

KEYWORDS

deep learning, brain tumors, convolutional neural networks, classification models

1 INTRODUCTION

Brain tumors constitute a heterogeneous group due to the different cells that originate them. In this sense, they can be divided into two groups: primary lesions and secondary lesions; the primary ones have a cellular origin belonging to the central nervous system, while the secondary ones originate in other areas of the body and can be implanted as metastases in the brain [1].

Worldwide, cardiovascular and cerebrovascular diseases occupy the first places in morbidity and mortality. These diseases develop from childhood and can occur suddenly in adulthood, negatively impacting society as it constitutes an important health problem that generates a greater burden than infectious diseases [2].

In Peru, it was determined that the mortality of brain cancer and other parts of the nervous system was 2.9% in 2016 and within the population cancer registry

Maquen-Niño, G.L.E., Sandoval-Juarez, A.A., Veliz-La Rosa, R.A., Carrión-Barco, G., Adrianzén-Olano, I., Vega-Huerta, H., De-La-Cruz-VdV, P. (2023). Brain Tumor Classification Deep Learning Model Using Neural Networks. *International Journal of Online and Biomedical Engineering (iJOE)*, 19(9), pp. 81–92. https://doi.org/10.3991/ijoe.v19i09.38819

Article submitted 2023-02-13. Resubmitted 2023-04-25. Final acceptance 2023-04-25. Final version published as submitted by the authors.

of Lima Metropolitan, the incident cases of brain tumors were 3.4% with a total of 2060 cases during the period 2010–2012 [3].

Early diagnosis of this disease is crucial, which is why different methods have been used, one of which is magnetic resonance imaging defined as a medical imaging technique that provides details of organs and tissues of the body thanks to the use of a magnetic field and radio waves generated by a computer. This technique is used for accurate diagnosis of the type of disease and observation of the damage caused by it. However, there are reasons such as lack of experience or theoretical bases, which cannot detect and/or locate anomalies that can lead to a wrong diagnosis and, as a result, a wrong treatment towards the patient.

To respond to these challenges, within the practice of medicine, many professionals have emphasized building accurate prediction models [4] capable of identifying patterns and classifying them to support the medical specialist with the diagnosis and treatment of the patient. This usually leads to advancing toward the idea of Deep Learning which is defined as a branch of Machine learning (Automatic learning) that carries out the attempt to model abstractions of high-level data using multiple layers of neurons (neural networks or NN) that constitute complex structures or non-linear transformations [5]. Recently, it has been argued that deep neural networks are convenient when extracting useful data and providing performance within predictions from training large and multimodal volumes of data [6].

These networks, in general, have high classification capacity that can surpass traditional techniques such as Logistic Regression, Random Forests, etc. [6].

In this sense, the present study proposes using deep learning techniques to classify brain tumors using Magnetic Resonance Imaging (MRI). The study aims to provide an accurate diagnosis and treatment of brain tumors in patients with the use of deep learning techniques.

2 RELATED WORK

Research related to detection-oriented machine learning models has been carried out, such as that of Cubas and Niño, who compiled the results of previous studies on detection algorithms for phishing websites and they created 6 different classification algorithms: K Neighbor-Nearest (KNN) which got an accuracy of 94.13, Decision Tree (DT) 96.52, Random Forest (RF) 97.02, Extra Tree (ET) 97.04, Ada Boosting (AB) 94.50 and Support Vector Machine (SVM) 94.50, choosing the Random Forest, Extra Tree, and Decision Tree algorithms as the most efficient and after a final optimization the accuracy of 97.56, 97.33 and 97.29 were obtained respectively, finally selecting the RF Model as the best by the highest score; Likewise, previous studies of 22 investigations were analyzed, among which neural networks are included, the one with the highest score being a deep learning neural network with an accuracy of 96%, being the difference in 1.56% with the algorithm Random Forest of the authors [7].

Likewise, there are studies oriented to disease detector classification models using convolutional neural networks that are within deep learning, one of them seeks the classification of melanoma as an indicator of skin cancer, using a methodology with the assembly of the EfficientNET and VGG-19 architectures, obtaining an accuracy of 92.85% with an improvement of 0.06% in accuracy and specificity. One of the conclusions of this research is the work with data augmentation of the images to generate a larger size in the database and obtain improvement in performance, as well as the assembly of the models that helps to improve accuracy [8].

There is research in which convolutional neural networks are used to solve a problem, as in the research of Alameri et al. where the objective was to reduce or eliminate noise in speech signals and increase intelligibility and quality. The improvement of the speech signal is performed by applying two machine learning models based on a convolutional neural network. on a convolutional neural network (CNN) and a fully connected neural network (FCN). It is concluded that, in general, the CNN is better than the FCN in terms of coloring than the FCN. FCN in terms of maximum coloration, maximum signal-to-noise ratio (PSNR), short-term objective intelligibility (STOI), and short-term objective error (STOI). intelligibility (STOI) and short-term objective intelligibility (STOI) and mean square error (MSE). mean square error (MSE) [9].

Following the research line, the use of Convolutional Neural Networks (CNN) models has been given for the classification of medical images. Sarwinda et al. used this network for the detection of colorectal cancer by training classifiers ResNet-18 and ResNet-50 on images of colon glands to distinguish between benign and malignant colorectal cancer [10]. Similar work was carried out by Mogotocoro who focused on classifying three types of brain tumors (meningioma, glioma, and pituitary) from magnetic resonance images using mathematical techniques and neural networks and then implementing an accessible graphical interface system; for this, he developed two CNN models, the first consisting of a traditional neural network while the other had data augmentation [11].

Wadood Abdul presents an automatic system for detecting and classifying lung cancer using deep learning (DL); for this, he used the convolutional neural network (CNN) model and he builded an ALCDC system to detect and classify whether the tumors found in the lungs are malignant or benign; the robustness and effectiveness of the proposed ALCDC system are validated with images from the Lung Image Database Consortium (LIDC) and the Image Database Resource Initiative (IDRI); concluding that the proposed ALCDC will be useful in medical diagnosis and health systems [12].

Balasooriya and Nawarathna used magnetic resonance imaging scans of the axial, coronal, and sagittal planes of the brain to train and test a CNN model to facilitate the classification and prediction of brain tumors in 5 different classes (astrocytoma, glioblastoma multiforme, oligodendroglioma, healthy tissue, and unidentifiable tumor); the proposed model was developed and trained using TensorFlow, TFLearn, Scikit-learn and other Python libraries used in machine learning [13].

Wadhah Ayadi et al. suggest a new neural network model for the classification of brain tumor MRI; their model takes advantage of multiple layers and a classifier called Softmax; the experimental study of the proposed model was carried out using a public dataset; the dataset used includes 3064 images and focused on the evaluation of accuracy; their technique requires minimal preprocessing and the experimental results proved the effectiveness of the model despite the small amount of training data [14].

Regarding the work with images, Al-Hazaimeh et al. indicate that the processing and analysis of images based on the human form can be carried out in a better way through a robust geometric model, which in their research was used to classify a fundus image as human or non-human, comparing the results of the proposed model with other models such as artificial neural networks, vector support machine and random forest, resulting in the proposed model being efficient and comparable to any of the other state-of-the-art models [15].

Al-hazaimeh et al. conducted an investigation on the application of image processing and machine learning for the detection of diabetic retinopathy and its classification into exudates, microaneurysms and retinal hemorrhages using fundus images through MATLAB software. Image processing was performed, applying HSI conversion (H hue, S saturation and I intensity), DE noise and other preprocessing steps that take place during this stage. Feature selection was also applied using a deep convolutional neural network (DCNN) architecture consisting of four convolutional layers and two fully connected layers, the third convolutional layer is fed directly from the previous convolutional layer after applying contrast normalization, clustering, and a nonlinear function. Once converted to a feature vector, the output of the convolutional layer is passed to the fully connected layers for classification. Segmentation and classification have been used in the proposed diabetic retinopathy detection and retinal disease categorization. For segmentation, the BCSPNFCM algorithm: Bias Corrected Separate Possibilistic Neighborhood FCM algorithm was used. The SVMGA algorithm was used for classification. In the evaluation, it was verified that the accuracy of the model with proposed machine learning methods was 98.80 percent [16].

3 METHODOLOGY

For this research, the CRISP-DM (Cross Industry Standard Process for Data Mining) was used, which is defined as a data mining process model that describes a way in which experts in this field approach the problem [17] in addition to being a process for developing iterative, open, adaptable and highly recognized data mining projects by the industry and academia [18]. This methodology includes a structured guide in six phases and can be reviewed between phases, so it is not necessary to follow a strict order to perform it. A dataset published on the Kaggle data science community web platform (https://www.kaggle.com/datasets/mateuszbuda/lgg-mri-segmentation) was used as input data, this has a total of 3847 magnetic resonance images of the brain along with manual anomaly segmentation masks from 110 patients, many of these images have dimensions of 256 × 256, of which 72% will be used for training, 13% for validating and 15% for testing.

3.1 Understanding the business

Regarding diseases caused by brain tumors, due to the high incidence of morbidity and the high costs of special and multidisciplinary treatment of their patients, it is necessary to have the necessary diagnostic and therapeutic tools to provide the best options for the patient and their family; in this way, it contributes to the patient being able to live the longest possible time with the best quality of life [19].

As part of the diagnostic tools, the detection of brain tumors is determined as a goal through the development of a deep learning-based model that can make a relatively reliable estimate based on the logic of the model and its accuracy. Therefore, the objective is to: develop a model that detects brain tumors from magnetic resonance imaging and evaluate the model through performance metrics obtained after its training.

Following this line, the construction of a said model with performance values that exceed 70% is established as a success criterion. To do this, a platform called

Google Colab is available, which provides tools for performing data mining tasks on a set of data stored within a Google Drive repository.

3.2 Understanding of data

The second phase of the CRISP-DM approach involved collecting data to establish an initial relationship with the problem, becoming familiar with the data and determining its quality, and identifying the most obvious relationships to formulate the first hypothesis.

First, a dataset published on the Kaggle data science community web platform was used as input data, which can be downloaded at the following link: https://www.kaggle.com/datasets/mateuszbuda/lgg-mri-segmentation, used by Buda et al. and Mazurowski et al. [20], [21]; This has a total of 3847 magnetic resonance images of the brain along with manual anomaly segmentation masks from 110 patients, many of these images have dimensions of 256 × 256 and a width and height of 256 pixels. A documentation file of each of the dataset elements that can be used when doing data mining is also included.

Secondly, the data that is stored in a Google Drive repository and organized by folders was described. Also, access to the names and direct links to the images are together; the first corresponds to the unique patient identifier, the second corresponds to the MRI link of the patient, the third corresponds to the image of the mask corresponding to the patient and the last can have values 0 and 1 depending on the presence of an anomaly in the MRI mask.

Subsequently, the data were explored, finding 2474 MRIs without anomalies in their masks and 1373 images that did have them, the total number of images was resized on a scale of 256 × 256 with the purpose that they all have the same size and reduce training run time.

3.3 Modeling

Since the Google Colab tool will be used for data mining, it is essential to use some of the modeling techniques that can be implemented in said platform according to the objectives of the project.

Thanks to the analysis of the current literature and the problems, it is concluded that to reach the proposed objectives, a model must be developed that will be specifically in charge of ruling out a tumor in the magnetic resonance. Since the challenge is to solve a detection problem and therefore the response to be predicted is categorical.

For the construction of the model, image processing, training, optimization, storage, and evaluation of the resulting model are performed.

3847 brain MRI images were used, of which 2770 images are used for the training, 500 images are used for the validating and 577 for the testing process; Figure 1 shows 4 images of the total dataframe with the labels MRI Brain, Mask-0, and Mask-1. Within the data partitioning stage, the train and test dataframes are created. From the training data, a data generator called datagen is created that is responsible for dividing pixels between 255 and a validation set to achieve that the model manages to generalize the images.

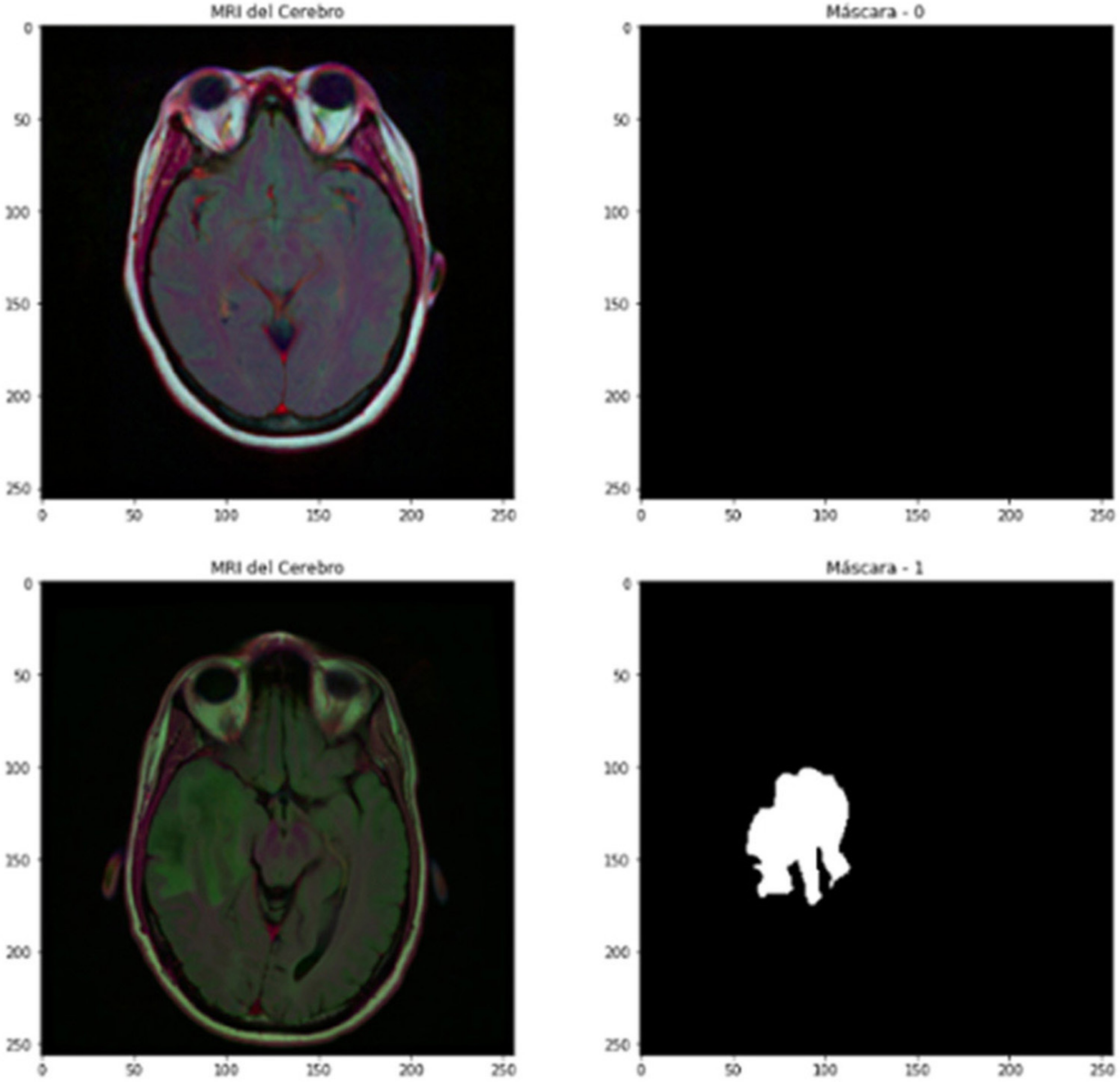


Fig. 1. Magnetic resonance imaging and masks

For this, the normalization of the images is carried out by changing the data of the pixels in a range from 0 to 1 necessary to improve the training parameters of the neural network. Subsequently, two generators are created to take the train and test dataframe created, for the first one two subsets are built using the datagen generator: train_generator and valid_generator, and for the last one the subset test_generator (previous creation of the data generator for test images: test_datagen). In addition, the batch size is set to 16, which means that the model will be run in batches of 16 images and is set to categorical mode (class_mode) since the output consists of two values. With these generators, the training, validation, and test images are loaded and processed in two classes that represent the value of the presence or absence of tumor.

Applying transfer learning, an already trained base model called ResNet50 is downloaded, which will carry the training weights used on the dataset without including the final dense layer that is in charge of prediction. A classification header is added to the model which consists in adding a MaxPooling2d type pooling layer that will reduce to a quarter the dimensions of the images that are flattened to obtain neurons one after the other. Then a first dense layer with a unitary linear rectifier is added to aggregate 256 neurons and the dropout technique is used which disables 30% of the neuron connections to avoid possible overfitting and improve the generalization capacity. Subsequently, a second dense and dropout layer with the same characteristics as the previous ones are added, and at the end of the convolutional neural network, 2 output neurons with softmax activation function were used (Figure 2). Finally, the input is taken from the base model ResNet-50 base, and from this the output will be set by the header, thus creating a new CNN model.

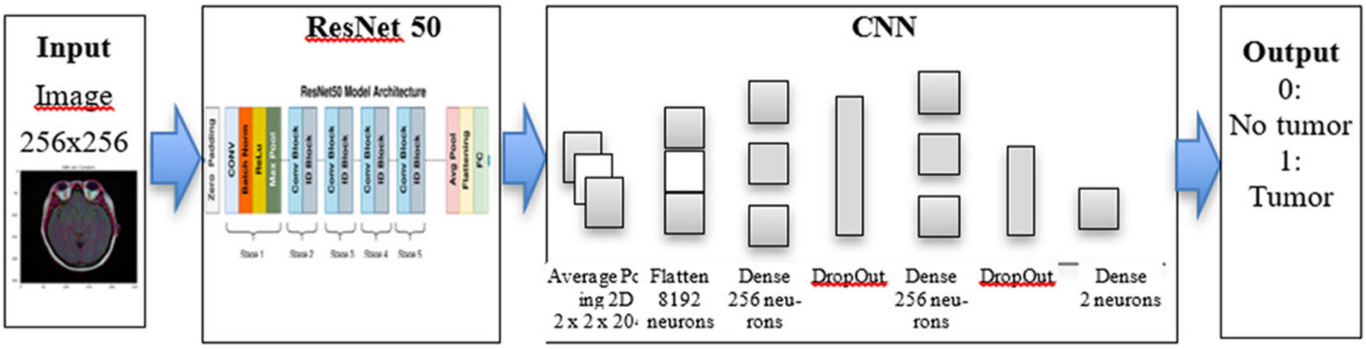
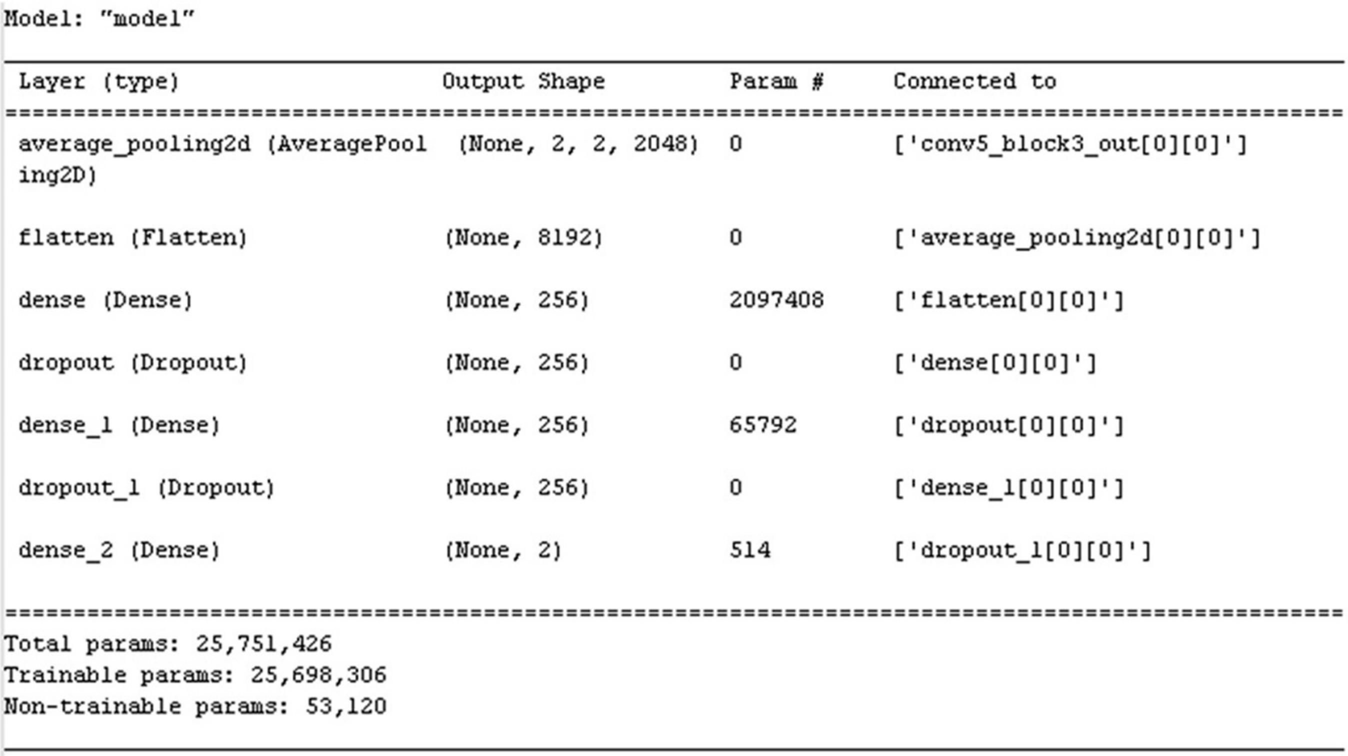


Fig. 2. Architecture of the convolutional neural network


```
Model: "model"
_______________________________________________________________________________________________
 Layer (type)                   Output Shape         Param #     Connected to
===============================================================================================
 average_pooling2d (AveragePool  (None, 2, 2, 2048)  0           ['conv5_block3_out[0][0]']
 ing2D)

 flatten (Flatten)              (None, 8192)         0           ['average_pooling2d[0][0]']

 dense (Dense)                  (None, 256)          2097408     ['flatten[0][0]']

 dropout (Dropout)              (None, 256)          0           ['dense[0][0]']

 dense_1 (Dense)                (None, 256)          65792       ['dropout[0][0]']

 dropout_1 (Dropout)            (None, 256)          0           ['dense_1[0][0]']

 dense_2 (Dense)                (None, 2)            514         ['dropout_1[0][0]']

===============================================================================================
Total params: 25,751,426
Trainable params: 25,698,306
Non-trainable params: 53,120
_______________________________________________________________________________________________
```

Fig. 3. Layers of the convolutional neural network

In the compilation, first, the loss function is configured, the use of the categorical crossover, the adaptive estimation by sectors (Adam) and it is configured with accuracy metrics (accuracy). Early stopping is also used to monitor errors in the validation phase and when the minimum takes more than 20 iterations without improvement, the execution is cut (patience parameter) to then save the model with the best validation loss. The history is created to adjust the number of steps given by the total number of images divided by 16, the set of validations, and the execution of the early stop and the checkpointer. When loading the history, the training architecture where the weights are placed is saved. The model is loaded and compiled to then carry out the prediction of the test generator obtaining the predicted class from the model; the number of images to be predicted is limited and the model hit rate is obtained.

3.4 Evaluation

Since the source dataset is unbalanced, a confusion matrix is prepared (Figure 4) where it can be seen that, of the 577 images to be predicted, 14 of them correspond to false positives (FP) and 23 to false negatives (FN), while 170 to True Positives (TP) and 370 to True Negatives (TN).

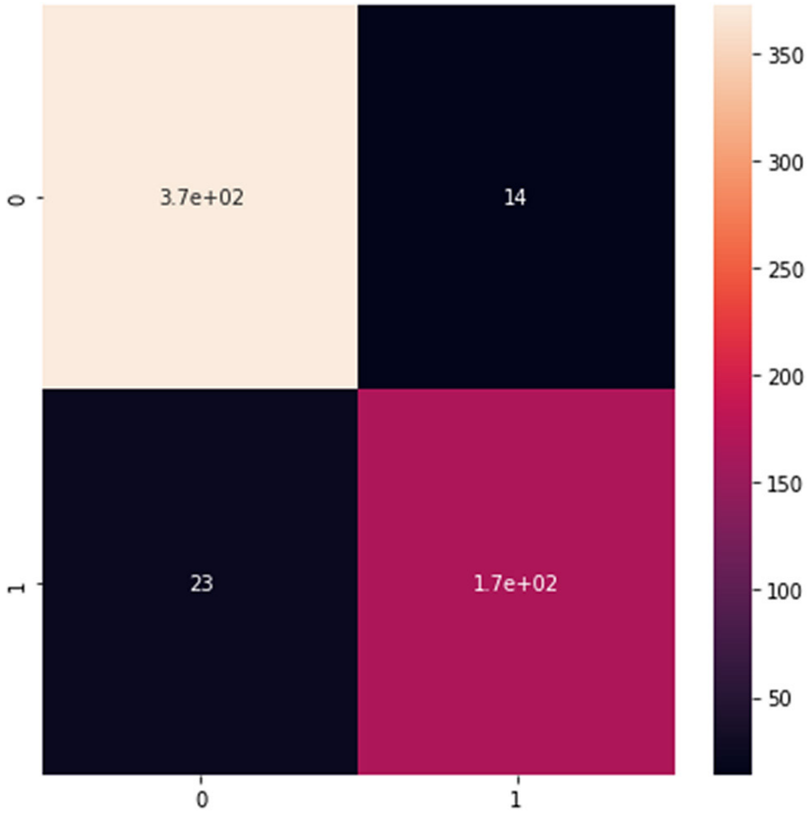


Fig. 4. Confusion matrix of the model in the detection of brain tumor

From a commercial perspective, key success criteria have been identified to develop a model based on the detection of brain tumors based on magnetic resonance imaging. In accordance with that objective, the calculations applied for the experiment results consist of accuracy and precision. The theorems for each of them imply the following [22]:

$$Precisión = \frac{TP}{TP + FP} \qquad Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Recall = \frac{TP}{TP + FN} \qquad F1 - Score = 2 * \frac{Recall * Precisión}{Recall + Precisión}$$

The brain tumor classification model developed in the present investigation yields 4 results (Table 1), from which a balance can be shown concerning previous works.

Table 1. Performance of the proposed model

Accuracy	Precisión	F1-Score	Recall
0.94	0.92	0.90	0.88

4 DISCUSSION AND RESULTS

When comparing the results obtained (Table 2) with the study of Sarwinda et al. [10], it is determined that there is a difference of 0.04 in favor of the proposed model concerning accuracy. likewise, it is known that 80% and 20% of the data were destined for training and testing, respectively; while in the proposed model 85% was destined for training and 15% for testing obtaining an accuracy of 0.92 mentioned above.

In the case of the study of the training results of the convolutional neural network of Mogotocoro [11], it was found that the training accuracy for the CNN model and CNN model with data augmentation was 0.76 and 0.85, respectively, applying 10 learning epochs with 5 validation steps. On the other hand, the difference between the two previous models and the proposed model lies in the learning epochs which, in the case of the latter, was greater (20 epochs) obtaining a training accuracy of 0.98 and a test accuracy of 0.94.

In the study of Wadood Abdul [12] the creation of the proposed model is carried out using seven thousand nodes (3500 of each, malignant and benign) as training to evaluate the proposed architecture and 1296 nodes (829 benign and 467 malignant) for testing; the model has been tested and trained with a learning rate of 0. 01, a kernel size of five and a kernel size of two and a clustering layer of 30 to 30 batches for 30 epochs; the proposed CNN architecture obtained an accuracy of 97.2%.

In the study OF Balasooriya and Nawarathna [13], the accuracy of the trained model was higher than that of the proposed one (99.46%); the model layers were programmed using TFLearn and once built Cross Validation was used to train and test the models by dividing the complete dataset into 6 folds; each data partition was trained in 10 epochs with a total of 60 epochs for the complete dataset obtaining a sophisticated configuration within the compilation.

While Ayadi et al. [14] obtained an accuracy percentage of 95.23% for Meningioma, 95.43% for Glioma, and 98.43% for pituitary tumors by training a total of 3060 images, of which 2093 images was classified correctly and 161 incorrectly, it provides 94.74% as the whole accuracy.

In this sense, under the above-mentioned, it can be asserted that input parameters such as the data rate for training, testing, number of learning epochs, and validation techniques can influence the search for greater precision and accuracy of training so that the compilation of the model using different values for the input parameters of the neural network should be considered.

Table 2. Performances of the related works and the proposed model

Investigation	Accuracy Training	Accuracy Testing	Epoch	Output
Proposed	0.98	0.94	20	0: Tumor, 1: Not_Tumor
Sarwinda et al. [10]	0.92	0.90	–	0: Tumor, 1: Not_Tumor
Mogotocoro [11]	CNN: 0.756	CNN: 0.719	10	0: Glioma 1: Meningioma 2: Pituitary
	CNN datos aumentados: 0.85	CNN datos aumentados: 0.885		
Wadood Abdul [12]	0.98	0.972	30	0: Tumor, 1: Not_Tumor
Balasooriya and Nawarathna [13]	–	0.995	10	0: Astrocytoma 1: Glioblastoma 2: Multiforme Oligodendroglioma 3: Healthy tissue 4: Unidentified tumor
Ayadi et al. [14]	–	94.74%	20	0: Meningioma 1: Glioma 2: Pituitary

5 CONCLUSIONS

Using the CRISP – DM methodology, it can be concluded that it provides a diversity of activities in all phases of the project, although due to the versatility it offers, it does not assign specific roles to the mentioned activities. A new CNN architecture was built by adding to a head-end with ResNet-50 architecture, a 7-layer network architecture: a pooling layer, a flatten layer, three dense layers and two dropout layers, for the detection of brain tumors from magnetic resonance images by training a dataset of variable size. At the end of the development, it was identified that, in the diagnostic performance of the model, the precision achieved for the precision was 0.92, the accuracy of 0.94, the F1-Score of 0.90 and the Recall or sensitivity of 0.88; considering that these values can be improved with the use of different parameters within the construction of the neural network.

6 REFERENCES

[1] L. E. Contreras, “Epidemiología De Tumores Cerebrales,” *Rev. Médica Clínica Las Condes*, vol. 28, no. 3, pp. 332–338, 2017, https://doi.org/10.1016/j.rmclc.2017.05.001

[2] J. Vigo-Ramos, “Muerte súbita y emergencias cardiovasculares: Problemática actual,” *Rev Peru Med Exp Salud Publica*, vol. 25, no. 2, pp. 233–269, 2008, [Online]. Available: www.cpr.com.pe

[3] Ministerio de Salud, “Prevencion y Control Del Cancer,” Minist. Salud, pp. 1–303, 2021, [Online]. Available: https://www.minsa.gob.pe/presupuestales/doc2021/ANEXO2_6.pdf

[4] J. Trujillano, J. March, and A. Sorribas, "Aproximación metodológica al uso de redes neuronales artificiales para la predicción de resultados en medicina," *Med. Clin. (Barc).*, vol. 122, no. SUPPL. 1, pp. 59–67, 2004, https://doi.org/10.1157/13057536

[5] X. Hao, G. Zhang, and S. Ma, "Deep Learning," *Int. J. Semant. Comput.*, vol. 10, no. 3, pp. 417–439, 2016, https://doi.org/10.1142/S1793351X16500045

[6] F. Wu, M. Li, L. Kurgan, and L. Rueda, "Neurocomputing guest editorial : Deep neural networks for precision medicine," *Neurocomputing*, vol. 469, pp. 330–331, 2022, https://doi.org/10.1016/j.neucom.2021.06.095

[7] J. V. Cubas and G. L. E. Maquen-Niño, "Machine learning model in the detection of phishing websites," *RISTI – Rev. Iber. Sist. e Tecnol. Inf.*, vol. 2022, no. E52, pp. 161–173, 2022.

[8] H. Vega-Huerta et al., "Convolutional Neural Networks on Assembling Classification Models to Detect Melanoma Skin Cancer," *Int. J. online Biomed. Eng.*, vol. 18, no. 14, pp. 59–76, 2022, https://doi.org/10.3991/ijoe.v18i14.34435

[9] B. M. Alameri, I. J. Kadhim, S. Q. Hadi, A. F. Hassoon, M. M. Abd, and P. Premaratne, "Convolutional deep neural network and full connectivity for speech enhancement," *Int. J. Online Biomed. Eng.*, vol. 19, no. 04, pp. 140–154, 2023, https://doi.org/10.3991/ijoe.v19i04.37577

[10] D. Sarwinda, R. H. Paradisa, A. Bustamam, and P. Anggia, "Deep learning in image classification using Residual Network (ResNet) variants for detection of colorectal cancer," *Procedia Comput. Sci.*, vol. 179, no. 2019, pp. 423–431, 2021, https://doi.org/10.1016/j.procs.2021.01.025

[11] O. Mogotocoro, "Clasificación De Tumores Cerebrales Meningioma, Glioma, Pituitary a Partir De Imágenes De Resonancia Magnética Mediante Wavelet E Inteligencia Ingenieria Electrónica," pp. 20–100, 2022.

[12] W. Abdul, "An Automatic Lung Cancer Detection and Classification (ALCDC) System Using Convolutional Neural Network," pp. 443–446, 2023.

[13] N. M. Balasooriya and R. D. Nawarathna, "A sophisticated convolutional neural network model for brain tumor classification," *2017 IEEE Int. Conf. Ind. Inf. Syst. ICIIS 2017 – Proc.*, vol. 2018, pp. 1–5, 2018, https://doi.org/10.1109/ICIINFS.2017.8300364

[14] W. Ayadi, W. Elhamzi, and M. Atri, "A new deep CNN for brain tumor classification," *2020 20th Int. Conf. Sci. Tech. Autom. Control Comput. Eng.*, pp. 266–270, 2020, https://doi.org/10.1109/STA50679.2020.9329328

[15] O. M. Al-Hazaimeh, M. Al-Nawashi, and M. Saraee, "Geometrical-based approach for robust human image detection," *Multimed. Tools Appl.*, vol. 78, no. 6, pp. 7029–7053, 2019, https://doi.org/10.1007/s11042-018-6401-y

[16] O. M. Al-hazaimeh, A. A. Abu-Ein, N. M. Tahat, M. A. Al-Smadi, and M. M. Al-Nawashi, "Combining artificial intelligence and image processing for diagnosing diabetic retinopathy in retinal fundus images," *Int. J. Online Biomed. Eng.*, vol. 18, no. 13, pp. 131–151, 2022, https://doi.org/10.3991/ijoe.v18i13.33985

[17] V. Galán, "Aplicacion de la metdologia CRISP-DM a Un Proyecto de Minería de datos en el entorno universitario," 2015. [Online]. Available: https://e-archivo.uc3m.es/handle/10016/22198

[18] A. Fernandez, "Algoritmos heurísticos y metaheurísticos basados en busqueda local aplicados a problemas de rutas de vehiculos," Universidad de Valladolid, 2016.

[19] E. Hodelin Maynard, M. Cardona Castillo, G. Maynard Bermúdez, and R. Maynard Bermúdez, "Aspectos epidemiológicos, clínicos y quirúrgicos de los tumores cerebrales metastásicos," *Rev. Inf. Científica*, vol. 98, no. 4, pp. 524–539, 2019, [Online]. Available: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1028-99332019000400524

[20] M. Buda, A. Saha, and M. A. Mazurowski, "Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm," *Comput. Biol. Med.*, vol. 109, pp. 218–225, 2019, https://doi.org/10.1016/j.compbiomed.2019.05.002

[21] M. A. Mazurowski, K. Clark, N. M. Czarnek, P. Shamsesfandabadi, K. B. Peters, and A. Saha, “Radiogenomics of lower-grade glioma: algorithmically-assessed tumor shape is associated with tumor genomic subtypes and patient outcomes in a multi-institutional study with The Cancer Genome Atlas data,” *J. Neurooncol.*, vol. 133, no. 1, pp. 27–35, 2017, https://doi.org/10.1007/s11060-017-2420-1

[22] D. Sarkar, R. Bali, and T. Sharma, *Practical Machine Learning with Python*. Bangalore: Apress, 2018. https://doi.org/10.1007/978-1-4842-3207-1

7 AUTHORS

Dra. Gisella Luisa Elena Maquen-Niño is a Professor at Pedro Ruiz Gallo National University, Lambayeque, Peru, researcher specialized in Artificial Intelligence and Machine Learning. Hired Teacher at the Universidad Mayor de San Marcos. Doctor’s degree in Computer Science and Engineering from the Señor de Sipan University and master’s degree in Information Technology and Educational Informatics. https://orcid.org/0000-0002-9224-5456.

Ariana Ayelen Sandoval-Juarez is a Systems Engineering student at Universidad Nacional Mayor de San Marcos. Training in Database Management, Analysis and Information Management in the banking sector. Handling of Python, .NET, C++ and Java programming languages. https://orcid.org/0000-0002-5534-6370 (email: ariana.sandoval@unmsm.edu.pe).

Robinson Andres Veliz-La Rosa is a Systems Engineering student of the Universidad Nacional Mayor de San Marcos, Lima, Peru, with experience in machine learning and artificial intelligence. https://orcid.org/0000-0003-2064-6629 (email: robinson.veliz@unmsm.edu.pe).

Dr. Gilberto Carrión-Barco is a Professor at the Pedro Ruiz Gallo National University, Lambayeque, Peru. Researcher in data science and digital transformation, educational informatics and process innovation for public management. PhD in Computer Science and Systems from the Señor de Sipán University and Master in Systems Engineering from the Pedro Ruiz Gallo National University. https://orcid.org/0000-0002-1104-6229 (email: gcarrion@unprg.edu.pe).

Mg. Ivan Adrianzén-Olano is a Computer and Informatics Engineer, Master’s in Education Sciences with a mention in Teaching and University Management, completed Master’s studies in Systems Engineering with a mention in Information Systems at the Antenor Orrego Private University of Trujillo. Professor at the Toribio Rodriguez de Mendoza National University of Amazonas, Bagua, Peru. https://orcid.org/0000-0002-1910-2854 (email: ivan.adrianzen@untrm.edu.pe).

Dr. Hugo Vega-Huerta is a Professor at San Marcos University, Lima, Peru, researcher specialized in Artificial Intelligence and Business Intelligence, with more than 30 publications in nationals and internationals journals indexed. Doctorate degree in Systems Engineering at UNFV and master’s degree in business administration at UNMSM. In 2012 he obtained the recognition “Scientific Merit Award”. Google Scholar ID: cffnecwaaaaj, https://orcid.org/0000-0002-4268-5808 (email: hvegah@unmsm.edu.pe).

Mg. Percy De-La-Cruz-VdV is a Full-time professor at the Universidad Nacional Mayor de San Marcos, Lima-Peru, specialist in computer science, research professor with more than 33 years of experience. Member of the Research Group “Innovating Intelligent Systems (YACHAY)”. Research area includes Knowledge Management, Technology Management, Software Engineering, AI-Learning. https://orcid.org/0000-0002-4943-7620 (email: pdelacruzv@unmsm.edu.pe).