Improvement of Deep-Learning Algorithms for Disease Detection: The Case of Cerebral Hemorrhage

Lail Dauris MADAMA[1], Bouchra NASSIH[2] and Aouatif AMINE [3]

[1]* Engineering Sciences Laboratory, National School of Applied Sciences, Ibn Tofail University, Kenitra, Morocco.

[2]Faculty of Economics and Management, Ibn Tofail University, B.P. 241, university campus, Kenitra, Morocco.

[3]Engineering Sciences Laboratory, National School of Applied Sciences, Ibn Tofail University, Kenitra, Morocco.

laildauris.madama@uit.ac.ma

bouchra.nassih@gmail.com,

Amine_aouatif.@yahool.fr

Keywords: CNN, Classification, Cerebral hemorrhage (CVA), Deep learning, Vgg16, Vgg19, Computed tomography (CT).

Abstract

Cerebral haemorrhage is a serious condition and a major public health issue that requires immediate and accurate care to guide doctors in their treatment decisions. This study developed three deep learning models to accurately identify and classify images of haemorrhages based on normal brain images from computed tomography (CT) scans. These models include two pre-trained models (VGG16 and VGG19) and a custom convolutional neural network (CNN). Due to the severe effects of this disease (paralysis, disability, long-term death) and the challenge of identifying and interpreting it for healthcare professionals, the research considered using these models with a dataset comprising two classes: haemorrhagic and normal. The three models were tested under the same conditions, and the results demonstrated each model's ability to generate data. VGG19 showed 99.8% accuracy, 3% loss, 99% detection and classification capability, and 99% sensitivity. The pre-trained VGG16 model generated an accuracy of 99.7%, an estimated error margin of 30%, a detection capacity rated at 99% and a sensitivity of 98%. The custom CNN model performed the worst, with an accuracy of 89%, an error rate of 88%, a recall of 91% and a lower sensitivity level estimated at 84%. The VGG19 approach performed better than the other models.

1. Introduction

Artificial intelligence has proven remarkably effective in the healthcare sector. The integration of innovative machine learning and deep learning technologies has made it possible to correct human errors and provide reliability, confidence and tangible explanations when interpreting the results of several diseases, such as cerebral haemorrhage, predicted by opaque systems. In recent years, health experts have made extensive use of artificial intelligence to conduct concise studies based on cranial imaging (Schemit and al., 2022). Cerebral haemorrhage is bleeding observed in the brain following the rupture of a nerve. It is a major cause of high mortality and morbidity, requiring both accurate diagnosis and rapid treatment to optimise outcomes (Selon Kang and al., 2025). Globally, this condition accounted for 10% of deaths. This death rate could be explained by the difficulties in predicting and stopping blood flow (Umapathy and al.,2023). The severity of a cranial cerebral haemorrhage is, in many cases, defined by blood loss (Ahmed and al.,2023). The incidence of cerebral haemorrhage can take two forms: either immediate or acute. The immediate form can occur at any time in an individual's life, but it is often observed after trauma. Unlike the acute form, it occurs hours or even days after the brain sends signals to the body, but the individual tends to ignore them in most cases (Flaquer-Pérez and al.,2023). This condition manifested itself through several neurological factors such as severe headaches, visual disturbances, and fainting in some cases (Burduja and al.,2020). Other symptoms such as hypertension, malformations, trauma, age, and diet also contributed to its manifestation (Wang and al.,2021). To avoid any damage, surgery based on CT scans is recommended to assess the situation and establish a rapid treatment strategy as soon as possible (Selon Santhoshkumar and al.,2021). In the case of emergency diagnoses involving the nervous system, CT scans are widely used due to their low cost, sensitivity and accuracy in locating haemorrhagic areas (Castro Marcias and al.,2024, Chang and al.,2023). They also facilitate treatment, and certain medical intervention techniques are defined for use in emergencies (Wang and al.,2021). Advanced machine learning and deep learning techniques are used for their ability to process data carefully, evaluate performance and reduce errors. These advantages make them powerful tools in the healthcare sector (Castro Marcias and al.,2024).

However, how can cerebral haemorrhage be effectively detected and classified while minimising errors?

This research, focused on the problem of classification and detection, aimed to design an advanced Deep Learning model that is powerful, reliable, accurate, and robust, based on CNNs and pre-trained VGG19 and VGG16 approaches, capable of predicting, detecting, and classifying cerebral haemorrhage without errors. This work began with an introduction, followed by section 1, which clearly reviewed the literature based on previous work. Section 2 presented the equipment and various methods used to carry out this work. Finally, section 3 combined the results, discussion, future work and conclusion.

2. Related Works

Shweta and R, in their recent article (Swetha and al.,2024), aim to revolutionise the healthcare sector by automating tasks such as classification and detection of cerebral haemorrhage and many other pathologies. Using machine learning and deep learning approaches, they are deploying intelligent systems capable of meeting this challenge. In order to achieve better results, they conducted their research in an organised manner. First, they applied machine learning approaches such as Random Forest, Decision Trees, REPTree and K-Nearest Neighbors (K-NN) to a large dataset comprising a total of 2,501 CT images divided into two classes: the normal class contained 50 cases, 30 cases for the haemorrhagic class and historical data (text). Three different techniques were used to extract relevant information from the images and medical reports (the cosine technique, the wavelet technique and an imaging technique known as GLCM). The models were tested with several K values. Random Forest outperformed all other ML techniques with an estimated accuracy of 86.41%. Secondly, the researchers tested a hybrid CNN-LSTM deep learning approach on the same database with a logical classifier that predicted a categorical value specifying the patient's status (haemorrhagic or normal) at the output. The results were satisfactory. The Deep Learning approach outperforms all machine learning approaches with an accuracy rating of 94%. In the recent article (Chang and al.,2023), locating cerebral haemorrhages was a very difficult task that required intense concentration and the application of effective segmentations techniques. During the segmentation and classification task, they were able to meticulously identify the infected parts of the brain in minute detail. Joonho, Inchul, and Minho designed a hybrid R-CNN model based on or derived from U-NET. The R-CNN accurately located the source of the bleeding in the brain. To better locate the area, they used a colour-coded marking system as a reference to delineate the region of the image. They then used U-NET specifically to segment the pixels of the images in order to extract the relevant information. The researchers then combined U-NET and CSD for a semi-prediction of the haemorrhagic areas. Finally, with all this information, they performed a global classification using CNN to predict haemorrhagic or normal status. The hybrid R-CNN model showed an accuracy of 69.7%, close to 1, and a segmentation deviation of 12.9%. In (Selon Chen and al.,2024), intracranial haemorrhage is seen as a potential threat to public health. The devastation caused by this condition changed people's lives on a daily basis and could not go unnoticed. Looking at the high annual mortality rate, 50% of deaths worldwide were caused by this condition. Accurate diagnosis and rapid treatment contributed to first aid and reduced all fatal risks. Yu-Ruei, Chih-Chieh, Chang-Fu and Ching focused their work on new technologies using deep learning models. In the hope of finding a model capable of extracting characteristics down to the smallest details without overlooking anything, an advanced convolutional neural network model was designed to improve accuracy and double the credibility of the output results. Two control databases, CQ500 and PhysioNet, were used for testing and validation on 5,682 images. The model showed promise with an estimated accuracy of 95.2%. These results demonstrated the model's good interaction with the data. In (Ahmed and al.,2024), cerebral haemorrhage was a serious neurological condition that required rapid detection and monitoring. In this study, Samia, Jannatul, Rahman, Shamim, and others focused on an in-depth and comparative study of existing work on the detection and classification of cerebral haemorrhage over the period 2012-2023. This period was marked by a study of 54 projects in which numerous machine learning and deep learning approaches were deployed on different databases. Several extraction, segmentation, class balancing and normalisation techniques were applied (filtering, windowing, partitioning, resizing, image conversion). The aim was to produce a detailed report on the involvement of ML and DL algorithms in the classification task. By comparing these two giants, the summation of the results showed that DL algorithms such as CNN, VGG, ResNet50, and CNN-LSTM far outperformed all ML approaches (KNN, SVM, ANN, random forest, decision tree) to date. In (Mahjoubi and al.,2023), cerebral haemorrhage represented a serious threat to public health worldwide. This alarming threat was a topical issue that required rapid detection and classification by doctors in order to better control the situation. Recognising the significance of this disease and the difficulties healthcare professionals face in recognising haemorrhagic images, researchers Mohamed, Soufiane and others developed two advanced CNNs (Vgg16, Vgg19) to help doctors detect the areas of the skull affected by bleeding and then classify them accurately. The models were trained and evaluated on several metrics. The VGG16 model performed best, with an accuracy of 99.10%. In contrast, VGG19 had an accuracy of 97.21%. VGG16 minimises errors better and is well suited to classification tasks. In the recent article (Qdaih and al.,2024), a stroke is defined as the rupture of a blood vessel due to the accumulation of blood clots or an increase in blood pressure, thereby slowing blood flow to the brain. This situation was concerning and marked by the fact that doctors had difficulty accurately detecting the affected area. Wanting to help medicine take a step forward in imaging detection and classification, Ibrahim, Omar, Suluman, Diala and others proposed a hybrid HEDL model that combined DenseNet121 with edRFL. In order to obtain the best results, they used a specialised method (MedSAM-LOVE5) to extract and segment the image into several parts to extract the relevant information. The model was evaluated for the first time, but, disappointed with the results, they relied on a SPEM correction technique to improve the quality and readability of the images by increasing the brightness to optimise the results. The model was tested again and achieved an accuracy of 92.5%, up from 87%. These results highlighted the contribution of the improvement technique (SPEM). In (Bishi-Il and al.,2024), the health registry classified cerebral haemorrhage as one of the leading causes of death worldwide. Its location and classification required a high level of expertise and high-quality scanning tools. The researchers set up a deep convolutional neural network. The idea was to integrate it with several techniques for feature extraction (windows), region localisation (CAM) and relevant information assembly (feature combination) so that the model would perform well without errors in classification and localisation tasks. The model was trained and evaluated on an ASNR database of 75,000 images and achieved a measured accuracy of 97.3%. This model showed promise and better regulates the data.

Although previous work had pushed scientific thinking to unprecedented heights, it was significantly hampered by several issues, such as: firstly, the latency observed when interpreting results; the integration of an automated interpretation tool helped doctors to make decisions (Chen and al.,2024). Second, the lack of radiological equipment limited the availability of data. The staff at radiology centres increased the availability of data. In addition, there were problems with interpretation and transparency (Qdaih and al.,2024). Finally, there were shortcomings in the extraction and segmentation systems used to identify the affected area of the image. The use of high technology such as 3D scanners should be a viable solution (Swetha and al.,2023, Chen and al.,2024).

Table 1: Summary of previous works

https://doi.org/10.5194/isprs-archives-XLVIII-4-W19-2025-85-2026 | © Author(s) 2026. CC BY 4.0 License.

Study	*Year*	*Datasets*	*Methods*	*Optimization*	*Accuracy*
[1]	2024	TDM	CNNLSTM	SGD	94
[7]	2023	CT-BRAIN	R-CNN	SGD	70
[20]	2024	RSNA/CQ500	CNN-2D	ADAM	95.2
[3]	2023	CT-BRAIN	VGG19	ADAM	99.10
[6]	2024	CT-BRAIN	HEDL	ADAM	92.5
[21]	2024	ASNR	CNN	ADAM	97.3

3. Materials and Methods

3.1. Methods Proposed model

Pioneering work on the detection and classification of cerebral haemorrhage carried out in the healthcare sector has encountered several problems, including: the difficulty for healthcare experts to identify the haemorrhagic area, which influenced decision-making. Secondly, the use of limited data. However, the more diverse the data, the better the classification model learned and optimised results while minimising errors. To overcome all the previous difficulties, this work proposed the implementation of pre-trained VGG16 and VGG19 models and developed a customised CNN-2D model [6]. At the input stage, this model was equipped with a three-dimensional layer comprising three windows to filter the information, with only the relevant information being retained. This process ensured effective cleaning and better preservation of relevant information. At this stage, the data was normalised for the first time using the Min-Max Scaler technique to bring it to the same scale in order to avoid any size imbalance. The CAM region marker is used to materialise and visualise the haemorrhagic area before moving on to the segmentation of the marked areas. The data is then divided to ensure its integrity and avoid overfitting. Finally, an advanced SPEM technique is applied to make the images more readable down to the smallest detail. At the end, the three models (CNN-2D, VGG16 and VGG19) are compared to determine which one has been the most promising and reliable for real-time deployment in healthcare centres to help doctors make good decisions. This model solves the problem of accurately locating haemorrhagic areas, reduces human error, improves decision-making and ensures patient privacy, which should increase the deployment of large databases (Bishi-Il and al.,2024).

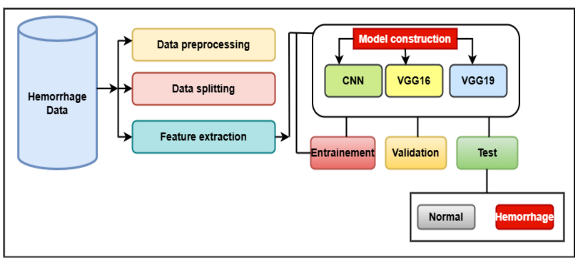

Figure 1: Proposed model

3.2. Data collection

Detecting and classifying cerebral haemorrhages has always been a very difficult task requiring well-structured data. This research was based on data from Kaggle. This dataset included two classes: a haemorrhagic class, labelled with zero (0), which was used during the training phase, and a normal or control class, which was used to validate and test the results (Agarwal and al.,2023).

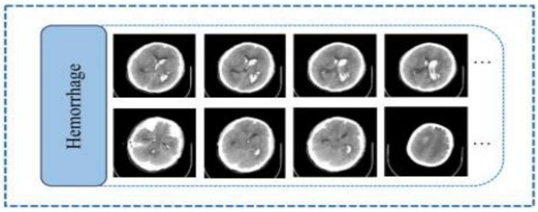

Figure 2: Viewing database classes (Wang H and al.,2024).

For more detailed information about our database, we can describe its content and organization more clearly as follows:

Table 2: Organization of our database	*Nbr of hemorragique images*	*Nbr of normal images*
Labels of the dataset	2598	4105
Labels of the training set	2078	3284
Labels of the validation set	260	411
Labels of the test set	260	410

3.3. Data Preprocessing

The process began with data collection, followed by pre-processing to extract the most relevant features. To do this, several methods were generally applied to reduce the least relevant information, eliminate noise and remove random variations in the images in order to obtain high-quality images that could perfectly optimise or further improve our classification results. Among these methods, a few are worth mentioning: first, the standardisation and normalisation process, which consisted of bringing the data to the same scale and size while standardising it. Next, rotation and flipping, which are very important techniques that provided a broad representation of the features. Finally, the creation of the appropriate model, training, testing, validation and parametric adjustment, which promoted a clear view of the data, thereby improving the contours of the images (Jingya and al.,2024).

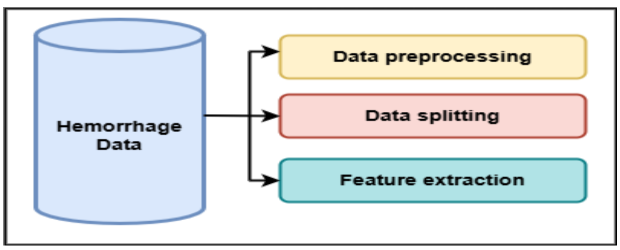

Figure.3: Data Preprocessing

3.4. Convolutional Neural Networks (CNN)

It is a revolutionary approach to deep learning commonly used in image detection and classification tasks for its ability to enrich, diversify, and make data tangible for successful classification. This architecture was developed in 1979 by K. Fukushima. It works in a similar way to the human brain (biological neuron), and its numerous layers stacked on top of each other are information processing centres that produce organised and uniform information. This algorithm is renowned for its accuracy in diagnosing neurological diseases. Its strength lies in its meticulous feature extraction, which has made it easy to learn and automate binary prediction tasks (Umapathy and al.,2023). In most cases, CNNs used CT images to perform numerous tasks such as segmentation, classification, detection, pre-processing, reducing human error and learning time (Anjum and al.,2023, Kousar and al.,2025). CNN was structured and organised into several layers: convolutional layers, pooling layers, normalisation layers, fully connected layers, and a ReLu activation function. This model consists of 13 convolutional layers comprising 3 filters. At the output of the convolutional layers, a maximum pooling resizing layer is applied to preserve relevant information. To generate a final prediction, the sigmoid function is used. One of the distinctive features of CNN is its ability to reduce the number of connections between neurons (Li and al.,2020).

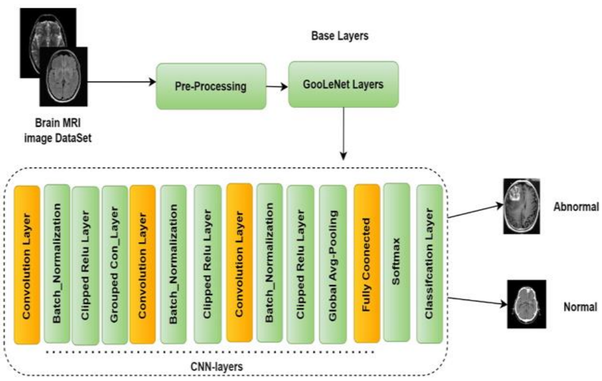

Figure 4: Visualisation of the CNN architecture (Amran and al.,2021)

3.5. The VGG16 Model

Pre-trained model or advanced convolutional neural network based on large calculations. In VGG16, each block consists of several layers of convolutions and max pooling. The pre-trained VGG16 model is used for various tasks such as classification, detection and segmentation. The unique feature of this network was its increasing performance. This advanced network has a well-structured layered organization: its thirteen (13) convolution layers, reinforced by input filters, enabled it to accurately extract image features. These maximum pooling layers sectioned the images, resized them and retained the most agile information for binary prediction. All this information was then fed into the dense layers for a final prediction of cerebral haemorrhage (Kousar and al.,2024).

3.6. The VGG19 Model

VGG19: Like VGG16, VGG19 is a pre-trained model comprising nineteen (19) layers, including 16 convolution layers, three (03) fully connected layers, as well as maximum pooling layers and an activation function. During the extraction phase, the convolution layers received the raw information, and the passage of information from one layer to another underwent processing. This extraction is reinforced by three-dimensional filters, contributing to a thorough analysis. The subsampling layers performed a second extraction by resizing the images, with the aim of retaining useful information. The operating principle remained the same as that of VGG16, with only the number of layers differing. VGG19 is used for its high capacity to extract and carefully process images, which was an advantage for it. VGG19 is used for classification, segmentation and disease detection tasks based on scanner imaging (Kousar and al.,2024).

3.7. Hyperparameters

Optimising the learning and adjustment performance of the VGG16, VGG19 and CNN models required the use of several hyperparameters.

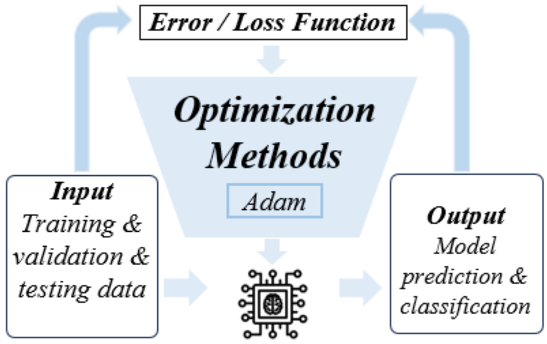

Figure 5: Deep Learning Model Workflow Diagram

Figure 5 illustrates a typical workflow for a DL model. It begins with input data, which is generally divided into three parts: training, validation, and testing. The approach developed uses this data to make predictions, which are then evaluated using an error or loss function. The optimisation method, in particular Adam (a popular optimisation algorithm), is used to optimise the model's performance while reducing the margin of error. The final result expected from the model following the iterative process is prediction and classification. The diagram provides an overview of the main components involved in the progression and refinement of a DL model.

Table 3: Best Hyperparameters Used for Our Models

Models	*Learnig Rate*	*Optimizer*	*Batch Size*	*Epochs*
CNN	0.005	ADAM	16	24
VGG16	0.005	ADAM	16	24
VGG19	0.005	ADAM	16	24

This table shows the hyperparameters used to train and optimise the prediction results.

4. Description of Results

4.1. Accuracy Based on CNN VGG16 and VGG19 for Brain Hemorrhage Classification and Detection

These graphs are simply the results of our various models, which were trained on 80% of the data set, 10% of the validation data, and 24 epochs each. These graphs allowed us to observe the different variations in the models and their great ability to reduce errors. According to these graphs, in Figure 7 of the CNN model, from the first (01) epoch onwards, the model began its learning process with the training curve, while the validation curve sought to better understand the data. From

https://doi.org/10.5194/isprs-archives-XLVIII-4-W19-2025-85-2026 |

the second (02) to the fifth (05) epoch, this model experienced a sharp increase in accuracy. After the fifth (05) epoch, the training curve remained constant until the end of the process. The validation curve then decreased sharply. Subsequently, accuracy increased and followed the training curve. In Figure 8, we observed that training began in the first epoch and that a sharp increase in accuracy was visible until the fifth (05) epoch. Between epochs 5 and 7, the VGG16 model began learning from the first epoch with a slight delay on the validation curve, and a sharp increase in accuracy was observed from the first five epochs onwards. A decrease in accuracy began between the seventh and eighth epochs, then the validation curve increased as in its initial state, and from epoch 9 until the end, the two curves collided, thus stabilising the curves to migrate in the same direction. And the VGG19 model in Figure 9, unlike the VGG16 and CNN models, showed that both the training and validation curves began learning from the first epoch. The more epochs we increased, the more the accuracy tended to reach the peak of perfection (01) with a very low margin of error. The two curves progressed together and followed the same direction. Of all the existing models, VGG19 is the most promising model with good error minimisation without overfitting.

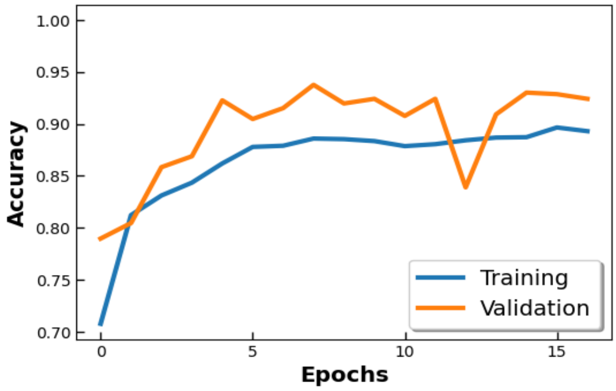

Figure 6: CNN model accuracy

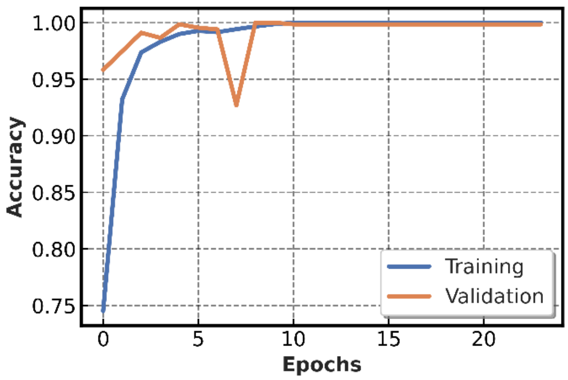

Figure 7: VGG16 model accuracy

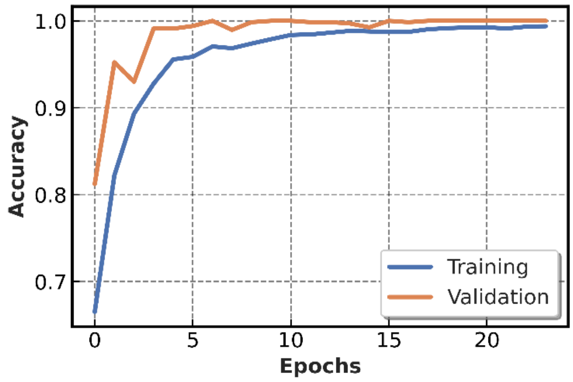

Figure 8: VGG19 model accuracy

4.2. Loss Based on CNN VGG16 and VGG19 for Brain Hemorrhage Classification and Detection

Graphs 10, 11, and 12 of the CNN, VGG16, and VGG19 models presented results that involved both the training and validation curves to highlight the models' ability to minimise errors. From the first epoch, the models began to learn well from the data and reduce errors. The increase in accuracy was dependent on the number of epochs. Increasing the number of epochs increased accuracy and significantly reduced losses. Hence, there was a positive and increasing relationship between the increase in epochs and accuracy. And a negative and decreasing relationship between the increase in epochs and errors. However, the validation curves of all models were lower than the training curves, which allowed for an evaluation in terms of losses for each model. The average error of the CNN model was estimated at 88%, with a 30% margin of error for VGG16 and 3% losses for VGG19. These results demonstrated the ability of the VGG19 model to better adapt to medical data for detection and classification tasks while offering better generalisation of cerebral haemorrhage (Kousar and al.,2024).

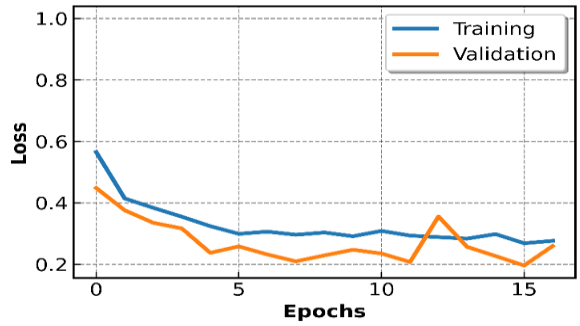

Figure 9: CNN model loss

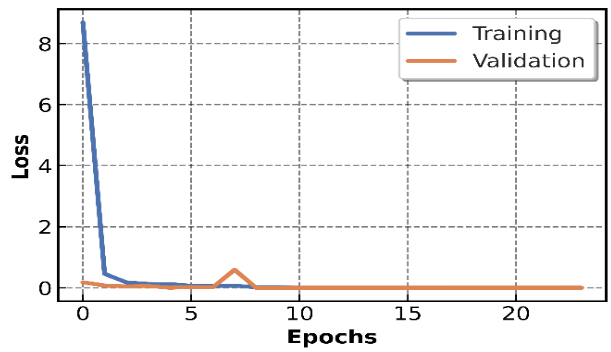

Figure 10: VGG16 model loss

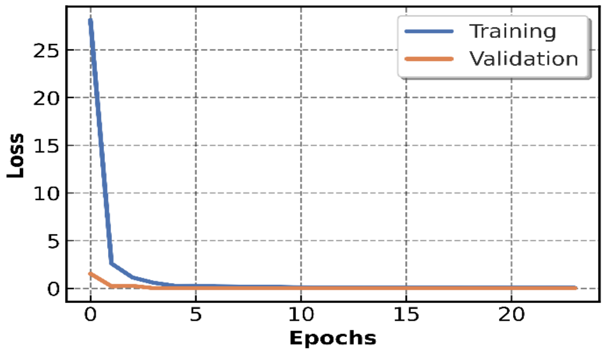

Figure 11: VGG19 model loss

4.3. ROC Based on CNN VGG16 and VGG19 for Brain Hemorrhage Classification and Detection

The area under the curve is generally used to evaluate the performance of detection and classification models. Three models (CNN, VGG16, VGG19) were evaluated: the CNN approach performed well with an estimated AUC of 87% for both classes: haemorrhagic and normal. It was observed that the two curves of this model were ineffective in reaching the left axis, which significantly impacted accuracy and made it unable to recognise images. On the other hand, the VGG16 and VGG19 models stood out with remarkable performance, curves that reached the left axis, and an area under the curve (AUC) with near-perfect accuracy (1.00%). These results highlighted the effectiveness of the models in differentiating diseased images from healthy images (Mahjoubi and al.,2023). The curves of the two VGG models had much better accuracy than the CNN curves

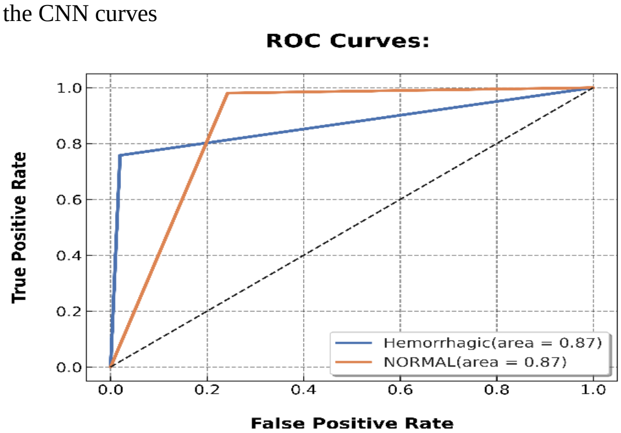

Figure 12: CNN model Roc curve

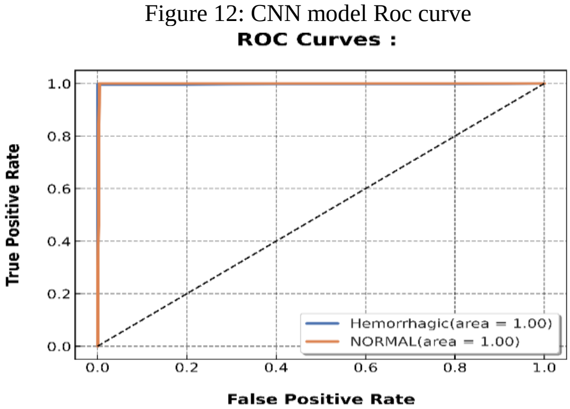

Figure 13: VGG16 model Roc curve

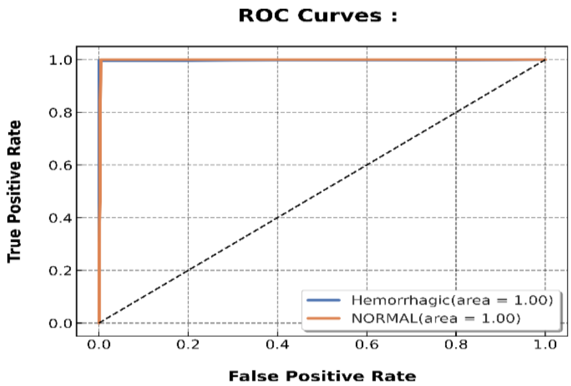

Figure 14: VGG19 model Roc curve

4.4. Confusion Matrix Based on CNN VGG16 and VGG19 for Brain Hemorrhage Classification and Detection

The performance of prediction models is generally evaluated using the confusion matrix technique. Figure 16 showed a high error rate estimated at 71 in total. These observed errors demonstrated the limitations of CNN in correctly identifying haemorrhagic cases from normal cases. CNN does not learn data well during training, which could explain these errors. The pre-trained VGG models demonstrated great insight in correctly classifying normal and haemorrhagic cases with a reduced margin of error. Figure 17 showed three (3) errors, including one (1) normal case detected as haemorrhagic and two (2) haemorrhagic cases predicted as normal. Figure 18, corresponding to the VGG19 approach, demonstrated excellent performance with only one (1) error observed. This model learned well from the training and validation data, which justified its effectiveness in accurately predicting false positives and false negatives. In terms of performance, the VGG approaches outperformed the custom CNN model in the task of detecting and classifying cerebral haemorrhage. VGG19 surpassed the other two models with promising, reliable and robust performance. This made it the best of all.

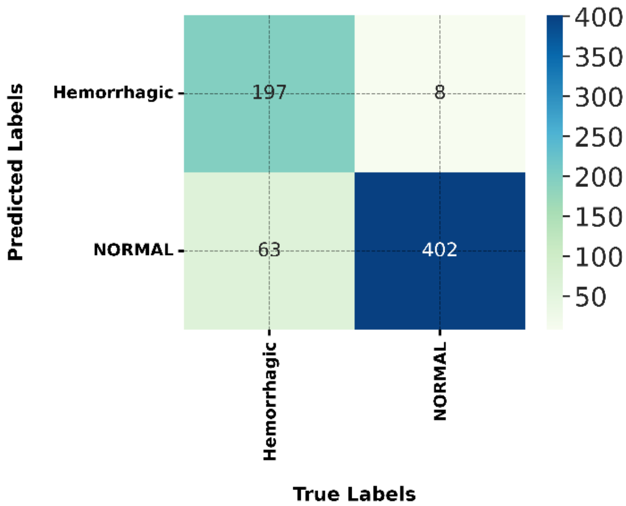

Figure 15: CNN model Confusion matrix

https://doi.org/10.5194/isprs-archives-XLVIII-4-W19-2025-85-2026 | © Author(s) 2026. CC BY 4.0 License.

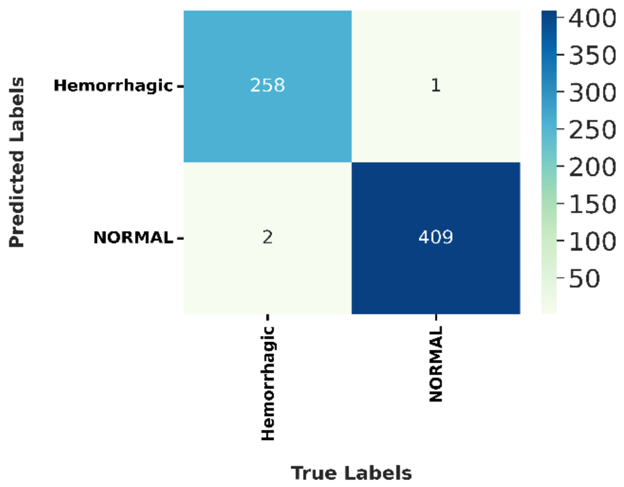

Figure 16: VGG16 model Confusion matrix

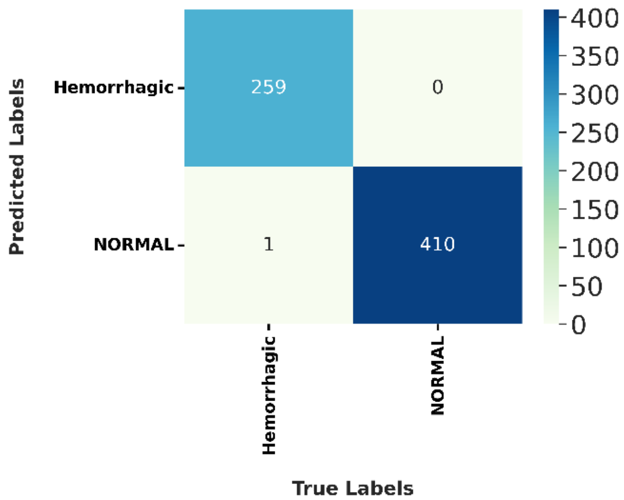

Figure 16: VGG19 model Confusion matrix

4.5. Models Evaluation

This comparative table corresponds to the prediction results of different Deep Learning approaches that were trained, evaluated and tested using a clinical database. Several parameters were used to optimise the results. Pre-trained models performed better than the custom CNN model. Indeed, the metrics used to evaluate model performance showed that the pre-trained VGG19 approach had an accuracy of 99.8%, an error rate reduced to 3%, a recall of 99.9% and an F1 score of 99%. Under the same conditions, the VGG16 approach presented the following metrics: 99.7% accuracy, 30% error margin, 99% recall and 98% F1-score. Finally, the customised CNN model demonstrated acceptable performance, but was less effective than the other models. CNN had an accuracy of 89%, losses of up to 88%, a recall of 91% and an F1 score of 84%. Based solely on these results, the pre-trained VGG19 approach was better suited to clinical data, while minimising errors. This suitability, combined with its performance, made it the best and most promising of all.

Table 4: Summary of Analysis Results

Models	*Accuracy*	*Loss*	*Recall*	*F1-score*
CNN	89%	0880	91%	84%
VGG16	99.7%	0.030	99%	98%
VGG19	99.8%	0.003	99.9%	99%

5. Discussion

The detection and classification of cerebral haemorrhage highlighted several results that made it possible to evaluate the capacity of three models: VGG19, VGG16 and CNN. Figure 9 shows that the VGG19 model began learning from the first epoch, with both curves (training and validation) increasing together in terms of epochs and accuracy, while following the same path until the end. The VGG16 model, shown in Figure 8, began learning with a slight delay in the validation curve compared to the training curve. A significant increase in accuracy was observed in the first five (5) epochs. From the sixth (6) to the ninth (9) epoch, a sharp decrease in accuracy was observed, allowing the two curves to collide and follow the same direction. In Figure 7, CNN training was marked from the first epoch. Indeed, from this initial epoch, the training curve took the lead over the validation curve. During the learning phase, both curves performed well, with an increase in accuracy from the first five (5) epochs. Between the eleventh (11) and thirteenth (13) epochs, a large decrease in accuracy is observed. At the fourteenth (14) epoch, accuracy increased and followed the training curve. In terms of error evaluation, the training and validation curves in Figures 10, 11, and 12 showed that learning was successful overall, without overfitting. VGG19 had 3% errors, VGG16 had 30% errors, and CNN had 88% errors. The VGG19 and VGG16 models (Figures 14 and 15) demonstrated an estimated area under the curve (AUC) of 1.00 for both. The customised CNN approach (Figure 13) had an AUC of 87%. Furthermore, in terms of classification, VGG19 showed that zero normal cases were detected as haemorrhagic, and one haemorrhagic case was predicted as normal (Figure 18). The VGG16 approach, on the other hand, presented one normal case predicted as haemorrhagic and two (2) haemorrhagic cases predicted as normal. Finally, CNN presented eight (8) normal cases predicted as haemorrhagic and sixty-three (63) haemorrhagic cases predicted as normal. After the testing phase, the models presented several results. VGG19 had 99.8% accuracy, 99.9% recall, and an F1 score of 99%. VGG16, on the other hand, had 99.7% accuracy, 99.9% recall, and an F1 score of 84%. CNN demonstrated the following performance: 89% accuracy, 91% recall, and 84% F1 score. These results highlighted several relationships: increasing the number of epochs affected both accuracy and errors. In other words, as the number of epochs increased, accuracy also increased and errors decreased. Furthermore, the results depend on the quality of data processing. Good processing yields good results, as in the case of VGG19, and poor processing yields poor results, as in the case of CNN. The three models were tested under the same conditions, but VGG19 stood out in terms of its results. Healthcare remains a highly sensitive sector where decision-making is crucial. The errors recorded had the effect of misleading healthcare professionals in their decision-making. And the use of limited data restricted the scope of the results on a global scale. Based on the results obtained, we can say that the pre-trained VGG models are far superior to the customised CNN model. VGG19 has proven itself in data generalisation. Its adaptation to clinical data makes it a model that effectively predicts cerebral haemorrhage while minimising errors. When compared to previous research, specifically Article 3 by Mahmed Amine Mahjoubi, our work is vastly superior with remarkable performance. Our pre-trained VGG19 and VGG16 models have 99.8% and 99.7% accuracy, compared to 97.21% and 99.1%. In terms of detection and classification, our confusion matrices show one (1) error for the VGG19 model and three (3) errors for the VGG16 model, compared to twenty-eight (28) errors for the VGG19 model and nine (9) errors for the VGG16 model. These results highlight the robustness, reliability and high capacity to correctly detect and classify haemorrhagic cases from normal cases.

Limitations and future work

The use of non-diversified data led to confusion in some of our models during classification. This explained the average performance and large margin of error observed for the custom CNN. Despite the limited data size, our VGG models excelled with a very small margin of error. Future work could focus on a number of areas: the use of hybrid models, exploring all possible combinations in order to not only optimise results but also make them more reliable. This could lend credibility to the use of AI in several hospital structures. In addition, they should consider combining optimisers to speed up the process and reduce transfer time [Villringer and al.,2025]. Finally, they could explore this fundamental avenue: using automatic 3D scanners combined with sensors for better image capture. The image remains the starting point for this type of task.

6. Conclusion

The study focused on the issue of effective detection and classification of cerebral haemorrhages and minimising errors. The research was guided by the concept of finding an effective, safe, accurate and robust learning model based on CNN and the pre-trained VGG19 and VGG16 methods. The model can effectively predict, detect and classify cranial haemorrhage with accuracy. To achieve this, these models were compared. The results showed that the pre-trained VGG19 model was the most effective of all in terms of its ability to interact better with clinical data. In this case, VGG19 responded effectively to the problem with the results provided. Careful analysis of this approach showed the following performances: accuracy: 99.8%, losses: 3%, recall: 99.9% and an F1 score of: 99%, with only one classification error observed by the confusion matrix. The VGG16 model, on the other hand, had 99.7% accuracy, 30% margin of error, 99% recall, 98% F1 score and a total of three (3) estimated classification errors. Finally, CNN demonstrated the following performances: accuracy: 89%, errors: 88%, recall: 91%, F1 score: 84%, and seventy-one (71) classification errors were recorded.

References

Schmitt, N., Mokli, Y., Weyland, C. S., Gerry, S., Herweh, C., Ringleb, P. A., & Nagel, S. (2022). Automated detection and segmentation of suspected intracranial haemorrhage hyperdensities in non-contrast CT scans of patients with acute stroke. European Radiology, 32, 2246–2254. https://doi.org/10.1007/s00330-021-08352-4

Kang, D.-W., Kim, M., Park, G.-H., Kim, Y. S., Han, M.-K., Lee, M., Kim, D., Ryu, W.-S., & Jeong, H.-G. (2025, mars 21). *Deep learning-assisted detection of intracranial hemorrhage: Validation and impact on reader performance*. **Neuroradiology**. https://doi.org/10.1007/s00234-025-03560-x

Umapathy, S., Murugappan, M., Bharathi, D., & Thakur, M. (2023). Automated computer-assisted detection and classification of intracranial hemorrhages using ensemble deep learning techniques. *Diagnostics*, *13*(18), 2987. https://doi.org/10.3390/diagnostics13182987

Ahmed, S. N., & Prakasam, P. (2023, juillet). *A systematic review on intracranial aneurysm and hemorrhage detection using machine learning and deep learning techniques*. **Progress in Biophysics and Molecular Biology**, *183*(1), 1–16. https://doi.org/10.1016/j.pbiomolbio.2023.07.001

Flaquer-Pérez de Mendiola, J. M., Arboix, A., García-Eroles, L., & Sánchez-López, M. J. (2023, 16 janvier). *Acute spontaneous lobar cerebral hemorrhages present a different clinical profile and a more severe early prognosis than deep subcortical intracerebral hemorrhages: A hospital-based stroke registrystudy*.**Biomedicines**,*11*(1),223. https://doi.org/10.3390/biomedicines11010223

Burduja, M., Ionescu, R. T., & Verga, N. (2020). Accurate and efficient detection and classification of intracranial haemorrhage subtypes in 3D computed tomography scans using convolutional and long-term memory neural networks. Sensors, 20(19), 5611. https://doi.org/10.3390/s20195611

Wang, X., Shen, T., Lan, J., Xu, Y., Wang, M., Zhang, J., & Han, X. (2021, 11 août). *A deep learning algorithm for automatic detection and classification of acute intracranial hemorrhages on head CT scans*. **NeuroImage: Clinical**, *33*, Article 102785. https://doi.org/10.1016/j.nicl.2021.102785

Santhoshkumar, S., Varadarajan, V., Gavaskar, S., Amalraj, J. J., & Sumathi, A. (2021, October 21). *Machine learning model for intracranial hemorrhage diagnosis and classification*. **Electronics**,*10*(21),2574.https://doi.org/10.3390/electronics10212574

Castro-Macias, F. M., Morales-Alvarez, P., Wu, Y., Molina, R., & Katsaggelos, A. K. (2024). Hyperbolic secant representation of the logistic function: Application to probabilistic multi-instance learning for intracranial haemorrhage detection in computed tomography (CT). Artificial Intelligence, 331, 104115. https://doi.org/10.1016/j.artint.2024.104115

Chang, J., Choi, I., & Lee, M. (2023). PESA R-CNN: Segmentation of haemorrhage and perihaematoma oedema using an adaptive scale-guided R-CNN. IEEE Journal of Biomedical and Health Informatics, 27(1), Article 3220820. https://doi.org/10.1109/JBHI.2022.3220820

Swetha Mucha, & Ramesh Babu, A. (2024). Classification of intracranial haemorrhage images (CT) using the CNN-LSTM method and image-based GLCM features. ICMED 2024, MATEC Web of Conferences, 392, 01075. https://doi.org/10.1051/matecconf/202439201075

Chen, Y.-R., Chen, C.-C., Kuo, C.-F., & Lin, C.-H. (2024). An effective deep neural network for automatic classification of acute intracranial haemorrhages in brain scans. Computational Biology and Medicine, 176, 108587. https://doi.org/10.1016/j.compbiomed.2024.108587

Ahmed, S., Esha, J. F., Rahman, M. S., Kaiser, M. S., Hosen, A. S. M. S., Ghimire, D., & Park, M. J. (2024). Exploration des approches d'apprentissage profond et d'apprentissage automatique pour la détection des hémorragies cérébrales. IEEE Access, 12, article 3376438. https://doi.org/10.1109/ACCESS.2024.3376438

Mahjoubi, M. A., Hamida, S., Siani, L. E., Cherradi, B., & Raihani, A. (2023). Deep learning for cerebral hemorrhage detection and classification in head CT scans using CNN. In Proceedings of the 2023 3rd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)(pp. …). IEEE. https://doi.org/10.1109/IRASET57153.2023.10153010

Qdaih, I., Daraghmeh, O., Thwib, S., Vicuna Polo, S., Atari, S., & Abu Al-Halawa, D. (2024). Hybrid deep learning model to improve the detection and classification of ischaemic strokes in clinical application. Journal of Imaging, 10(7), 160. https://doi.org/10.3390/jimaging10070160

Bishi-Il, Xu, Z., Zhou, D., & Zhang, L. (2024). Deep multi-scale convolutional learning for intracranial haemorrhage classification and weakly supervised localisation. Heliyon, 10, e30270. https://doi.org/10.1016/j.heliyon.2024.e30270

Agarwal, S., Wood, D., Grzeda, M., Suresh, C., Din, M., Cole, J., Modat, M., & Booth, T. C. (2023). Systematic review of artificial intelligence for anomaly detection in high-volume neuroimaging and meta-analysis of subgroups for intracranial haemorrhage detection. Clinical Neuroradiology, 33, 943–956. https://doi.org/10.1007/s00062-023-01291-1

Wang, H., Wang, G., Li, Y., & Zhang, K. (2024). YOLO-HV : A fast method based on YOLOv8 for measuring haemorrhagic volumes. Biomedical Signal Processing and Control, 90, 107131. https://doi.org/10.1016/j.bspc.2024.107131

Jingya, L. (2024). Towards intelligent detection of cerebral haemorrhages: advanced medical imaging technologies based on machine learning. In J. Lee (Ed.), Proceedings of the 1st International Conference on Engineering Management, Information Technology and Intelligence (EMITI 2024) (pp. 422–428).SCITEPRESS. https://doi.org/10.5220/0012939500004508

Anjum, N., Sakib, A. N. M., & Ahsan, S. M. M. (2023). Classification of Brain Hemorrhage Using Deep Learning from CT Scan Images. In A. Ghosh et al. (Eds.), *Proceedings of ICCIDE2022*(pp.xxx–xxx).Springer. https://doi.org/10.1007/978-981-19-7528-8_15

Kousar, T., Rahim, M. S. M., Iqbal, S., Yousaf, F., & Sanaullah, M. (2025). Applications of deep learning algorithms in the detection, segmentation, and classification of ischaemic strokes. Artificial Intelligence Review. https://doi.org/10.1007/s10462-025-11119-8

Li, J., Fu, G., Chen, Y., Li, P., Liu, B., Pei, Y., & Feng, H. (2020). A multi-label classification model for whole-brain computed tomography images. BMC Bioinformatics, 21(Suppl 6), 200. https://doi.org/10.1186/s12859-020-3503-0

Villringer, K., Sokiranski, R., Opfer, R., Spies, L., Hamann, M., Bormann, A., Brehmer, M., Galinovic, I., & Fiebach, J. B. (2025). An artificial intelligence algorithm integrated into the clinical workflow can ensure high-quality CT diagnosis of acute intracranial haemorrhage. Clinical Neuroradiology, 35, 115–122. https://doi.org/10.1007/s00062-024-01461-9

Amran, G. A., Alsharam, M. S., Blajam, A. O. A., Hasan, A. A., Alfaifi, M. Y., Amran, M. H., Gumaei, A., & Eldin, S. M. (2022). *Brain Tumor Classification and Detection Using Hybrid Deep Tumor Network. Electronics, 11*(21), Article 3457. https://doi.org/10.3390/electronics11213457