Intracranial Hemorrhage Detection Using Deep Learning and Transfer Learning

MSc Research Project
Data Analytics

Rohit Kumar
Student ID: X15004902

School of Computing
National College of Ireland

Supervisor: Jorge Basilio

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Rohit Kumar
Student ID:	X15004902
Programme:	Data Analytics
Year:	2020-2021
Module:	MSc Research Project
Supervisor:	Jorge Basilio
Submission Due Date:	16/8/2021
Project Title:	Intracranial Hemorrhage Detection Using Deep Learning and Transfer Learning
Word Count:	7019
Page Count:	22

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Rohit Kumar
Date:	23rd September 2021

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Intracranial Hemorrhage Detection Using Deep Learning and Transfer Learning

Rohit Kumar
X15004902

Abstract

Intracranial Hemorrhage (ICH) is a life-threatening medical occurrence that is associated with a poor result despite optimal care. Given that early detection and care of ICH can improve health outcomes, there is a need for a triage system that can promptly detect and speed the treatment process. Previously published work took a more traditional technique, comprising numerous steps of alignment, image analysis, image rectification, handmade image segmentation, and classification. This research work examines the intracranial hemorrhage detection problem and develops a deep learning model and transfer learning models to reduce the time required to identify hemorrhages. For classification of ICH sub types, we developed a convolutional neural network based on the Transfer learning Model. DenseNet121, Xception and CNNs were compared with using many evaluation criteria to ensure that the model's results are accurate and that it does an excellent job. As predicted, the system delivers impressive results, and the data reveal that Xception is more successful than competing models. For the identification and classification of ICH subtypes, the Xception model is used for the final output.

1 Introduction

A serious brain stroke caused by intracranial hemorrhage can happen at any age in anybody, but it is especially dangerous in people with high blood pressure because the blood flow to the brain is restricted also the blood leakage into the brain tissues from blood vessels that supply blood, it is a significant risk to a person's life and should be taken seriously. Several factors, including high blood pressure, abnormally unstable or dilated (aneurysm) blood vessels that rupture, mistreatment, and injury, can contribute to intracranial brain hemorrhaging. Several human beings who experience from an Intracranial hemorrhage (ICH) suffer symptoms related with those of a stroke, including muscular fatigue, speech problems, and sensations in their arms and legs. Early detection of ICH, especially within 24 hours, is critical in reducing patient death (Tomasz et al.; 2021). Aside from that, intracranial hemorrhage (ICH) is a brain disorder distinguished with blood vessel injury as well as infected cells which extend towards the brain ventricles throughout the period and are particularly bad for people (Sage and Badura; 2020). Overall amount of intracranial hemorrhage incidents that occur in a year is about 30,000. There appear to be nearly 795,000 strokes per year in the United States solely, as per available information (Burduja et al.; 2020). During the year 2008, since there were estimated 168 cases per 100,000 Vietnamese people, the country deemed ICH to be among the most serious illnesses.(Luong et al.; 2020). Intracranial Hemorrhage (ICH) is a serious consequence of traumatic brain injury (TBI). Hassan, N (2021).

It is possible to identify between ICH subtypes based on the location of the bleeding in the below figure 1 ICH subtypes are shown. Subdural hemorrhage (SDH) is defined as bleeding between both the dura and the arachnoid, whereas epidural hemorrhage (EDH) is defined as bleeding between the dura and the bony skeleton (EDH). Traumatic injuries frequently result in one or more of the following effects. A type of bleeding that occurs within the parenchyma of the brain, is known as Intraparenchymal hemorrhage(IPH), Intraventricular hemorrhage (IVH) is a type of hemorrhage that occurs within the ventricular system. Finally, blood in the subarachnoid space is a sign of subarachnoid hemorrhage (SAH) (Sage and Badura; 2020).

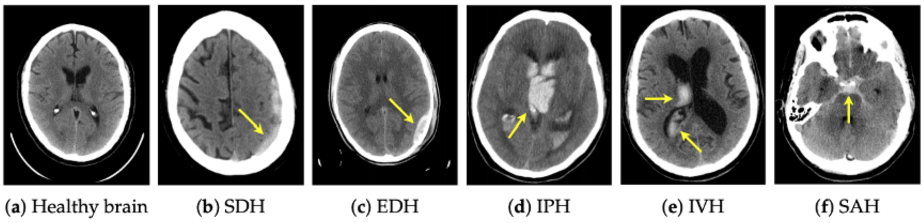

Figure 1: Non-contrast CT sub types of intracranial hemorrhage yellow arrows point to bleeding regions(Sage and Badura; 2020)

Machine learning (ML) improvements in object detection have expanded the usability of ML algorithms such as neural networks for image analysis. Without formal training, deep learning systems discover important components and characteristics inside big databases. This implies that they can be trained from the beginning to the ending. Computerized tomography (CT) scans that often require an expert, a radiologist using machine learning can predict and accomplish. As CT is faster, less expensive, and emits lower radiation, it is better equipped for emergency purposes than MR imaging (Luong et al.; 2020). Convolutional neural networks (CNNs) have risen in popularity mostly as a result of the precision and recall, and they have evolved as essential aspects in the tracking of medical users. In diagnostic imaging, deep learning has made tremendous progress in detecting hidden patterns and features extraction from them intelligently. As a result, it can promote improved data interpretation and oversight, assisting clinicians more effectively (Karthik et al.; 2020).

1.1 Research Question

This study's goal is to propose an alternative: - Is it achievable to use deep learning and transfer learning to diagnose intracranial hemorrhage and its sub types in the brain CT dataset?

1.2 Research Objective

Research Objectives	Description	Evaluation Metrics
First Objective	a critical analysis of this study looks for any previously implemented gap in the literature	-
Second Objective	Data generator for data visualisation and data balancing before the actual data processing	-
Third Objective	CNN model implementation and evaluation	Accuracy, precision recall loss function
Fourth Objective	DenseNet121 model implementation and evaluation	Accuracy, precision recall loss function
Fifth Objective	Xception model implementation and evaluation	Accuracy, precision recall loss function
Sixth Objective	Model Comparison	Accuracy

In this research project, deep convolutional neural network is employed to detect and classify distinct ICH on CT scans. For early identification of ICH and differentiation from other types of hemorrhage can lead to more effective therapy and reduce long-term brain damage as well as fatality. Based on a single head computed tomography (CT) picture, we created a predictive deep learning model using both CNN and Transfer learning networks.

2 Related Work

Over the last few decades, there has been a huge increase in scientific interest in intracranial hemorrhage detection because of advances in image processing. To find and classify intracranial hemorrhage, the researchers are utilizing multiple detection and classification strategies. This article is a concise overview of the deep learning-related research on intracranial hemorrhage. The proposed deep learning techniques for identifying intracranial hemorrhage and its sub types include CNN and transfer learning.

2.1 Deep Learning using CNN

The proposed research(Luong et al.; 2020) utilized deep learning, Hounsfield Unit (HU), and data clustering techniques to manage a classification model to detect brain hemorrhage using CT. Due to the capability of high-performance computing (HPC) technologies, Deep learning approaches to image processing have risen to prominence. The DICOM files within the database include two-dimensional axial representations from the patient's brain. Using deep learning, the ICH classification makes the assumption of the scanning of each image is contains ICH. The authors mentioned the side effects can be identified and reported using various types of conventional processing techniques if the prediction of the test is positive. The system was assessed on two distinct databases by the technologists. The model MobileNetV2 was trained on the RSNA dataset and then used to predict ICH at Vinh Long Province Hospital in Vietnam, with a 99 percent accuracy, 0.992 sensitivity, and 0.807 specificity(Luong et al.; 2020).

The identification of ICH was conducted in a research works (Anupama et al.; 2020) using a blend of image segmentation, picture pre-processing, and the Gabor filtering technique to pre-process the noise. Then came segmentation with Grab Cut algorithm and feature extraction with Synergic Deep Learning. The specificity and the sensitivity of the applied model was 97.78 percent and 94.01 percent. The obtain percentage of precision was 95.79 percent which was similar to the accuracy of 95.73 percent. The researchers advise using comprehensive DL models for image segmentation modification to improve prediction performance.

(Sage and Badura; 2020) employed CNN for feature extraction and multiple classifiers for detecting different subtypes of cerebral bleeding. Because a single CT slice might indicate multiple ICH subtypes, the multiclass technique is sometimes moderately inaccurate. Therefore, all subtypes of ICH were identified individually through separately qualified examples of the architecture suggested. Preprocessing is implemented before feature extraction and classification, including a skull removal algorithm. A ResNet-50 architecture is used to automatically extract classified features. The classification approach also uses SVM and random forest models. The results suggest using a random forest classifier with double-source features. The maximum detection accuracy of 96.7 percent was obtained when Intraventricular Hemorrhage was detected.

Convolutional and LSTM neural networks were utilized in the research (Burduja et al.; 2020) to identify intracranial hemorrhage in 3D CT images. The paper was using the 2019 RSNA Brain CT Hemorrhage database. CNN and LSTM models perform two classification tasks simultaneously. The first stage is to assess whether the CT scan format should be remapped, and the second classification assignment is to identify the subtype of hemorrhage (Burduja et al.; 2020). The researchers discovered that blending ResNeXt and BiLSTM models is much more effective in identifying hemorrhages than using a single ResNeXt model. Thus, they find that extraction of the feature for intracranial hemorrhage detection and sub type segmentation is both productive and convenient.

On the basis of a semi–attention–based multi–task model, the researchers present throughout Paper (Wang et al.; 2020) a revised u-net and curriculum learning approach. Al Hilla Teaching Hospital in Iraq has gathered and identified CT images from 36 ICH cases from the Radiological Society of North America (RSNA) ICH dataset. (Wang et al.; 2020) applied u-net architecture for three experiments: pre-training the encoder with massive identification data, then performing feature extraction on a small number of features using transfer learning. Next, supervised training utilizing only labelled data. Finally, the suggested semi-supervised multi-task focus method tested with large unlabelled and smaller labelled databases. Both approaches resulted in unexpected responses, leading to early training loss. They offered a research technique that would decrease when the difficulty was overcome. Their inability to apply these responses to different settings and data was caused by the unexpected nature of artificial intelligence from limited data sets.

The paper (Guo et al.; 2020) suggest an all-round, end-to-end multi-tasking network for simultaneous classification and intracranial hemorrhage segmentation. It defines a similar encoder for classification and segmentation, and a convolutive long-term storage module (ConvLSTM) enables contextual data gathering. The study utilized 1176 CT images from three hospitals. They proposed utilizing an FCN architecture called ICHNet to integrate classification and segmentation in identifying hemorrhages.For the segmentation and branch classification, a convolutional model with LSTM is considered. The authors showed that ICHNet architecture performs basic classification and segmentation models

On a non-contrast head CT, (Ginat; 2020) proposed deep learning technology to diagnose acute hemorrhage. Applying Aidoc neural network software, researchers discovered that emergency hemorrhage identification accuracy was greater than hospital hemorrhage detection (96.5 percent versus 89.4 percent). Deep learning software output for acute intracranial hemorrhage detection differs by area, according to this paper. The findings may help strengthen AI-driven clinical workflows (Ginat; 2020) .

Another study used the 3605 non-contrast head CT to comprehensively present the results of a deep learning DSS for identifying intracranial hemorrhage (ICH) (Voter et al.; 2021). The researchers observed that, to a certain extent, age, gender, previous neurosurgery, ICH type, and the quantity of ICHs were all associated with diagnostic accuracy. The probabilistic distinctions in image quality had no connection to the diagnostic outcome, and the errors were completely consistent with the sources of false positives assessment of textual analysis.

With an FCN architecture, the research(Imran et al.; 2021) attempted to divide the site to identify hemorrhage and describe the characteristics of ICH detected. Including 82 CT scans, 36 are of ICH patients. Following additional study, the technologies implemented was VGG architecture and RPN(Regional Proposal Network).

Throughout the paper (Wang et al.; 2021), A group of 8898 ICH and regular CT images were fed into three neural networks (VGG-19, DenseNet11, and Resnet-101) that were trained to detect intracranial hemorrhages. From that training set, a particular CT image from the 2019 Kaggle Intracranial Hemorrhages Challenge was selected to detect the intracranial hemorrhage. Grad-CAM was the technique for this research. The map showed the pixels from the graphics uploaded which took the place of the system, making it easier for the system to understand the concept for the particular course and acquire a clear understanding of the diagnostic stage. All three pre-trained models were calculated utilizing training loss curve, speed comparison models, the grad CAM map, the ROC curve and the AUC score. Studies have found that DenseNet-201 has a highest AUC score of 0.92 among the three forms. In the test data, DenseNet-201 performs best(Wang et al.; 2021)

A new deep learning-based brain tumour and intracranial hemorrhage detection approach was proposed from the research (Kirithika et al.; 2020) Firstly, the data is pre processed three stage, with bilateral filtering, contrast limited, and skull stripping, in order to enhance the image resolution. For the segmentation, the invariant feature transform and AlexNet models are utilized, and tumors in the brain and intracranial hemorrhage can be classified employing two different models. Authors applied the DLAN-RF, CNN-VGG16, CART, RF, K-NN, Linear SVM, WEM-DCNN, CNN, and SVM models to determine that the Deep learning AlexNet-Random Forest (DLAN-RF) model was sensitive to 92.41percent, specific to 100 percent, and 94.26 percent accuracy.

The objective of this research (Hebbar et al.; 2020) was to develop a comprehensive solution for detecting brain hemorrhage in a CT image by utilizing convolutional neural networks (CNNs). The data set includes 100 healthy brains and 100 hemorrhage CT images. The authors evaluated CNN models because of its capacity to recognize the spatial and temporal dependencies of a picture using a mixture of several different filter types that enabled them to surpass other learning algorithms in computer vision tasks. (Hebbar et al.; 2020) stated the deep leaning model provided has an accuracy rate of 90 percent for test data and an accuracy rate of 99, 29 percent for validation data. They found that it takes 1.5 seconds to detect a single scan and the effectiveness of the suggested approach was proved.

As mentioned by (Rao et al.; 2021)an FDA-approved Artificial Intelligence (AI) approach to analyse the efficacy of ICH using a convolutional neural network (CNN). (Rane and Warhade; 2021) discusses intracranial acute hemorrhage and its subcategories in detail. Transfer learning is applied to diagnose hemorrhage patterns. The paper includes the RSNA Brain CT Hemorrhage database on DICOM files, and windowing approaches until they are supplied to the system. The effectiveness plot indicates a 0.944 increase over fewer epochs without a window. Precision with windowing (0.956) shows a higher expectation to learn.

Methodologies and studies presented have an outstanding capability to increase deep learning techniques to acquire high accuracy in the classification of brain hemorrhage. In a study (Remedios et al.; 2020), the possibility of deep-learning models to identify bleeding in the brain area was examined.

2.2 Deep Learning using Transfer Learning

For identification and classification of various ICH on unaugmented CT scans, the researchers used a DCNN (Deep convolutional neural network) in the study (Tomasz et al.; 2021). RSNA Data with 752,803 labelled DCM files containing cross-sectional images of the brain is used in the paper. Windowing is employed to extract features, and the target cells' stated window level and window width were used in its implementation. Only components with the required Hounsfield units (HU) are mapped into the three channels known as the brain window, blood window, and bone window. The images are then resized into a 512×512 size. As a feature extractor, the researchers employed the convolutional base model, on top of which they layered a 2-layer fully connected classifier consisting of ReLU and sigmoid activation. They discovered that ImageNet weights are an appropriate initialization weight for the model. Adam Optimizer was used with a learning rate of 1.105 and ten epochs was considered to train the model, which took 48 hours to correspond. The model achieved an accuracy of 93.3 percent per image and a recall of 87.5 percent on average.

Based on a single head computed tomography (CT) picture, the researchers identified a forecasting deep learning model (Ko et al.; 2021). An analysis of 727,392 head CT images revealed 5 different forms of ICH, including intraparenchymal, intraventricular, subarachnoid, subdural, and epidural. To address data imbalance, windowing is first applied to three graphics, then data augmentation. The three graphics are used as input for the training and have been adjusted to 299x299x3. The classification accuracy ranged from 92 to 93 percent. Authors suggested using deep CNNs like Inception and VGG to discover optimal models (Ko et al.; 2021).

To avoid repeated hand-tuning processes (Majumdar et al.; 2021), employed a deep convolutional neural network to learn features and classify. Applying data augmentation to generate the mean output for random left-right flipping and rotation of the input image increases efficiency. On the other hand, post processing improves specificity.134 CT cases including the database. The researchers examined k1,2 and obtain comparable results. 2x2 max-pooling is applied to reduce samples after the first 4 blocks. Then comes 2x2 nearest neighbour expansion, 3x3 convolution, batch normalization, and ReLU optimizer. On the testing data set, 81 percent sensitivity and 98 percent specificity were achieved.

Brain uniformity and Machine Learning were also employed by researchers to identify intracranial haemorrhage, which was integrated into hospital's pre-existing diagnostic process. (Aboutaleb et al.; 2020) developed a deep learning model based on investigations of patients admitted to the Emergency Room with a significant risk of Acute Ischemic Stroke. 568 patients brain CT studies were used to train the system. After that, the model was evaluated against an additional sample. With an accuracy of 81 percent, the model was quite efficient in determining whether there was any severe intracranial hemorrhage.

(Karki et al.; 2020) proposed a novel approach in which a deep convolutional neural network (DCNN) is trained in combination with a CT window estimator module for higher forecasting radiology predictions. Rather than depending on a single window to detect a hemorrhage lesion, graphics are analyzed across several windows. The CNN window estimator network calculates the ideal window settings by each image. The authors provided a set of probable optimal windows and re-examined the data at various windows, blending the outcomes in a comparable cascade approach and another aggregate approach. A window level (WL) and window width (WW) control the picture's brightness and contrast. The training and validation data were 5-fold cross validated. A scaling layer applying different source data to gray scale images using the estimated window values. For classification, the gray scale images are converted into a deep convolutional network. To improve lesion classification by transfer learning, the pretrained models were applied. According to the researcher, a mixture of window width and levels outperformed a single or flexible window.

(Karthik et al.; 2020)found that deep learning models effectively identify and segment brain strokes. Research was aimed at 113 previous academic research papers. The study focuses on enhancements in diagnosing and segmenting stroke symptoms. Stroke lesion diagnosis methods require careful analysis of input image data. The Deep Convolutional Neural Network - Xception method was created to divide ischemic lesion areas better accurately utilizing deep learning technology. The 36-layer included in the deep network extracts precise features. To prevent overfitting, the paper proposes utilizing patch-based training collected from several perspectives can improve the trained network's discriminative ability. The trained network's performance must be evaluated every epoch. Measures such as batch normalization of input data and early stopping can help address this issue.

(Viriyavisuthisakul et al.; n.d.) used InceptionV3 to identify the best window setting for diagnosing Hyperacute and Acute Ischemic Stroke in NCCT (noncontrast cranial computer tomography) images without CTp (computer tomography perfusion). The survey concentrated on hyperacute and acute stages because these require prompt treatment and are the most recoverable. The NCCT radiographs of 49 participants were gathered from Siriraj Hospital in Thailand. Each modified image is obtained to improve the size of the dataset and equalize the quantity of data in the 2 categories. The model is trained using 5-fold cross validation. A hemorrhage classifier is used to define the initial weight. It was trained with Inception V3 and ImageNet dataset. The model's overall accuracy was 82.20 percent, and to reach the best accuracy, the window level and window width had to be manually modified.

A new classifier model VGG-16 was implemented on Kaggle online platform using the keras library. The VGG16 architecture pre-trained on the ImageNet data set is identified by examining the bleeding. For quicker and efficient training and substitutes the negative number of the convolutional with 0, the Relu activation is utilized in all convolution layers. The research also showed that data inequality is a further concern for medical analyses to overcome. The authors (Rane and Warhade; 2021) identified of multi-label learning algorithms have considered to be a potentially feasible method of exploring such classifications on multi-label. In most base deep neural networks, the fine-tuning setting system was often more effective than the fixed extraction setting. With only 3/3 scale convolutional filters, the VGG architecture revealed great capability of a deep CNN. The researcher has assigned the Inception model, which blends maps for characteristics to address scale inconsistencies from different convolutionary filters. As mentioned by (Ho and Kim; 2021), This had an F1-score accuracy of 0.90 with regard to nine popular deep CNNs, as well as a soft precision of 0.97.

3 Methodology

The CRISP-DM methodology is being used for the research project as shown in the below figure 2. CRISP-DM is the market standard and process model for the development of data-mining initiatives. The research indicates that CRISP-DM has been applied in a wide range of data mining projects since 2017 (Schröer et al.; 2021). The methodology of this research is drawn from the findings that were derived from research work. This procedure will be discussed in further detail in subsequent subsections.

Figure 2: CRISP-DM data mining process Diagram

3.1 Business Understanding

Manually diagnosing intracranial hemorrhage from the computed tomography (CT) DICOM images is difficult and the evaluation of hemorrhage using operations is also an expensive procedure. Neurologists and physicians have indicated that identifying the sub type of hemorrhage in a fast way can influence their decisions while assessing emergency patients or detecting some unambiguous results that may be critical in the last stage of treatment. According to a recent assessment by the American Medical Colleges Association, the nation physician availability will become more restricted, estimated to be up to 121000 physicians by 2032 (Hebbar et al.; 2020). So, to decrease the load on doctors we need a strong model that can identify the hemorrhage in its beginning stages, the goal of the project is to develop a best deep learning approach that assists physicians in diagnosing the intracranial hemorrhage and its sub types that that helps doctors diagnose and raise the chance of patient survival. This also helps them detect bleeding at their early stages. As previously indicated in the literature review, deep learning is extremely effective at identifying intracranial hemorrhage.

3.2 Data Understanding

The data source used for this research is the Radiological Society of North America (RSNA) data set (Kaggle; 2021) , which is openly accessible in Kaggle using the Kaggle API and does not contain any clinical information about patients, making it appropriate for use in this study. Because it does not contain any medical information about a patient, so using it for analysis is ethical. The dataset contains 762131 DICOM-format images of various subtypes of intracranial hemorrhage. The training data is given as a number of image IDs and numerous label, with each of five hemorrhage sub types, including an additional label, which must always be valid, if any of the subtypes label is valid. Every file is marked for one or more types of hemorrhage or no hemorrhage at all. The Figure 3 below shows that the majority of the scans show no signs of hemorrhage. Label were provided on the basis of a CSV file with six lines for each patient ID, relating to a kind of bleeding and a Boolean value to show if the form of hemorrhage is present in the picture. The CT scans in DICOM format take up a significant amount of system memory. All DICOM images has been converted to PNG format in order to make sure that no pixel quality is lost during the conversion process. This is addressed in better detail during the data processing processes. The data sets available are divided into three categories: training data sets, testing data sets, and validation data-set.

type	filename	any	epidural	intraparenchymal	intraventricular	subarachnoid	subdural
0	ID_000457398.png	0	0	0	0	0	0
1	ID_0004a5701.png	1	0	0	0	0	1
2	ID_0006058f6.png	0	0	0	0	0	0
3	ID_0006d4f2c.png	0	0	0	0	0	0
4	ID_000908614.png	0	0	0	0	0	0

Figure 3: of the data frame where no hemorrhage in the image

3.3 Data Pre-processing

X-rays are being used to perform computed tomography (CT) scans. The higher the pixel intensity, the more X-rays are attenuated by the skin cells (Tomasz et al.; 2021). Pre-processing is the early phase in removing the images blurriness. It assists in identifying the object of interest among the various images. There are a variety of obstacles associated with preparing DICOM images for pre-processing; trouble observing and detecting diseases is common in these images due to interference, blurriness, and a lack of contrasts. The image format also differs in terms of brightness, contrast, sharpness, and visibility of structures depending on the specifications and manufacturer of the medical imaging device such as ultrasound, MRI etc. The large amount of storage space that DICOM images take up in any system must be dealt with by first transforming them into a layout that is medically reliable and preserving the data. Often, radiologists look at many types of picture formats, including JPEG, PNG, TIFF, and GIF. As a result, the first and most important step is to convert the image to a windowed state.

3.3.1 Image conversion:

The PNG pictures are great for computers, delivering device compatibility and picture resolution. When contrasted to the .dcm file extension, it is far less in size. The PNG images don't lose any crucial information while using encoding. The sample of DICOM images included in the data set have a resolution of 512x512 pixels. DICOM images typically contain 12–16 bits per pixel, which correlates to 4,096 to 65,536 shades of Gray (Viriyavisuthisakul et al.; n.d.). However, most computing devices are restricted to 8 bits or 256 shades of Gray. 50000 training image and 5000 test DICOM image is selected for the conversation of DICOM (.dcm) file to a PNG file to ensure there were no complications in the design stage.

3.3.2 Windowing of Images:

The goal of using windowing is to condense the 256 shades of grey into a small number of Hounsfield units(HU) that represent the essential thicknesses of the tissue that we are examining while diagnosing. This helps users to easily see the diverse features of tissues by concentrating on a particular area and emphasizing minor distinctions between cells. The paper (Majumdar et al.; 2021) suggested linear transformation is required for DICOM images. Following the removal of the HU, the appropriate windowing measures can be applied to improve visibility of any potentially aberrant locations. After intensity windowing, images are standardized to a certain range. The paper (Rane and Warhade; 2021) discovered that employing windowing of the images resulted in a significant increase in precision. To compute the upper and lower grey level, we employed window level (WL) and window width (WW) to create different types of windows depending on the type of hemorrhage. Window width refers to the range of CT numbers that an image comprises while Window width refers to the range of CT numbers that an image comprises; window center refers to the midway of the range of CT values shown; as the window level is reduced, the CT picture becomes brighter, and inversely . Windowing optimizes the slight variation between features for feature extraction. Only tissues with the required HU are mapped into the three channels of the input tensor to use the defined range of WL and WW for the desired tissues as shown in the figure 4 below. The data for window width and window level are extracted from DICOM tags with the help of the pydicom package, which is highly effective for processing.dcm format images. We have fixed the outliers in 16-bit DICOM files before constructing our windows. The image pixels are then clipped between the highest and lowest observable levels, allowing us to focus just on a small area where the irregularity may be visible. Which indicates that any pixel value larger than (WL+WW/2) will seem white, while any value less than (WL-WW/2) will appear black. The bone window and the brain window are two main window settings for head CT. The complexities of soft tissues like the brain, which have a lesser density than bones, are lost in the bone window settings. (Jones; 2021) Subdural WL=80 and WW=200 and brain matter WL=40 and WW=80 windows were chosen because we were seeking bleeding inside the skull. Despite the fact where both give details about soft tissues, their HU ranges are distinct. For Bone WL=40 and WW=380 was chosen as our third feature to identify abnormalities involving the skull.

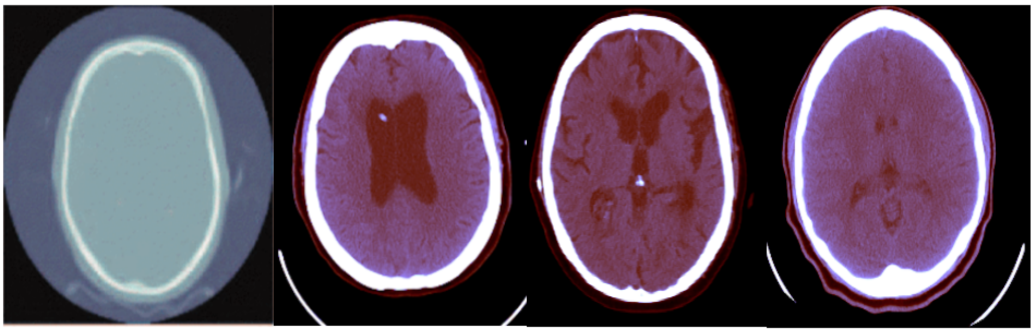

Figure 4: pre-processing sequence : scan without contrasts, Brain window, subdural window and one window

3.3.3 Resizing of the Data:

The sample of DICOM pictures contained in the data set has a resolution of 512x512 pixels, which is the highest possible resolution. At this point, we attempted to achieve a balance in the number of each single sub type of ICH images. At the end, we downsized the three images to a size of dimensions 128x128 pixels, which would be used as input for the training process. From OpenCV, the cv2 package is used to resize and process the image.

3.4 Modelling

Three distinct models are utilized in this step to utilize the pre-processed data. The novelty of this study is it employs a pre-trained CNN model (Transfer Learning) to detect intracranial bleeding and its sub types. Transfer learning is applied to the processed data using pre-trained models such as DenseNet121 and Xception. The figure 5 below shows the design of process flow.

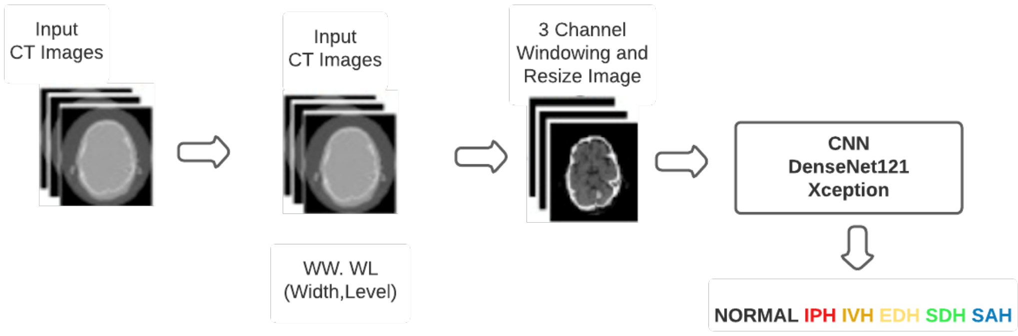

Figure 5: Design Process Flow

3.4.1 CNN (Convolutional Neural Network):

A CNN, is a type of deep neural network design that is particularly well suited to interpreting image database features. CNN based designs have proven to be the most effective in visual data analysis when compared to other deep learning models (Sage and Badura; 2020) (Ho and Kim; 2021).(Rao et al.; 2021) (Kirithika et al.; 2020) The researcher (Wang et al.; 2020) mentioned CNN architecture outperformed most machine learning classifiers, including Gaussian Naive Bayes, Generalized Linear Models, Gradient Boosting Classifier, AdaBoost, Random Decision Forests, and Extra Trees Forest methods. CNN uses filtering, also known as a set of kernel, to learn the shared features by combining the picture across several axes or channels depending on multiple classifiers such as padding, stride and so on. The model learns to generalize picture features from simple to complex through a sequence of deep convolutional layers. To simplify the set of features and data flow, CNN employs a pooling layer among the convolutional layers as shown in the below figure 6. The max pooling techniques examine a matrix of images pixel, usually 2x2 and take into account the region's greatest activation value. This procedure assures that a scalar outcome detects the occurrence of a feature. The number of neurons in the end resulting layer is the same as the number of classes in the dense layer. In the following part, we'll go through the specifics of how to use a CNN model as an image features generator.

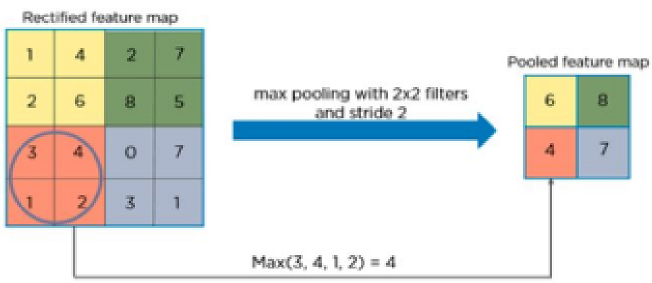

Figure 6: CNN Maxpooling operation

3.4.2 Pre-trained models (Transfer Learning):

The lack of information obtained from the given dataset could be remedied by employing the methods of Transfer learning, in which information gathered through one task is applied to another. Transfer learning is studied in the research work (Wang et al.; 2020) for picture segmentation on a smaller dataset. The researchers identified more accurate findings can be obtained by increasing the number of layers and building a thick structure. Transfer learning was applied to predict ICH subtypes with good results using an ensemble of four well-known CNN algorithms (Anupama et al.; 2020).Transfer Learning is a form of deep learning that has been successfully applied to picture categorization, visual classification, and object prediction studies (Kirithika et al.; 2020).A transfer learning strategy is used in a similar work (Rane and Warhade; 2021) to recognize the bleeding types. Transfer learning appears to perform effectively with medical picture data, according to a research(Li et al.; 2020).

- DenseNet121 – The advantage of DenseNet is the improved rebuilding of features, which makes the design incredibly efficient. DenseNet has the best AUC score in the study and the best results (Wang et al.; 2021). DenseNet has been steadily improving in terms of accuracy while simultaneously improving in terms of features without introducing overfitting and performance loss. Another feature of DenseNet is that it can save significant amounts of time by increasing the number of features that can be reused in future projects. Because of these changes, DenseNet was evaluated in this study. DenseNet121 was trained using ImageNet weight prior to its use in research. DenseNet121 networks have been chosen because it is deeper, denser, and more accurate.

- Xception Model – (Karthik et al.; 2020) (Ko et al.; 2021)The Xception model is a CNN architecture based on deeply convolutional layers. Xception is stated as the assumption for the initial module that produces cross channel correlations and space interaction that can be totally disconnected within the CNN characteristic mapping. This concept is regarded as a more advanced variant of the Inception model. The Xception architecture has 71 layers, which serve as the framework for extracting features. The layers below are divided into 14 segments. These are the depth layers of distinct convolution layer that, apart from the first and final module, are layered linearly with the residual batch normalization. Xception also uses a technique based on the three different flows, in which the usable data passes first through the input flow, then through the middle flow, which is replicated 8 times and then it goes through the output flow. This flexible aspect of design simplifies the definition and modification with the help of a range of sophisticated libraries such as TensorFlow and Keras.

4 Proposed Design Architecture

As illustrated in the diagram below, a three-stage framework (Choudhury; 2020) be adopted for this study, which consists of a database layer, an application layer, and a presentation layer. The first stage of the research project is to use the Kaggle API to gather all of the relevant data. Additionally, DICOM data is translated to PNG format through image conversion, windowing, and resizing at the business logic layer. After that execute EDA (exploratory data analysis) to find out how widely distributed the data set is. Deep learning and transfer learning methods are then used to the data set. The final level is the presentation layer, which uses Python libraries to visualize the presence or absence of an intracranial hemorrhage subtypes.

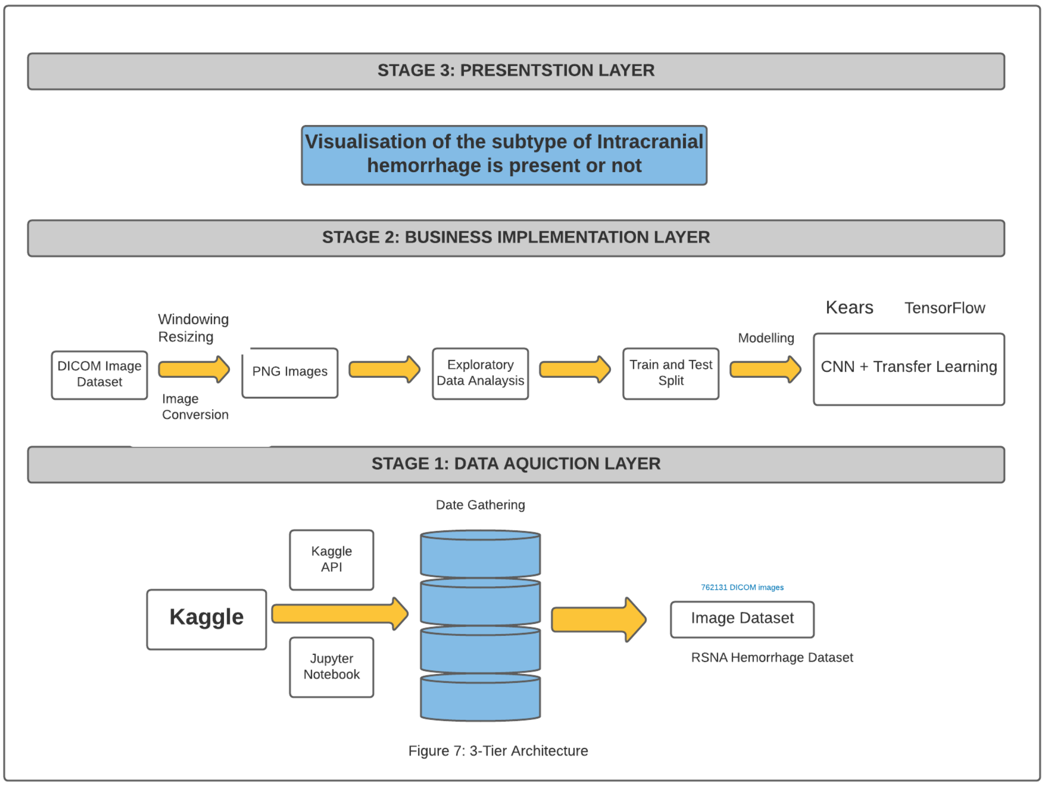

Figure 7: Design Architecture for the proposed research

5 Implementation

In this section, we describe the general implementation of the study results. The initial setup, data conversion stages, and model development are all covered in this section, as well as the tools that were utilized to complete the task.

5.1 Environment Setup

The Windows 10 Home operating system with a 64-bit operating system and 8 GB of RAM was used in the implementation. Data storage is accomplished through the use of an external 500GB hard disk.

- Python packages – Python is a powerful and user-friendly language. Several IDEs support this language. It also has a large network and tools for machine learning and other data analytics algorithms. This prompted the usage of Python for the tool's implementation. The system employed during the processing included Python libraries such as machine learning, data cleansing, data conversion and so on.

- Anaconda Jupyter Notebook – This is an IDE that operates with the least amount of resources possible. It is compatible with the latest edition of Python.

5.2 Data Transformation

Data transformation is the process of extracting data in accordance with the model's requirements. The entire data set, which includes six unique hemorrhage sub types, was retrieved using the Kaggle API. Because the data is compressed into a zip file, we must first unzip the folder before proceeding. Image conversion is implemented because DICOM pictures take up too much storage space. PNG images have a superior image quality due to the lossless compression. Due to the limited hardware available for implementation, modelling is performed on 50000 random images selected from the training data set file and 5000 random images from the test dataset file. In both the train and test data sets, duplicate files have been dropped. to validate the model results the article (Burduja et al.; 2020) (Hebbar et al.; 2020) employed techniques for separating training and testing data .To validate the results, 15 percent of the training data is used. Following that, the datasets are fed into image conversion.The research work (Wang et al.; 2021) used a python library called 'pydicom' to extract the data from the DICOM files. The "pydicom" package in Python handles the image conversion. The article (Ko et al.; 2021), (Luong et al.; 2020) and (Rane and Warhade; 2021) used windowing for feature extraction and present distinct images. The values in each window are normalized to (0, 1). With CT pictures, three alternative windowing techniques are used. After windowing, the images are resized to the size of (128x128) using the "cv2" library and saved in a distinct folder, which will be used as an input for the training in the process stage. From Keras, the "ImageDataGenerator" package is used to resize the images. The resized data file directory's main path was assigned in the data generator function, and the x column was assigned to the images column and the y column to the remaining Boolean columns. The image's target size was set to (128x128) and the batch size set to 32. In Python, deep learning models were created using the Keras neural network library, which is built on top of the TensorFlow framework.

5.3 Convolutional Neural Networks (CNN)

Convolutional layers with ReLU activation, pooling layers, fully connected layers, and a loss layer make up the architecture of the CNN model. Pooling layers and a convolutional layer (Conv2D with kernel size 3x3) with ReLU activation (set the negative value to 0) are incorporated after the Sequential model is the initialization process. The convolutional layer generates a feature map that outlines the input's distinct features (Burduja et al.; 2020). It is increased by modifying a filter to the layer's input. This method compresses the image while also increasing the feature map's depth. (Ko et al.; 2021) used the (299, 299, 3) structure in their research because the dimensions of the input image were large (299 x 299). The source images structure of this research is as follows: (128, 128, 3) because the image dimensions is (128 x 128). To effectively decrease the network's special feature sizes, the MaxPool2D pooling layer with pool size = 2 (Hebbar et al.; 2020) has been layered between convolutional layers. The model's MaxPool2D layer reduces data size by pooling only the most important characteristics. This combination is then repeated before implementing a flatten layer, which converts the input into a one-dimensional array. For the flattened layer result, a fully connected layer with two hides of a dense layer and a hidden, dense, sigmoid-activated layer (Hebbar et al.; 2020) (Imran et al.; 2021) on 6 neural (Ko et al.; 2021) is provided.

5.4 Transfer Learning based DenseNet121 and Xception

5.4.1 DenseNet121

In the research, we used a deeper, denser, and more accurate DenseNet121 network. (Burduja et al.; 2020) discovered that CNNs trained on ImageNet can be successfully used to diagnose ICH and (Ho and Kim; 2021) suggested that for feature extraction, pre-trained ImageNet models could be used. The first step is to initialize the model as transfer learning via the Keras application by downloading the ImageNet database. Once the model has processed the data and identified the input as an image, we have supplied the model to the input shape of the image as (128x128). The max pooling layer as well as Global Average pooling are taken into consideration while reducing the size of feature maps. With the 6 Dense layer of prediction, a sigmoid activation function was used, and the Adam optimizer (Burduja et al.; 2020) (Guo et al.; 2020) (Ko et al.; 2021) was used with a learning rate of 0.0001 (Luong et al.; 2020) to achieve the best results.

5.4.2 Xception Model

The Xception model from the Keras application was utilized in conjunction with the Xception function, which utilizes ImageNet pre-trained weights. This model requires an input image of size 128*128 and so requires that any input images that are not of this size be scaled before being provided as input. We used the global average pooling 2D mode on the last convolutional block to extract features. Dropout is adjusted to 0.15 to avoid overfitting. A sigmoid activation function was utilized. The Adam optimizer (Burduja et al.; 2020) (Guo et al.; 2020) was used with a learning rate of 0.0005. For callbacks with a defined decay argument, a learning rate reduction is also implemented; it reduces the learning rate from the previous epoch by the supplied fixed amount. The Batch size (Ho and Kim; 2021) is set to 32 and the epochs is set to 10 and model.fit() was utilized to execute the model.

6 Evaluation

There are numerous approaches for determining the success rate of any classifications or predictions model. The assessment techniques used are based on the information contained in the confusion matrix. (Viriyavisuthisakul et al.; n.d.) employed a confusion matrix to determine the actual and predicted window levels and widths. The confusion matrix data is used to determine the success of our created approach. True positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) comprise the confusion matrix.

6.1 Accuracy and Loss Variation

(Rane and Warhade; 2021) evaluated the performance of the learning curve of a full range image with and without windowing using training and validation accuracy plots. In the study (Hebbar et al.; 2020), the accuracy of testing and training was used to compare two deep learning models. The graphs illustrated below show the testing and training dataset accuracy and loss variation for all the models that were implemented in this research.

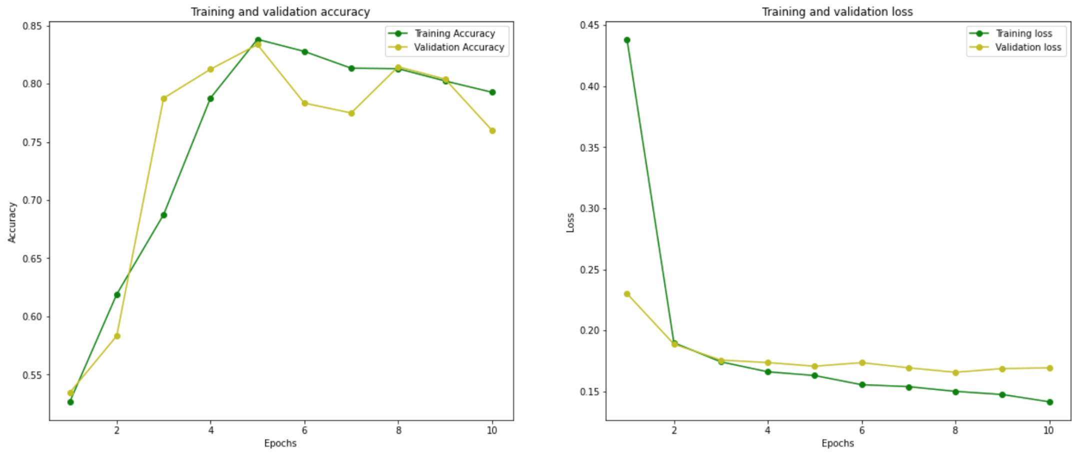

Figure 8: CNN Accuracy and Loss Graph

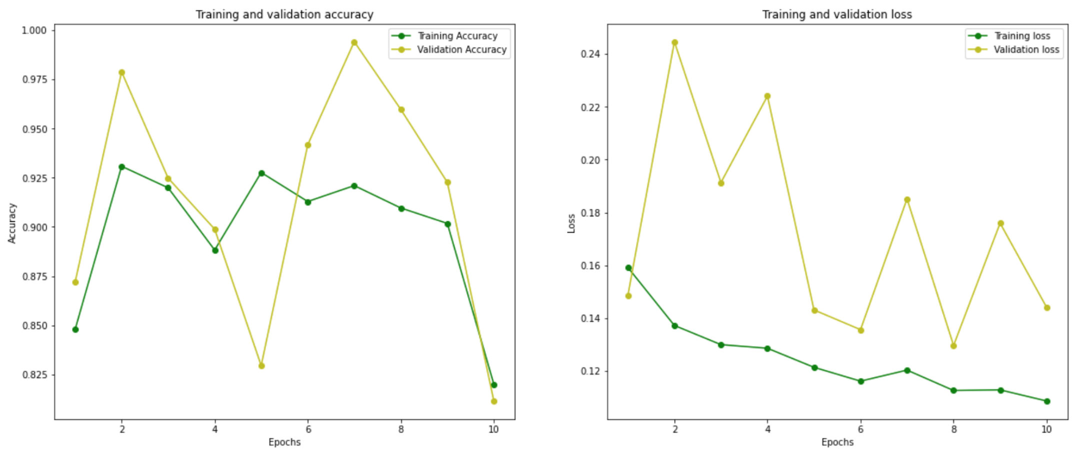

Figure 9: DenseNet121 Accuracy and Loss Graph

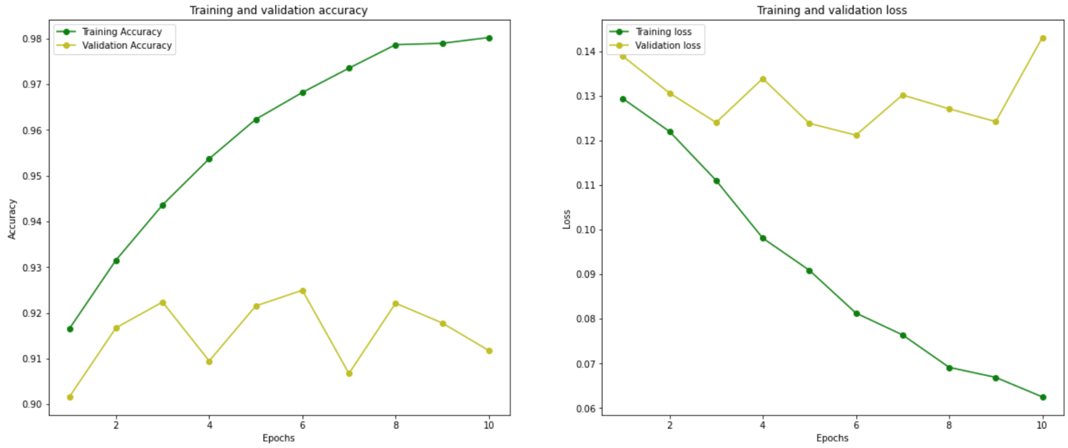

Figure 10: Xception Accuracy and Loss Graph

Figure 8, Figure 9, and Figure 10 exhibit respective model's overfitting with regard to curve fit. Also, in the model CNN, there is a little reduction in overfitting since the curves are positioned further apart from each other than of another. The paper (Luong et al.; 2020) utilized the accuracy of training and validation to balance the sensitivity and specificity values. If the model's accuracy during training continues to improve, the model's specificity will increase. However, the sensitivity will decrease significantly. Using CNN results in a lower loss of both training and validation, as well as a smaller improvement in training and validation accuracy. As the number of iterations for training dataset increases, the loss for both training and validation datasets decreases. Though there are minimal variations in terms of the correctness of the validation dataset.

6.2 Metrics obtained from confusion matrix

- Accuracy of a model is calculated by calculating the overall number of instances that are accurately predicted, as well as the total number of instances that are forecasted. This displays the model's efficiency. The accuracy and precision of the model was evaluated in the research work (Burduja et al.; 2020) (Ginat; 2020) (Ho and Kim; 2021) (Guo et al.; 2020).

- Precision measures how much of the images we've categorised as having hemorrhage and without hemorrhage represent actual bleeding.

Model	Accuracy	Precision	F1- Score	Recall
CNN	86	0.89	0.92	0.96
DenseNet	90	0.95	0.94	0.94
Xception	91	0.95	0.95	0.95

From the above table, we can see Xception model performed well than other two model.

6.3 Actual vs Predicted Analysis

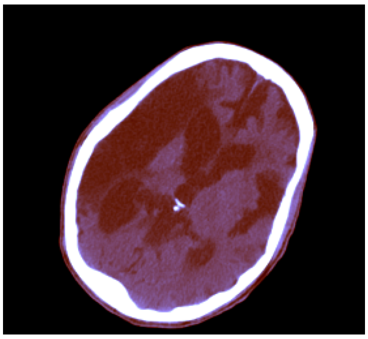

Figure 11: Actual Image

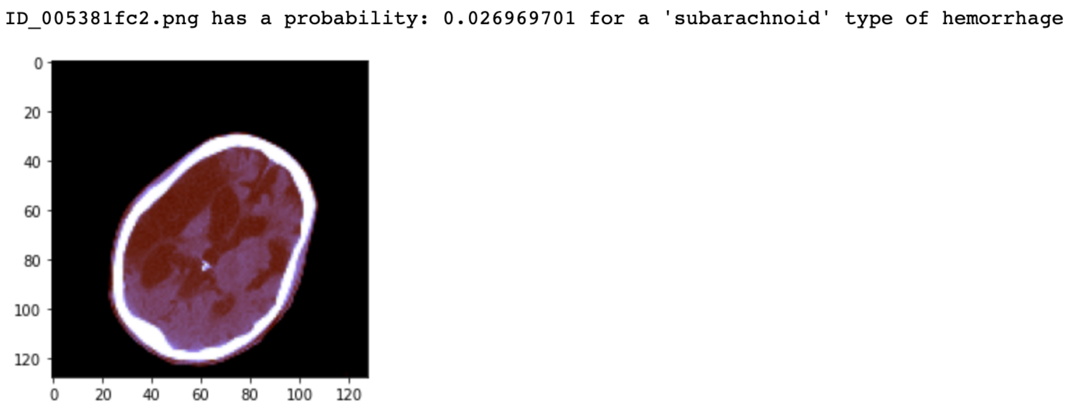

Figure 12: Predict Image

As we can see in the above predictive image the model works well and detects the subarachnoid hemorrhage on a lower probability. The Xception model is capable of identifying the subtype of intracranial hemorrhage with a decreased risk of failure.

7 Conclusion and Future Work

We developed a framework which is capable of analysing single frame CT scans with superior performance to a radiologist and able to identify and classify ICH as intraparenchymal, intraventricular, subarachnoid, subdural, or epidural. Our model was trained on the RSNA dataset(Kaggle; 2021), a massive labelled dataset consisting of 762131 DICOM format CT images. The CT images are transformed to 16 bit DICOM pictures with vectors of floating-point integers standardized in the (0, 1) range using windowing. This study makes use of only 55000 scaled PNG format images in the (128x128) dimensional format. We used the CNN and transfer learning model (Xception model and DenseNet121) and our proposed approach improves ICH detection and classification with an accuracy of 91 percent and can help physicians in the analysis of head CT images, The Xception model is capable of identifying the subtypes of intracranial hemorrhage with a decreased risk of failure. For future study, the model has a very good potential of improving accuracy when applied to a big image collection.

References

Aboutaleb, P., Barman, A., Lopez-Rivera, V., Lee, S., Vahidy, F., Fan, J., Savitz, S., Giancardo, L. and Sheth, S. A. (2020). Abstract wp405: Automated detection of hemorrhagic stroke from non-contrast computed tomography: A machine learning approach, *Stroke* **51**(Suppl$_1$).

Anupama, C. S. S., Sivaram, M., Lydia, E. L., Gupta, D. and Shankar, K. (2020). Synergic deep learning model–based automated detection and classification of brain intracranial hemorrhage images in wearable networks, *Personal and Ubiquitous Computing* .

Burduja, M., Ionescu, R. T. and Verga, N. (2020). Accurate and efficient intracranial hemorrhage detection and subtype classification in 3d ct scans with convolutional and long short-term memory neural networks, *Sensors* **20**(19): 5611.

Choudhury, M. D. (2020). *Automated Identification of Painters Over WikiArt Image Data Using Machine Learning Algorithms*, PhD thesis, Dublin, National College of Ireland.

Ginat, D. T. (2020). Analysis of head ct scans flagged by deep learning software for acute intracranial hemorrhage, *Neuroradiology* **62**(3): 335–340.

Guo, D., Wei, H., Zhao, P., Pan, Y., Yang, H.-Y., Wang, X., Bai, J., Cao, K., Song, Q., Xia, J. et al. (2020). Simultaneous classification and segmentation of intracranial hemorrhage using a fully convolutional neural network, *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, IEEE, pp. 118–121.

Hebbar, N., Patil, H. Y. and Agarwal, K. (2020). Web powered ct scan diagnosis for brain hemorrhage using deep learning, *2020 IEEE 4th Conference on Information & Communication Technology (CICT)*, IEEE, pp. 1–5.

Ho, N. and Kim, Y.-C. (2021). Evaluation of transfer learning in deep convolutional neural network models for cardiac short axis slice classification, *Scientific reports* **11**(1): 1–11.

Imran, R., Hassan, N., Tariq, R., Amjad, L. and Wali, A. (2021). Intracranial brain haemorrhage segmentation and classification, *iKSP Journal of Computer Science and Engineering* **1**(2): 52–56.

Jones, J. (2021). Ct head (an approach) — radiology reference article — radiopaedia.org.
URL: *https://radiopaedia.org/articles/ct-head-an-approach?lang=gb*

Kaggle (2021).
URL: *https://www.kaggle.com/c/rsna-intracranial-hemorrhage-detection/data*

Karki, M., Cho, J., Lee, E., Hahm, M.-H., Yoon, S.-Y., Kim, M., Ahn, J.-Y., Son, J., Park, S.-H. and Kim, K.-H. e. a. (2020). Ct window trainable neural network for improving intracranial hemorrhage detection by combining multiple settings, *Artificial Intelligence in Medicine* **106**: 101850.

Karthik, R., Menaka, R., Johnson, A. and Anand, S. (2020). Neuroimaging and deep learning for brain stroke detection - a review of recent advancements and future prospects, *Computer Methods and Programs in Biomedicine* **197**: 105728.

Kirithika, R. A., Sathiya, S., Balasubramanian, M. and Sivaraj, P. (2020). Brain tumor and intracranial haemorrhage feature extraction and classification using conventional and deep learning methods, *European Journal of Molecular & Clinical Medicine* **7**(7): 237–258.

Ko, H., Chung, H., Lee, H., and Lee, J. (2021). Feasible study on intracranial hemorrhage detection and classification using a cnn-lstm network.
URL: *https://ieeexplore.ieee.org/document/9176162*

Li, X., Yang, H., Lin, Z. and Krishnaswamy, P. (2020). Transfer learning with joint optimization for label-efficient medical image anomaly detection, *Interpretable and Annotation-Efficient Learning for Medical Image Computing* pp. 146–154.

Luong, K. G., Duong, H. N., Van, C. M., Thi, T. H. H., Nguyen, T. T., Thoai, N. and Thi, T. T. T. T. (2020). A computer-aided detection to intracranial hemorrhage by using deep learning: A case study, *Advances in Intelligent Systems and Computing* pp. 27–38.

Majumdar, A., Brattain, L., Telfer, B., Farris, C. and Scalera, J. (2021). Detecting intracranial hemorrhage with deep learning.
URL: *https://ieeexplore.ieee.org/document/8512336*

Rane, H. and Warhade, K. (2021). A survey on deep learning for intracranial hemorrhage detection, *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)*, IEEE, pp. 38–42.

Rao, B., Zohrabian, V., Cedeno, P., Saha, A., Pahade, J. and Davis, M. A. (2021). Utility of artificial intelligence tool as a prospective radiology peer reviewer—detection of unreported intracranial hemorrhage, *Academic radiology* **28**(1): 85–93.

Remedios, S., Wu, Z., Bermudez, C., Kerley, C. I., Roy, S., Patel, M. B., Butman, J. A., Landman, B. A. and Pham, D. L. (2020). Extracting 2d weak labels from volume labels using multiple instance learning in ct hemorrhage detection, *Medical Imaging 2020: Image Processing*, Vol. 11313, International Society for Optics and Photonics, p. 113130F.

Sage, A. and Badura, P. (2020). Intracranial hemorrhage detection in head ct using double-branch convolutional neural network, support vector machine, and random forest, *Applied Sciences* **10**(21): 7577.

Schröer, C., Kruse, F. and Gómez, J. M. (2021). A systematic literature review on applying crisp-dm process model, *Procedia Computer Science* **181**: 526–534.

Tomasz, L., Kumar, M., Hong, R. and Wu, W. (2021). Intracranial hemorrhage detection in ct scans using deep learning.
URL: *https://ieeexplore.ieee.org/document/9179504*

Viriyavisuthisakul, S., Kaothanthong, N., Sanguansat, P., Haruechaiyasak, C., Le Nguyen, M., Sarampakhul, S., Chansumpao, T. and Songsaeng, D. (n.d.). Evaluation of window parameters of noncontrast cranial ct brain images for hyperacute and acute ischemic stroke classification with deep learning.

Voter, A. F., Meram, E., Garrett, J. W. and John-Paul, J. Y. (2021). Diagnostic accuracy and failure mode analysis of a deep learning algorithm for the detection of intracranial hemorrhage, *Journal of the American College of Radiology* .

Wang, J. L., Farooq, H., Zhuang, H. and Ibrahim, A. K. (2020). Segmentation of intracranial hemorrhage using semi-supervised multi-task attention-based u-net, *Applied Sciences* **10**(9): 3297.

Wang, Z., Wu, L. and Ji, X. (2021). An interpretable deep learning system for automatic intracranial hemorrhage diagnosis with ct image, *Proceedings of the 2021 International Conference on Bioinformatics and Intelligent Computing*, pp. 338–357.