Medical Image Analysis (2023)

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Multi-dimensional feature fusion network for stroke infarct segmentation from non-contrast CT

Given-name1 Surname1[a,*], Given-name2 Surname2[a,1], Given-name3 Surname3[1], Given-name4 Surname4[1]

[a]*Dept. of Computer Science, Sri Sivasubramaniya Nadar College of Engineering, Chennai, India*

ARTICLE INFO

Article history:

2000 MSC: 41A05, 41A10, 65D05, 65D17

Keywords: Keyword1, Keyword2, Keyword3

ABSTRACT

Brain stroke is among the most common reason for death around the world. Brain imaging methods like magnetic resonance imaging (MRI), diffuse optical imaging, functional MRIs, and CT are quite helpful for initial screening. However, sophisticated imaging techniques like MRI require high operating cost and a well-trained operator. Often times, CT is the most convenient and quick imaging modality. Toward the diagnosis of strokes, physical and manual examination of the patient is performed to determine an appropriate course of treatment, aided by brain images. This process is made complex by the contrasting and divergent treatments required for the two kinds of stroke – ischemic and hemorrhagic – requiring accurate lesion type identification. Furthermore, an accurate segmentation of the infarct location is particularly useful for subsequent treatment. However, manually performing this diagnosis is extremely time-consuming for the already overburdened radiologist, and is prone to human-error. Thus, there is a need for an effective automated system for diagnosing strokes that can help physicians quickly start treatment after stroke onset. Hence, in this work, we proposed to develop a deep learning -based computation pipeline to process CT slices to detect stroke type and localize infarct location. Furthermore, we propose to reconstruct the CT slices into volumetric models while correcting errors in infarct localization, and potentially draw out diagnostic measurements that could help with subsequent treatment. Finally, we also propose a path planning algorithm to treat strokes using a six degree of freedom (DOF) arm. This Q-learning based method is simulated qualitatively on benchmark virtual environments.

1. Introduction

After heart disease, brain stroke is the most common reason for death around the world Members et al. (2012). A majority of survivors need to live with changeless or long-term injury. Brain imaging methods like magnetic resonance imaging (MRI) and CT are quite helpful for a doctor in order to start the initial screening of the patient. There are also many imaging modalities for the analysis of brain, which may include X-ray imaging, diffuse optical imaging, magnetoencephalography, functional MRI and positron emission tomography Durduran et al. (2004); Zhang et al. (2005). However, all these imaging techniques require high operating cost and well-trained operator, hence most of these imaging methods may not be available in all the clinics and hospitals.

Image classification is widely used in medical imaging Pereira et al. (2020). However, for better accuracy of the classification system, results should be close to the manual diagnosis.

*Corresponding author: Tel.: +0-000-000-0000; fax: +0-000-000-0000;
e-mail: author3@author.com (Given-name3 Surname3)
[1]This is author footnote for second author.

Nowadays, deep learning has been extensively used as a classification method because it automatically calculates features within the convolutional layers of the deep system Mousavi et al. (2019); Başaran et al. (2020). The main advantage of using deep learning is that it outperforms other conventional methods for image classification Suk et al. (2014). Many deep learning methods have come into existence such as recurrent neural networks, long short-term memory (LSTM) Hochreiter and Schmidhuber (1997), CNNs Ouhmich et al. (2019), deep belief nets (DBN) Hinton et al. (2006), etc. Among these methods, CNN has been generally utilized in computer vision and medical image processing problems like ImageNet, face recognition, house numbers digit classification, patch classification from medical images etc. Classification methods other than deep learning methods are random forest (RF), -nearest neighbors (NN), decision tree (DT), multilayer perceptron (MLP), support vector machine (SVM) and many more Wan et al. (2012); Subudhi et al. (2020); Toğaçar et al. (2020).

In this research, an automatic classification method has been proposed to predict the category to which brain CT scan image belongs, along with the locations of the stroke infarcts for hemorrhagic and ischemic stroke types, and normal tissue. *Hemorrhagic stroke* occurs due to fragile blood vessel which burst and drains into the neighboring brain tissues. On the other side, *ischemic stroke* occurs when blood supply in the brain stops due to the presence of blood clots. Brain hemorrhage can also occur after ischemic stroke which results in a serious complication Matesin et al. (2001). Sample images of the two types of stroke on CT slices are depicted in Figure **??**, with infarct regions demarcated in red.

Toward the diagnosis of stroke, doctors first confirm symptoms from the patient or family members. The most important thing which will help in accurate identification is the stroke history. This is usually followed by neuroimaging of the patient. The most commonly used neuroimaging method used by experts is CT scan for diagnosing brain strokes, on account of their greater affordability. The prediction of stroke from CT scan images serve as the initial step towards the proper diagnosis of a patient. These images are then sent to a cardiovascular radiologist to identify the stroke type. Thereafter, physical and manual examination of the patient is performed to determine an appropriate course of treatment. However, this manual diagnosis procedure is extremely time-consuming for the already overburdened radiologist, and is prone to human-error. Many researchers have worked in this area for providing a computer-aided diagnosis (CAD) system, in spite of the significant requirement, there is no clinically accepted CAD system for stroke Nowinski et al. (2014); Peixoto and Rebouças Filho (2018). The system proposed by Nowinski et al. (2014) was not automated and requires significant performance improvements to warrant automation. Peixoto and Filho Peixoto and Rebouças Filho (2018) have worked on the classification of hemorrhagic and ischemic stroke CT scan images. Their method was developed for small dataset where each case had 100 images and for future work they have to test their method on large datasets. Thus, there is a need for an effective automated system for diagnosing strokes that can help physicians quickly start treatment after stroke onset. Hence, in this work, we proposed to develop a deep learning -based computation pipeline to process CT slices to detect stroke type and localize infarct location. Furthermore, we propose to reconstruct the CT slices into volumetric models while correcting errors in infarct localization, and cast the volume into a computationally pliable 3D model that can be used to draw out diagnostic measurements as well as for treatment simulations.

2. Related Work

In general, processing and analysis of medical images using computer-aided diagnosis (CAD) systems can help physicians in early identification of diseases Nowinski et al. (2014); Gupta and Bhavsar (2018); Jain and Salau (2019); Wu et al. (2012). Herein, the proposal briefly describe some important image preprocessing and classification methods used so far in order to analyze brain strokes from medical images.

Saatman et al. Saatman et al. (2008) have discussed various classification techniques for traumatic brain injury so that proper treatment can be provided for that targeted injury. They discussed classification based on the severity of brain injury where various parameters which include Glasgow Coma Scale (GCS) have been used to identify the severity, however these procedures require many clinical trials. Other techniques which they have discussed are pathoanatomic classification, classification by physical mechanism, pathophysiology and prognostic modeling. According to them, for avoiding several clinical trails of patients new tools can be developed like dataset development, data sharing, data mining, and bioinformatics, etc. In 2008, Mozqueda et al. Torres-Mozqueda et al. (2008) have tested the hypothesis on clinical data of patients that acute ischemic stroke classification from CT or MR angiography images into major and minor stroke can be done by using Boston Acute Stroke Imaging Scale (BASIS) instrument. They have proved that BASIS classification instrument is better than Alberta Stroke Program Early CT Score (ASPECTS) in predicting stroke.

Chawla et al. Chawla et al. (2009) presented an automatic method for the detection and classification of an abnormality from low-contrast CT images into three categories — hemorrhage, chronic and acute infarct. Their proposed technique comprises of three steps: enhancement of the image, recognition of mid-line symmetry and classification of abnormal slices. The main limitation of their work is that they have just shown results on the dataset and have not compared their results with any of the previous state-of-the-art methods by generating their results on their dataset. The accuracy obtained by them on their dataset was also not given and have just shown precision and recall values for the experiments. Griffis et al. Griffis et al. (2016) have also worked on ischemic stroke lesions identification where they utilized the T1-weighted MRI scans and predict the stroke region using Gaussian naïve Bayes classification method. Limitations of their method was that they were unable to detect very subtle WM lesions with very small extents. Shahangian and Pourghassem Shahangian and Pourghassem (2016) proposed a method to segment and classify hemorrhage, where segmentation was based on a modified version

of distance regularized level set evolution (MDRLSE). Thereafter, SVM classifier was used to classify images by extracting weighted grayscale histogram features where they achieved an accuracy of 94.13%. However, the complete dataset information is not given like under what parameters images have been acquired and the details of CT scanner. The other limitation is that, the number of images which they have used for classification were also very less for 5 different categories i.e. only 627 images. The segmentation of hemorrhagic stroke from CT scan images was proposed by Gautam et al. Gautam et al. (2018) using variant of fuzzy clustering named as modified robust fuzzy c-means clustering (MRFCM) and used it with distance regularization level set evolution (DRLSE) method for the segmentation. However, their method was not able to identify very small lesions from the images. They have also proposed the method for the classification of a brain CT scan images into hemorrhagic, ischemic and normal Gautam and Raman (2020). A new feature descriptor known as the local gradient of gradient pattern (LG2P) descriptor has been proposed by them. The highest classification accuracy achieved by them on 900 image dataset is 83.11% and 86.11% using fine NN and cubic SVM. However, this accuracy is very less and for better results it must be improved. Incremental gradient descent decision boundary feature extraction method was proposed by Woo et al. Woo and Lee (2019) for the classification of UCI dataset where they have shown results on Parkinson speech, chess, molecular biology, gas sensor, and array drift datasets of UCI. Recently, some researchers have published a literature survey on feature extraction methods and they can be found in Salau and Jain (2019); Chowdhary and Acharjya (2020), so readers are referred to these papers for more methods.

Subudhi et al. Subudhi et al. (2020) have proposed a new method for the segmentation and classification of brain stroke from MR images where they used expectation–maximization and random forest classifier. The accuracy achieved by them was 93.4% on the dataset of 192 brain images. The limitation of this paper is that they do not have any discussion of the dataset like from where they collected it and how. Ortiz-Ramón et al. Ortiz-Ramón et al. (2019) also worked on MR images where they proposed a texture analysis method to identify whether the patient has a stroke or not and also identify whether that stroke is lacunar or cortical. In this paper, they have used various methods proposed by different researchers in order to carry out their work and references for them can be found in Subudhi et al. (2020). According to them the limitation of their work is inability to combine stroke and aging (old stroke lesion) datasets for the analysis of images showing recent cortical or lacunar strokes. However, image normalization techniques can solve this issue by improving the image qualities.

Apart from the conventional machine learning methods which have been used very widely for the identification of diseases from the medical images, deep learning approaches are gaining attention because of the accurate diagnosis results. In 2012, Krizhevsky et al. proposed AlexNet as the CNN architecture and won the most difficult ImageNet large-scale visual recognition challenge (ILSVRC) Krizhevsky et al. (2017). Thereafter, it emerged as the revolution in the field of computer vision and machine learning. AlexNet has 5 convolutional and 3 fully connected layers where 2 dropout layers have been used with 0.5 probability. The best classification prediction with it can be achieved when the softmax layer is fed with the result of the last fully connected layer. In this architecture, the input image size should be $224 \times 224 \times 3$. However, in order to improve the accuracy many researchers have tried to develop new architectures. In 2014, VGG16 has come into existence, proposed by Simonyan and Zisserman Simonyan and Zisserman (2014). The main contribution of their network is that they have made a very deep CNN architecture for the recognition of very large image dataset by increasing the network depth to 16–19 layers with 3×3 convolution filter. VGG16 consists of five batches of 13–16 convolutional layers and 3 fully connected layers. In the first layer the number of filters starts from 64 and increases by a factor of 2 for each layer after the max-pooling till it becomes 512. The input image size for this architecture remains the same as that of AlexNet i.e $224 \times 224 \times 3$. The other CNN architecture ResNet50 proposed by He et al. He et al. (2016) is a very deep residual network and it won the first place in ILSVRC 2015. It is a 50 layer architecture instead of 16–19 layers, which was trained on the 1.28 million training images. The input image size for this architecture is $224 \times 224 \times 3$. After that, many variants of CNN architectures have come into existence in order to provide better results for the diagnosis of the patients. Recently, deep learning has shown good classification accuracy on both training and testing datasets of breast cancer histopathology images Gupta and Bhavsar (2018). The other work on the detection and classification of breast cancer makes use of four fully convolutional networks (FCN) Gecer et al. (2018). CNN has also been used to detect colorectal cancer from the histology images Yoon et al. (2019). Gao et al. Gao et al. (2018) used deep CNNs to classify interstitial lung diseases (ILD) from CT images by using the entire image as a holistic input.

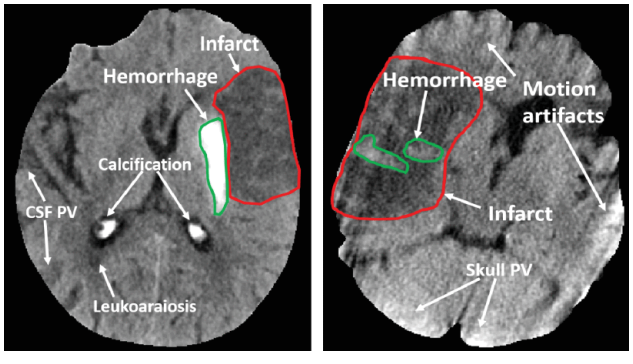

Fig. 1. Samples of CT slices to depict *overlapping infarcts* – infarcts of multiple types at the same image region.

However, the work on the classification and localization of origin of brain strokes from CT/MR images still requires improvements for clinical applicability. Furthermore, the presence of overlapping infarcts of the two types – ischemic and hemorrhagic – on the CT slides, as depicted in Figure 1 gives rise to a *multi-label* segmentation problem, where each image

pixel may belong to more than one class. Therefore, in this research, we propose to develop a new deep learning -based processing pipeline to classify brain stroke from CT scan images, and further localize the infarct through volumetric reconstruction. Moreover, to the best of the authors' knowledge, this work is the first attempt to implement simultaneous volumetric reconstruction and correction of segmentation to improve demarcation performance and evolve a pliable computational model that can be used to infer quantitative diagnostic measures and with further extensions, for simulating surgical procedures as well.

For thrombectomy path planning, Jia Y. et al. Jia et al. (2020) proposes a learning-based obstacle avoidance method using Q-learning for path planning in six-axis robotic arms using reinforcement learning for autonomous obstacle avoidance. The strategy prioritizes planning the obstacle avoidance path for the terminal point of the mechanical arm and then uses the calculated terminal path to plan the poses of the mechanical arm. If the mechanical arm cannot avoid obstacles within the limit of the safe distance for points on the terminal path, this strategy records those points as new obstacles and plans a new obstacle avoidance path for the terminal of the mechanical arm. This process is looped until the correct path is calculated. This method was adopted as a baseline with noteworthy modifications in our work.

3. Methods

In recent years, research in deep learning has demonstrated successful application of deep learning and convolutional neural networks for image classification, including medical image analysis. In this study, we propose to develop a deep neural network -based processing pipeline for classifying different stroke patterns on NCCT images, and localizing the infarct for downstream treatment formulation.

Toward solving the problem, we will collect NCCT imaging slides of study patients. Following this, the digital images will be randomly partitioned into two sets, one for model training and another as an independent test set for the final evaluation of the model. Care will be taken to avoid overlapping slides from a given patient across the training and test sets. Each slice will be manually assessed and its regions will be labeled as either *normal (0)*, *ischemic infarction (1)*, *hemorrhagic infarction (2)* by a trained cardiovascular radiologist. Thus, labeled slice image will form the input for training the deep learning -based computational pipeline. The model will be developed to take in square patches or tiles as inputs and output a prediction probability for each of the three classes: 0, 1, and 2 in a multi-label setting. Different model parameters including network weights, test for multiple network layers, learning rates, loss functions will be considered to fine-tune the model and formulate a model with optimal parameter settings. An unseen held-out subset of the slices data will be used for independently validating and evaluating performance and accuracy of the trained model. Finally, the models' ability to recognize the individual segmentation classes will be analyzed. As a final step, the processed image slides with infarct region annotations, as predicted by the developed computational model, will be used to perform volumetric reconstruction and infarct region correction. Furthermore, we will visualize the model's predictions overlaying color-coded dots on patches for which the model predicted patterns. This will aid in quickly identifying regions of the slice(s) from a given patient containing abnormal tissue. An overview of the proposed solution architecture is presented in Figure 2.

3.1. NCCT Scan Processing Module

An overview of the NCCT scan processing module is presented in Figure 3.

Data Collection.. To develop and evaluate our computational model for detecting distinctive infarct patterns in CTs, we will acquire volumetric NCCT slices in a digital image format, typically DICOM or NIfTI. Following this, the images will be randomly partitioned into two sets: (a) one for *model building* – i.e., training and validation (about three-fourths of the dataset), and (b) for *model evaluation* (about one-fourth of the dataset).

Slice-level Annotation to Establish Ground Truth.. All NCCT slice images will be manually labeled by one or more medical experts. Each image corresponding to the model building dataset will be annotated with scan regions as as either one of the two patterns or normal/non-applicable. These are: *normal (0)*, *ischemic infarction (1)*, *hemorrhagic infarction (2)*. Ideally, we would resort to multiple independent annotation results by the experts and correct for inter-rater biases to ensure high quality of the ground truth. For the purpose of annotations, a granulated user-friendly tool will be provided to the medical expert(s) to perform manual annotation at the super-pixel image region level, with the ability to select large regions of the image at a time and label them. Furthermore, an unsupervised clustering-based segmentation algorithm was applied initially to facilitate easy manual annotation of infarct regions on the NCCT slices. This algorithm is presented in Figure 3.1. This. In addition, case-specific comments were collected for downstream evaluation and analysis of the developed processing framework. The test set was used, once the model development was complete, to evaluate its performance and compare it against that of expert annotators.

Data Pre-processing.. The input data in the form of CT slices will be processed by a computation processing pipeline that comprise deep learning network to analyze images and draw inferences. Initially, basic pre-processing methods to ensure consistency of image quality and exposure statistics across the dataset will adopted, and further pre-processing workflows may be integrated at a later stage during model building by empirically determining the specific image artifacts to be managed and the method required to do so. Some of the basic methods include: automatic resizing of training images to contain a fixed dimension of pixels; apply color jittering on the brightness, contrast, saturation, and hue of each image to neutralize the differences across the dataset; and augmentation of image data to randomly flip and rotate them along the horizontal and vertical axes to obtain variability in the training data, and to ensure robust model development, among others.

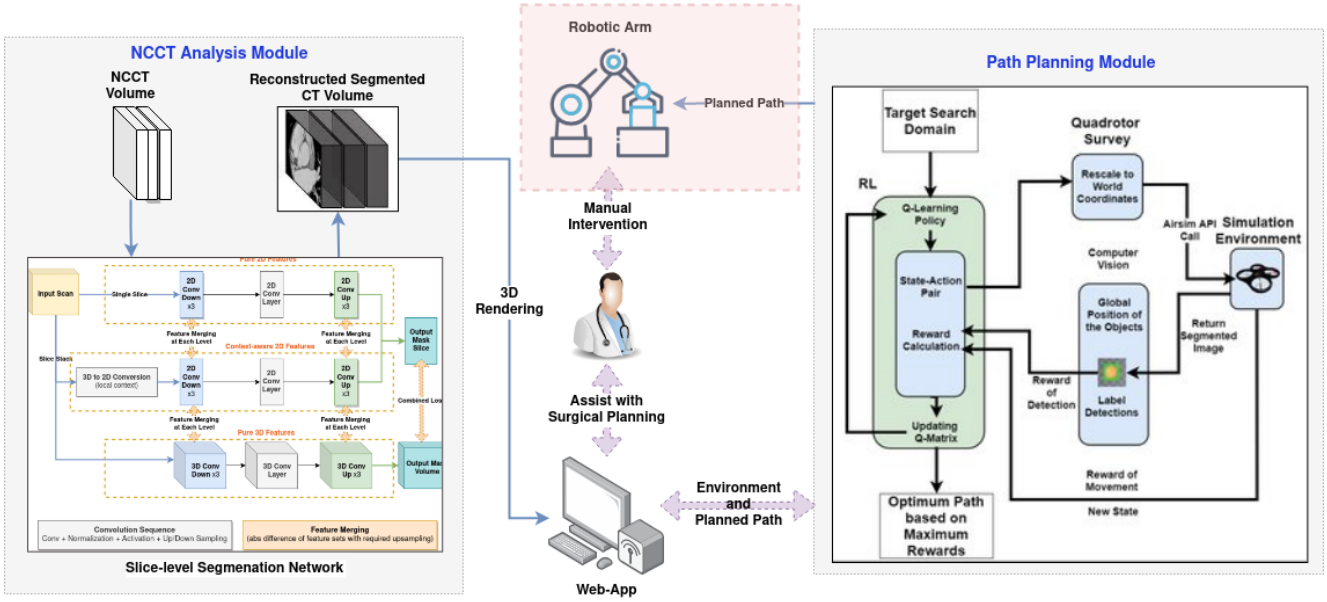

Fig. 2. Architecture of the proposed solution.

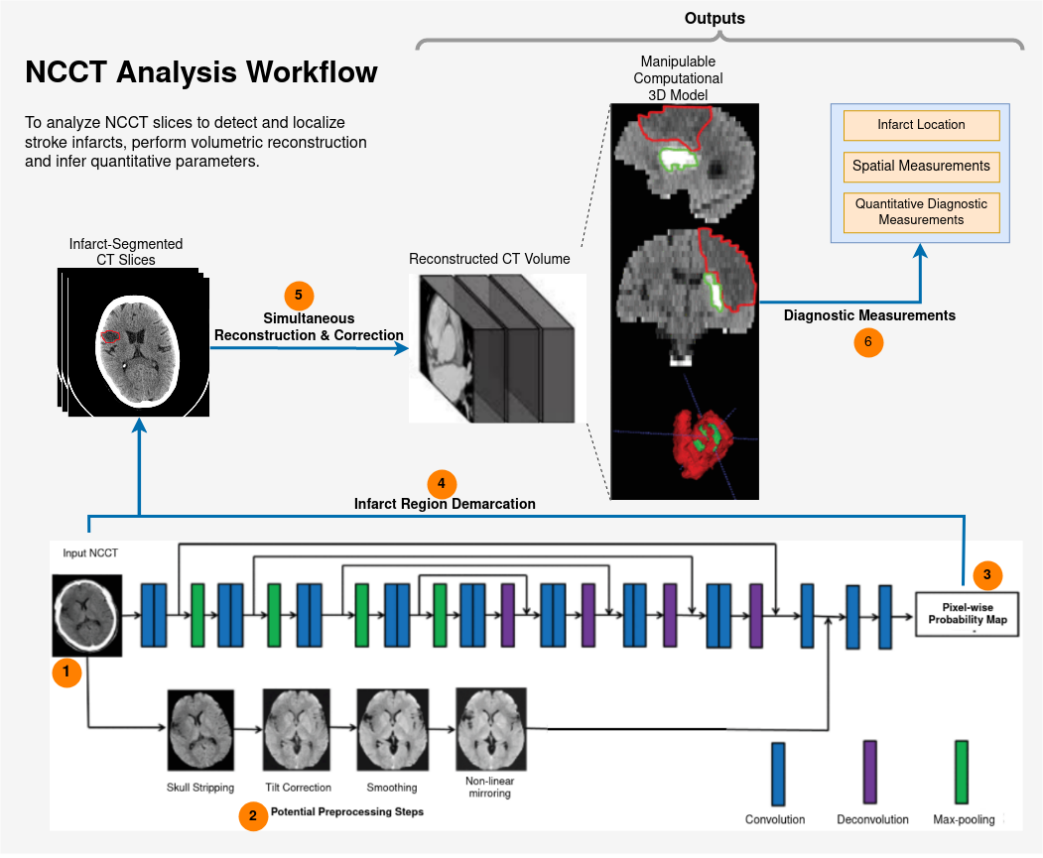

Fig. 3. Overview of the NCCT analysis workflow following steps 1 through 6. The network architecture depicted in the figure is illustrative only.

Segmentation Model Building.. Before training the classification and segmentation models, 90% of the model building data will be randomly selected for training and 10% for internal validation. The models will be implemented using the deep learning library Torch available for Python. Torch is a high-level neural networks API, capable of running on top of various backend architectures. The models will be trained under a supervised setting, with NCCT slices fed as input, and outputs corrected against their corresponding expert-annotated and inter-rater corrected NCCT slices. Different tests will be conducted to optimize for the number of epochs, initial learning rates as well as the optimal depth of the neural network and

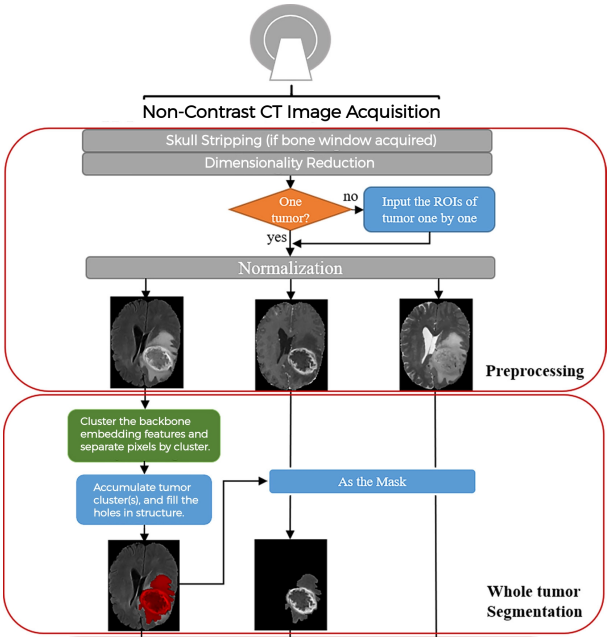

Fig. 4. An unsupervised approach to segment hemorrhagic infarcts from brain CTs.

other functionally configurable portions of the computational pipeline. The output of the models will be configured to output probability vectors p per stroke type, normalized to 1. Finally, the learning curves (accuracy vs. no. of iterations) for both training and testing will be plotted to determine if there is a generalization of the deep learning network model(s) to the unseen validation/test data i.e., no over-fitting has occurred. Two state-of-the-art slice-based image segmentation neural networks, namely Slice-Aware Net (SA-Net) Mei et al. (2022) and M-Net Mehta and Sivaswamy (2017) were adopted as improved models. In addition, a combination of the two models was also implemented to combine the strengths of these state-of-the-art approaches – long-range dependency capturing ability of SA-Net and spatial context capturing ability of M-Net. The architecture of this proposed architecture is presented in 5.

Volumetric Reconstruction and Correction.. With the infarct segmentation workflows in place, the subsequent steps involve reconstructing the volumetric CT of the patient's brain with annotated infarct regions, labeled by type – ischemic and hemorrhagic.This process will involve stacking the processed NCCT slides whilst simultaneously correcting the neighbors for potential errors in region segmentation. These corrections will be automated based on geometric, spatial, and shape-based constraints of adjacent slides such as contour evolution methods Kuang et al. (2019). The correction strategy will not only improve the overall localization accuracy of the ischemic strokes facilitating subsequent use of the reconstructed volume, but will save compute time due to simultaneous reconstruction and correction and improve model inference through a more interpretable basis for reconstruction. Ultimately, this step yields an annotated volumetric model of the brain region that is pliable for exploratory analysis on a computer.

Diagnostic Inference from Volumes.. The ultimate goal is to localize infarct types and locations in the spatial context. In addition, we also propose to estimate quantitative diagnostic measures. The annotated volume of the patient's brain can be easily exploited to draw out any form of geometric, spatial, or structural measurements, and even a combination thereof. Furthermore, the presence of each type of infarct can be individually quantified to present summary inferences of the patient diagnosis to help clinicians devise an ideal treatment strategy.

Validation from Radiologists.. At the completion of the complete processing pipeline, two sets of validations will be performed to ensure correctness of the overall computational pipeline – *(1) the volume-level annotations* will be validated by radiologists, and *(2) the slice-level annotations* of the held-out test-set will be compared to the model predictions. *Evaluation 1* will be qualitative in nature, owing to the tediousness of manually annotating volumetric regions. However, this evaluation will be supported by *evaluation 2*. In conjunction, the two sets of validations will provide sufficient evidence to verify the complete proposed computational pipeline. To evaluate the slide-level performance, the model's labels will be compared against those of the radiologists' by calculating an inter-rater reliability metric called Cohen's kappa score McHugh (2012), as it has been adapted as a standard metric when it comes to annotation of medical images. Thus, between every two sets of annotations from radiologists and/or the computational pipeline, both $_{pd}$ scores and kappa score per class will be computed. K_{pd} refers to the percentage of NCCT slices in which two annotators agreed on the same annotation. Kappa scores per infarct class, on the other hand, estimates the detection of a given infarct pattern or type between two sets of annotations. The annotations will be quantified using a suitable metric as deemed suitable during discussions with the experts.

3.2. Robotic Path Planning Module

Building atop the proposed Q-learning stratgey in Jia et al. (2020), and deriving from methods for obstacle avoidance in robotics, such as the artificial potential field-based industrial robot station conversion obstacle avoidance system, genetic algorithms, and RBF neural networks have limitations in complex environments and high-dimensional multi-constraint path planning problems. The approach prioritizes planning the obstacle avoidance path for the terminal point of the mechanical arm and then uses the calculated terminal path to plan the poses of the mechanical arm. If the mechanical arm cannot avoid obstacles within the limit of the safe distance for points on the terminal path, the strategy records those points as new obstacles and plans a new obstacle avoidance path for the terminal of the mechanical arm, looping this process until the correct path is calculated.

An overview of the path planning workflow is presented in Figure 6. The proposed method makes use of a novel distance function and simulated annealing techniques for exploratory actions.

The distance is calculated using a custom metric that incorporates the distance between the current and previous arm positions and the target point. The function uses vector algebra

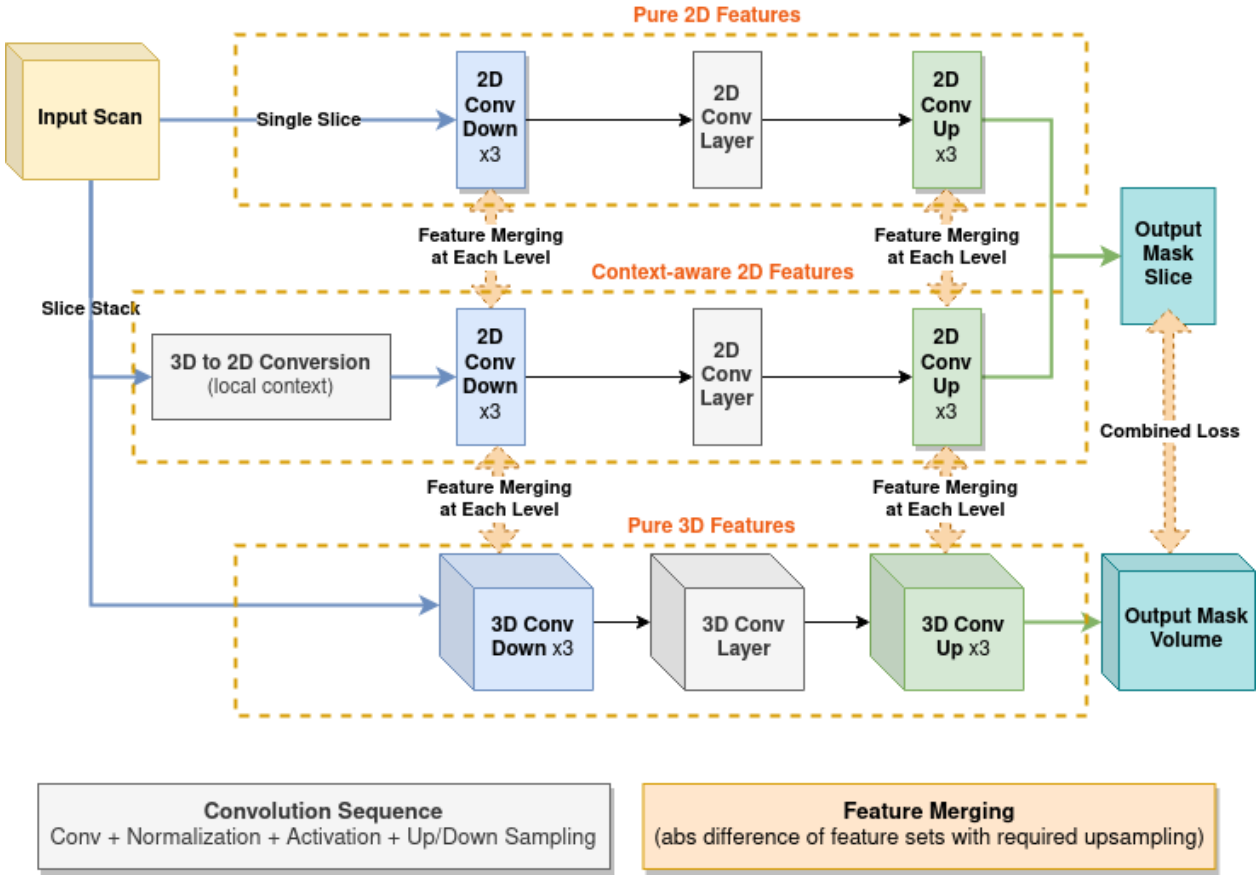

Fig. 5. Proposed segmentation network architecture, combining multi-dimensional image features.

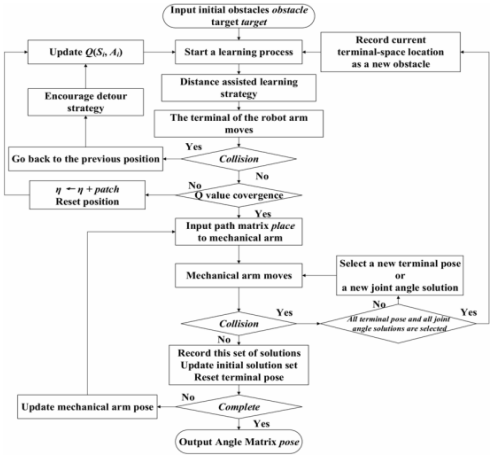

Fig. 6. Path planning workflow based on Q-learning for the 6-DOF robotic arm.

to determine the shortest distance between the line segment and the third point. Given two vectors $\mathbf{u}$ and $\mathbf{v}$ with n elements, the maximum distance between them using the Chebyshev and Minkowski distance metrics, with p defined as the difference between the Euclidean distances of the current arm position and previous arm position to the target point, described in Equation 1.

$$\max\left(\max_{i=1}^{n}|u_i - v_i|, \left(\sum_{i=1}^{n}|u_i - v_i|^p\right)^{1/p} - \left(\sum_{i=1}^{n}|u'_i - v'_i|^p\right)^{1/p}\right) \quad (1)$$

where $\mathbf{u}'$ and $\mathbf{v}'$ are the previous arm position and target, respectively, $\|\cdot\|_2$ denotes the Euclidean norm. Further, p is computed using Equation 2.

$$\|currentarmpos - target\|^2 - \|previousarmpos - target\|^2) \quad (2)$$

In addition, a temperature-based simulated annealing strategy is used to decide between greedy (maximizing reward) and exploratory (random) actions.

Let $E(s)$ be the energy (or cost) associated with state s, ΔE be the energy difference between the current state and the proposed next state, and $T(t)$ be the temperature at iteration t. The probability of accepting a worse state at temperature $T(t)$ is given by the Boltzmann distribution:

$$P(\Delta E, T(t)) = \exp\left(-\frac{\Delta E}{T(t)}\right)$$

To decide whether to take a greedy or exploratory action, start with a high initial temperature T_0 and gradually decrease it over time according to a cooling schedule, such as:

$$T(t) = T_0 \cdot \alpha^t$$

where α is the cooling rate, t is the iteration number, and $T(t)$ decreases exponentially as t increases. At each iteration, we can generate a random neighboring state of the current state, and compute the energy difference ΔE. If $\Delta E \leq 0$, we accept the new state as the current state. Otherwise, we accept the new state with probability $P(\Delta E, T(t))$.

4. Experiments and Results

Exploratory Data Analysis.. To prepare our processing and experimentation pipeline for the proprietary data we expect to acquire from *Chettinad Academy for Research and Education*, we reviewed closely-related NCCT datasets. To analyze the applicability of this data to our problem statement, and to ensure translation of developed methods to the proprietary dataset, we performed EDA to analyze these datasets. Based on these analyses, we chose to benchmark our experiments using the hemorrhage dataset Hssayeni et al. (2020). These analyses are available in our project codebase at: `https://github.com/karthik-d/Vision-For-Robot-Path-Planning`.

A peer-reviewed *intracranial hemorrhage dataset* Hssayeni et al. (2020) comprising 2500 brain window images and 2500 bone window images collected from 82 patient samples. This dataset provides annotations indicating the location of the hemorrhage on each slice.

Segmentation Models.. For segmentation, we first performed experiments using state-of-the-art deep-learning approaches on open-access datasets, particularly a peer-reviewed intracranial hemorrhage dataset Hssayeni et al. (2020), to assess the suitability and viability of these methods and to arrive upon a suitable course of action for the proprietary dataset. Furthermore, we explored the possibility of adopting an *unsupervised* approach for segmentation to minimize, or even eliminate the need for time-consuming manual annotation of stroke infarcts by expert radiologists. These experiments are summarized in Table 1.

In addition to the most commonly used supervised deep-learning methods, we also experimented with an unsupervised clustering-based approach to deal with sparse availability of lesion annotation masks. This workflow is presented in Figure 3.1. The approach can be further improved, for the proprietary dataset if appropriate, by integrating a high-performing pre-trained deep-learning backbone model as an autoencoder, or even in the few-shot learning regime. A qualitative comparison of performance is presented in Figure 7. The modified segmentation architecture is presented in 5.

A qualitative comparison of segmentation using the proposed network and other existing work assessed on the same dataset is presented in Figure 7. Table 1 presents a quantitative comparison.

It is tacit that 2D models generally miss out on small regions of hemorrhage. This is likely due to lack of context from neighboring scan slices, which is otherwise available to 3D processing models. The proposed method is able to draw out this contextual information to identify even small hemorrhagic regions well (for instance, row 3 in Figure 7). The 2D methods, at time, also produce segmentation regions that are starkly diminished or expanded in size

Furthermore, the 3D models tend to exhibit a smoothing effect of the segmented region. Such smoothing effect is often characteristic of naturally occurring hemorrhagic regions, as opposed to sharp-edged or ridged regions. However, this often causes closely-spaced hemorrhagic regions to get clubbed into a single region when processed by 3D segmentation models. The proposed approach, on the other hand, is able to sufficiently resolve even such closely spaced regions (for instance, row 5 in Figure 7), potentially due to the combined effect of multi-dimensional features, achieving the closest resemblance to the ground truth.

Path Planning for Thrombectomy.. To model robotic environments based on quantitative parameters measured from the proposed NCCT analysis pipeline, we conducted experiments to determine the most feasible simulation workflows and perform path-planning for a 6-DOF arm.

We then performed experiments to evaluate the proposed path-planning strategies, with the following parameter set,

$$\begin{aligned} initial\_temp &= 1.0; \\ cooling_rate &= 0.005; \\ episode &= 1; \\ finite_states &= 100000; \\ iterations &= 5; \end{aligned} \tag{3}$$

Qualitative results on one simulation environment is presented in Figure 4. It is worth noting that the proposed changes to the path-planning algorithm was able to improve the comparative baseline by successfully completing 6 of 7 target positions, as opposed to the baseline's 4 of 7.

The results show that the assisted learning strategies efficiently speed up the learning process and the mechanical arm can successfully avoid obstacles and reach the target point safely, even when limited by safe distance constraints. The proposed approach has potential applications in industries such as manufacturing, warehousing, and transportation.

Acknowledgments

The authors would like to thank SSN College of Engineering, Chennai, India for supporting this research.

References

Başaran, E., Cömert, Z., Çelik, Y., 2020. Convolutional neural network approach for automatic tympanic membrane detection and classification. Biomedical Signal Processing and Control 56, 101734.

Chawla, M., Sharma, S., Sivaswamy, J., Kishore, L., 2009. A method for automatic detection and classification of stroke from brain ct images, in: 2009 Annual international conference of the IEEE engineering in medicine and biology society, IEEE. pp. 3581–3584.

Chowdhary, C.L., Acharjya, D.P., 2020. Segmentation and feature extraction in medical imaging: a systematic review. Procedia Computer Science 167, 26–36.

Durduran, T., Yu, G., Burnett, M.G., Detre, J.A., Greenberg, J.H., Wang, J., Zhou, C., Yodh, A.G., 2004. Diffuse optical measurement of blood flow, blood oxygenation, and metabolism in a human brain during sensorimotor cortex activation. Optics letters 29, 1766–1768.

Table 1. Comparison of segmentation approaches on open-access datasets.

Approach	Backbone	Dataset	DSC	Mean IoU	AuROC
FPN	EffNet-B0	Peer-Reviewed Intracranial Hemorrhage Dataset Hssayeni et al. (2020)	41.18%	28.20%	-
UNet	EffNet-B0		46.73%	30.42%	-
PSPNet			40.21%	27.51%	-
DeepLabV3+			33.82%	17.43%	-
Best UNet *			44%	27.5%	-
ChanVese [2]*			70%	-	-
M-Net			70.41%	59.95%	86.13%
Proposed			**76.11%**	**64.52%**	**89.15%**
Clustering Analysis			**32.72%**	**21.09%**	-

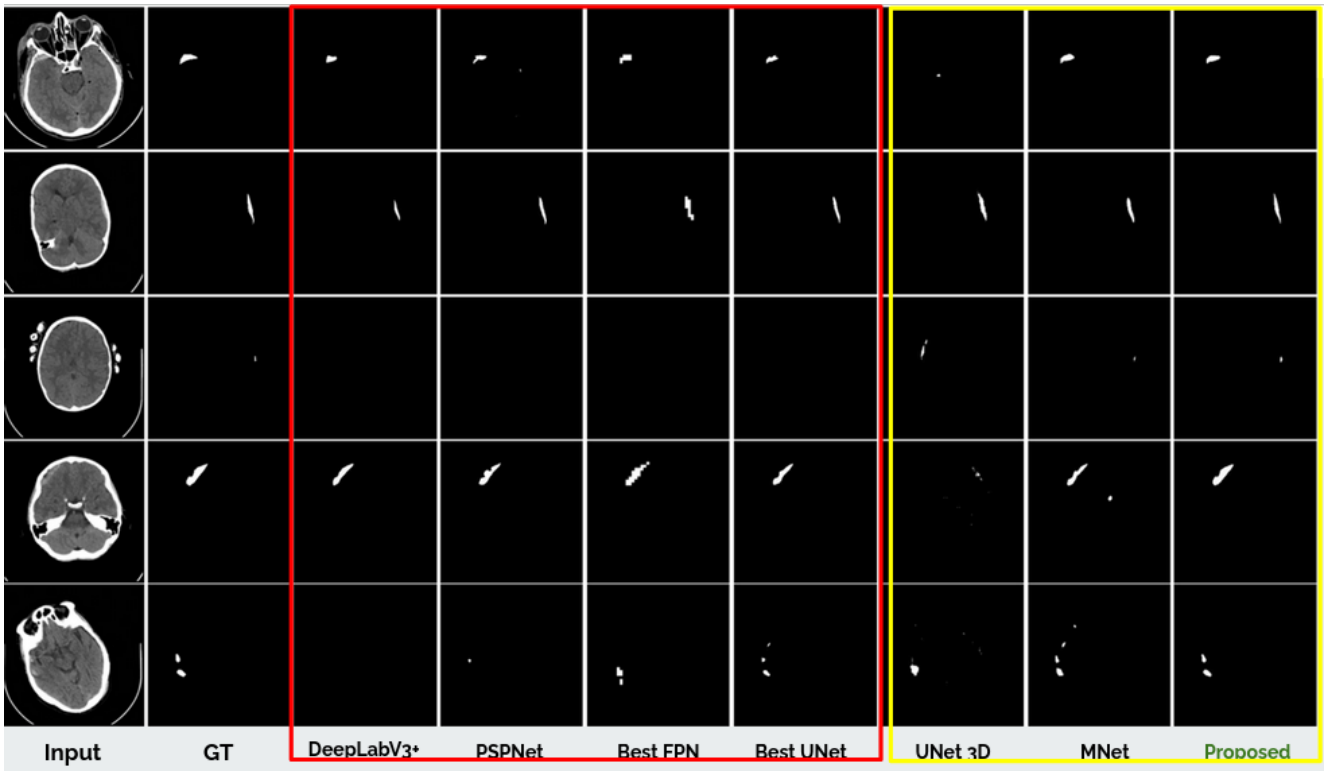

Fig. 7. Qualitative comparison of segmentation methods. 2D models are outlined in red. 3D models are outlined in yellow.

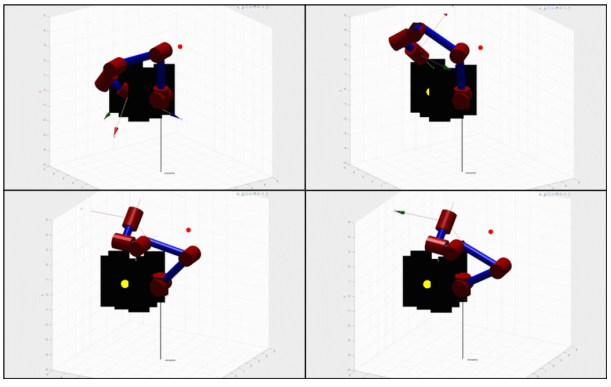

Fig. 8. Intermediate states of the 6-DOF robotic arm when implementing the planned path in one of the six successful environments successfully planned by the proposed algorithm.

Gao, M., Bagci, U., Lu, L., Wu, A., Buty, M., Shin, H.C., Roth, H., Papadakis, G.Z., Depeursinge, A., Summers, R.M., et al., 2018. Holistic classification of ct attenuation patterns for interstitial lung diseases via deep convolutional neural networks. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization 6, 1–6.

Gautam, A., Raman, B., 2020. Local gradient of gradient pattern: a robust image descriptor for the classification of brain strokes from computed tomography images. Pattern Analysis and Applications 23, 797–817.

Gautam, A., Raman, B., Raghuvanshi, S., 2018. A hybrid approach for the delineation of brain lesion from ct images. Biocybernetics and Biomedical Engineering 38, 504–518.

Gecer, B., Aksoy, S., Mercan, E., Shapiro, L.G., Weaver, D.L., Elmore, J.G., 2018. Detection and classification of cancer in whole slide breast histopathology images using deep convolutional networks. Pattern recognition 84, 345–356.

Griffis, J.C., Allendorfer, J.B., Szaflarski, J.P., 2016. Voxel-based gaussian naïve bayes classification of ischemic stroke lesions in individual t1-weighted mri scans. Journal of neuroscience methods 257, 97–108.

Gupta, V., Bhavsar, A., 2018. Sequential modeling of deep features for breast cancer histopathological image classification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 2254–2261.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.

Hinton, G.E., Osindero, S., Teh, Y.W., 2006. A fast learning algorithm for deep belief nets. Neural computation 18, 1527–1554.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural computation 9, 1735–1780.

Hssayeni, M., Croock, M., Salman, A., Al-khafaji, H., Yahya, Z., Ghoraani, B., 2020. Computed tomography images for intracranial hemorrhage detection and segmentation. Intracranial Hemorrhage Segmentation Using A Deep Convolutional Model. Data 5, 179.

Jain, S., Salau, A.O., 2019. Detection of glaucoma using two dimensional tensor empirical wavelet transform. SN Applied Sciences 1, 1–8.

Jia, Y., Li, Y., Xin, B., Chen, C., 2020. Path planning with autonomous obstacle avoidance using reinforcement learning for six-axis arms, in: 2020 IEEE International Conference on Networking, Sensing and Control (ICNSC), pp. 1–6. doi:10.1109/ICNSC48988.2020.9238112.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. Imagenet classification with deep convolutional neural networks. Communications of the ACM 60, 84–90.

Kuang, H., Menon, B.K., Qiu, W., 2019. Segmenting hemorrhagic and ischemic infarct simultaneously from follow-up non-contrast ct images in patients with acute ischemic stroke. IEEE Access 7, 39842–39851.

Matesin, M., Loncaric, S., Petravic, D., 2001. A rule-based approach to stroke lesion analysis from ct brain images, in: ISPA 2001. Proceedings of the 2nd International Symposium on Image and Signal Processing and Analysis. In conjunction with 23rd International Conference on Information Technology Interfaces (IEEE Cat., IEEE. pp. 219–223.

McHugh, M.L., 2012. Interrater reliability: the kappa statistic. Biochemia medica 22, 276–282.

Mehta, R., Sivaswamy, J., 2017. M-net: A convolutional neural network for deep brain structure segmentation, in: 2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017), IEEE. pp. 437–440.

Mei, J., Cheng, M.M., Xu, G., Wan, L.R., Zhang, H., 2022. Sanet: A slice-aware network for pulmonary nodule detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 44, 4374–4387. doi:10.1109/TPAMI.2021.3065086.

Members, W.G., Roger, V.L., Go, A.S., Lloyd-Jones, D.M., Benjamin, E.J., Berry, J.D., Borden, W.B., Bravata, D.M., Dai, S., Ford, E.S., et al., 2012. Executive summary: heart disease and stroke statistics—2012 update: a report from the american heart association. Circulation 125, 188–197.

Mousavi, Z., Rezaii, T.Y., Sheykhivand, S., Farzamnia, A., Razavi, S., 2019. Deep convolutional neural network for classification of sleep stages from single-channel eeg signals. Journal of neuroscience methods 324, 108312.

Nowinski, W.L., Qian, G., Hanley, D.F., 2014. A cad system for hemorrhagic stroke. The neuroradiology journal 27, 409–416.

Ortiz-Ramón, R., Hernández, M.d.C.V., González-Castro, V., Makin, S., Armitage, P.A., Aribisala, B.S., Bastin, M.E., Deary, I.J., Wardlaw, J.M., Moratal, D., 2019. Identification of the presence of ischaemic stroke lesions by means of texture analysis on brain magnetic resonance images. Computerized Medical Imaging and Graphics 74, 12–24.

Ouhmich, F., Agnus, V., Noblet, V., Heitz, F., Pessaux, P., 2019. Liver tissue segmentation in multiphase ct scans using cascaded convolutional neural networks. International journal of computer assisted radiology and surgery 14, 1275–1284.

Peixoto, S.A., Rebouças Filho, P.P., 2018. Neurologist-level classification of stroke using a structural co-occurrence matrix based on the frequency domain. Computers & Electrical Engineering 71, 398–407.

Pereira, P.M., Fonseca-Pinto, R., Paiva, R.P., Assuncao, P.A., Tavora, L.M., Thomaz, L.A., Faria, S.M., 2020. Skin lesion classification enhancement using border-line features–the melanoma vs nevus problem. Biomedical Signal Processing and Control 57, 101765.

Saatman, K.E., Duhaime, A.C., Bullock, R., Maas, A.I., Valadka, A., Manley, G.T., 2008. Classification of traumatic brain injury for targeted therapies. Journal of neurotrauma 25, 719–738.

Salau, A.O., Jain, S., 2019. Feature extraction: a survey of the types, techniques, applications, in: 2019 International Conference on Signal Processing and Communication (ICSC), IEEE. pp. 158–164.

Shahangian, B., Pourghassem, H., 2016. Automatic brain hemorrhage segmentation and classification algorithm based on weighted grayscale histogram feature in a hierarchical classification structure. Biocybernetics and Biomedical Engineering 36, 217–232.

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 .

Subudhi, A., Dash, M., Sabut, S., 2020. Automated segmentation and classification of brain stroke using expectation-maximization and random forest classifier. Biocybernetics and Biomedical Engineering 40, 277–289.

Suk, H.I., Lee, S.W., Shen, D., Initiative, A.D.N., et al., 2014. Hierarchical feature representation and multimodal fusion with deep learning for ad/mci diagnosis. NeuroImage 101, 569–582.

Toğaçar, M., Cömert, Z., Ergen, B., 2020. Classification of brain mri using hyper column technique with convolutional neural network and feature selection method. Expert Systems with Applications 149, 113274.

Torres-Mozqueda, F., He, J., Yeh, I., Schwamm, L., Lev, M., Schaefer, P., Gonzalez, R., 2008. An acute ischemic stroke classification instrument that includes ct or mr angiography: the boston acute stroke imaging scale. American journal of neuroradiology 29, 1111–1117.

Wan, C.H., Lee, L.H., Rajkumar, R., Isa, D., 2012. A hybrid text classification approach with low dependency on parameter by integrating k-nearest neighbor and support vector machine. Expert Systems with Applications 39, 11880–11888.

Woo, S., Lee, C., 2019. Incremental feature extraction based on gaussian maximum likelihood, in: 2019 34th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), IEEE. pp. 1–4.

Wu, B., Khong, P.L., Chan, T., 2012. Automatic detection and classification of nasopharyngeal carcinoma on pet/ct with support vector machine. International journal of computer assisted radiology and surgery 7, 635–646.

Yoon, H., Lee, J., Oh, J.E., Kim, H.R., Lee, S., Chang, H.J., Sohn, D.K., 2019. Tumor identification in colorectal histology images using a convolutional neural network. Journal of digital imaging 32, 131–140.

Zhang, L., Samaras, D., Tomasi, D., Volkow, N., Goldstein, R., 2005. Machine learning for clinical diagnosis from functional magnetic resonance imaging, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), IEEE. pp. 1211–1217.