Detecting and Segmenting Hematoma in Stroke Patients

Julian P. Stys
Department of Computing Science
University of Alberta
jstys@ualberta.ca

Colin D'Amore
Department of Computing Science
University of Alberta
cdamore@ualberta.ca

Yiheng Fu
Department of Computing Science
University of Alberta
yiheng5@ualberta.ca

Abstract

Hemorrhage detection has been at the forefront of medical research since the resurgence of machine learning. In this project, we are interested in studying CT-scans from hemorrhagic stroke patients to train a model that can automatically detect and segment these hematoma (collection of blood outside blood vessels in the brain). We are given a dataset from 88 patients, comprising of a total of 3600 labelled images to train from. We employ 2 collaborative techniques to solve the problems we wish to address throughout this project. After training on a Fully Convolutional Network (FCN) and statistically optimizing using a Conditional Random Field (CRF) method, we achieve a pixel-wise accuracy of 99.5% on a validation set, and a Jaccard Score of 0.946 between the predicted hematoma label and the true label provided to us.

1 Introduction

Strokes have historically been one of the leading causes of death, accounting for about 9.7% of fatalities worldwide [1]. Of the two primary stroke types, hemorrhagic strokes occur when blood from arteries bleed into the brain and although these types are uncommon, they have a very high mortality rate, reporting that only 26.7% of patients that suffer from them survive for 5 years after they had the stroke [2]. We develop a program that can detect and segment blood pools in stroke patients which can aid doctors and researchers to better diagnose and study the physiological effects of hemorrhagic strokes on patients.

1.1 Computer Tomography

Computer tomography (CT) is an imaging method which acquires cross-sectional volumetric images from a specific region in the human body [3]. CTs work by taking angled X-ray measurements around a person's body, then interpolating between these measurements in 3-dimension to extract a stacked series of images of constant thickness. CT-scans are commonly used to diagnose hemorrhagic stroke patients, as calcium rich hemorrhages affect incoming X-rays, and therefore look different from the cerebrum of one's brain. As a result, CT-scans are used to successfully image these hemorrhages, and provide doctors with diagnosis as how to further treat their patients.

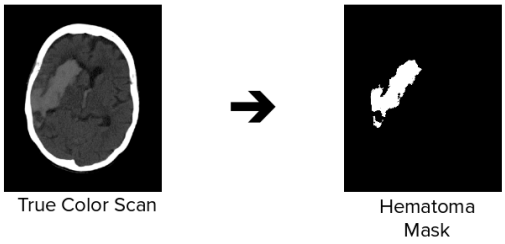

Figure 1: **The goal we wish to achieve.** Given a CT scan slice we want to output a 0/1 mask of where the hematoma is located in the scan.

1.2 Our Task

Given a CT scan of the brain, we aim to detect the contours of hematoma present in the scan as shown in Figure 1. This aims to remove the need for this task to be done by hand by a medical specialist, and could allow for a doctor to be able to more quickly locate present hematoma in the brain. Our model can also be used as a binary classification task that outputs a completely black mask when there is no hematoma present, and a segmented outline of the hematoma when there is hematoma in the brain.

1.3 Related Work

A good deal of similar work has been done in the past relating to hematoma detection/segmentation. For example, Chilamkurthy et al [4] have developed a model which analyzes CT-scans and predicts which type of trauma injury is present (intracranial hematoma, midline shift, etc.). These researchers trained from about 300,000 CT-scans to produce a high accuracy classification for 9 different types of head injuries that are most common in trauma patients. Fu et al [5] have done work with segmenting mid body organs such as stomach, kidney and liver using MRI scans. They used a similar approach to ours, using a convolution neural network (CNN) in addition to the aforementioned CRF to obtain accurate segmentations.

In addition to research papers, there is some commercial software available which locates hematoma in CT-scans. Accipio Ix [6] is one such example, and provides a method for location by circling the hematoma in a scan. The difference from our work and the above examples comes from the fact that we are interested in both finding the location and the shape of the hematoma. Chilamkurthy et al provides a purely classification algorithm on an image level, and as such have found success using a deep CNN. Accipio Ix on the other hand provides hematoma location, however not the shape and pixel values like we are interested in studying.

2 Methods

So far, we have studied 2 different methods to achieve our desired goals for this project. The first is an FCN, a trained model that provides for a rough segmentation, and the second a CRF, a statistical algorithm for result optimization. We found that applying these 2 results together have provided the most accurate segmentation results.

2.1 Fully Convolutional Network (FCN)

Based on the 2016 paper [7] by Evan Shelhamer, Jonathan Long, and Trevor Darrell. It consists of a neural network that is composed of fully-convolutional layers without any fully-connected layers. What this means is that each layer is learning filters, even the output layer. As a result, the network is able to perform classifications by pixel instead of by image. This is done by downsampling the input image into many learnable filters, then upsampling the filters to the same dimensions as the image and then using an activation function to evaluate the pixel-by-pixel output. The FCN architecture differs from the usual CNN through its upsampling step which allows the network to classify per pixel instead of per image, as can be seen in Figure 2. Using this upsampling step, we designed our

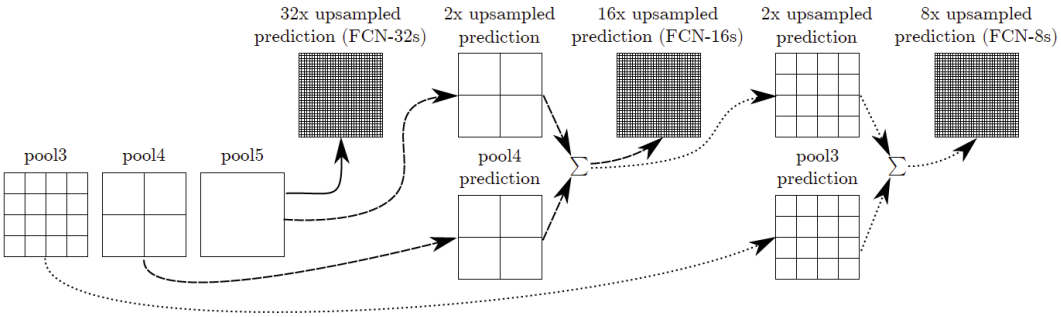

Figure 2: **FCN-32 vs. FCN-16 vs. FCN-8.** A few different kinds of upsampling methods that take and combine the output of different pooling layers and upsample them back to the dimensions of the input image. [7]

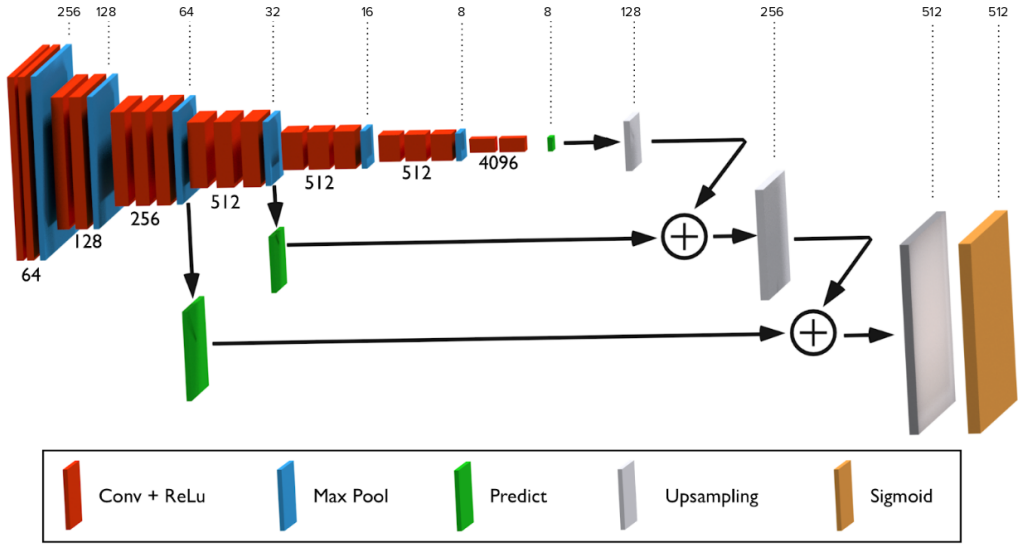

Figure 3: **Our FCN architecture in full.** It takes at takes a 512x512xN image, where N is an odd number of CT Scan slices. There are then 7 convolutional layers that produce the learnable filters of the model. We then take pooling layers 3,4, and 7 as part of the upsampling process to output a 512x512x1 matrix that is evaluated pixel-wise using the sigmoid activation function.

FCN (see Figure 3) to downsample with convolutional layers and then upsample using pooling layers 3,4, and 7 to output an image of the same dimensions of the input.

We used a sigmoid cross-entropy loss function to calculate the loss between the predicted and true masked pixels, which is defined by

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}[-y^{(i)}\log(h_\theta(x^{(i)})) - (1-y^{(i)})\log(1-h_\theta(x^{(i)}))]. \quad (1)$$

This function calculates the performance of the model whose output is a probabilistic number between 0 and 1. This is then applied to each of the 512x512 pixels in a given image to produce a total loss for each image.

2.2 Fully Connected Conditional Random Field (CRF)

Even though the dependencies between neighbouring pixels are not modelled directly by the FCN, the soft segmentation results are smooth since they share substantial spatial context [8]. However, noise such as imprecise manually labelled image and imperferct convergence in training can still

result in some specious output. Therefore, A 2D fully connected CRF[ref] is applied to refine the resulted segmentation from FCN and generate the hard segmentation maps for each CT-scan.

Fully connected CRF establishes pairwise relations between each pair of pixels in terms of their original grey scales and positions in the image. In the fully connected pairwise CRF model. The corresponding Gibbs energy is

$$E(\mathbf{x}) = \sum_i \psi_u(x_i) + \sum_{i<j} \psi_p(x_i, x_j), \tag{2}$$

where i and j range from 1 to 512. The unary potential $\psi_u(x_i)$ is the negative log-likelihood $\psi_u(x_i) = -\log P(x_i|\mathbf{I})$ where $P(x_i|\mathbf{I})$ is the result value from FCN for pixel i. The pairwise potentials $\psi_p(x_i, x_j)$ have the form

$$\psi_p(x_i, x_j) = \mu(x_i, x_j) k(\mathbf{f}_i, \mathbf{f}_j), \tag{3}$$

where $\mu(x_i, x_j)$ is $k(\mathbf{f}_i, \mathbf{f}_j)$ is defined with grey scale vectors I_i and I_j and positions p_i and p_j for pixel i and j:

$$k(\mathbf{f}_i, \mathbf{f}_j) = w^{(1)} exp\left(-\frac{|p_i - p_j|^2}{2\theta_\alpha^2} - \frac{|p_i - p_j|^2}{2\theta_\beta^2}\right) + w^{(2)} exp\left(-\frac{|p_i - p_j|^2}{2\theta_\gamma^2}\right). \tag{4}$$

In (3), parameters θ_α, θ_γ and θ_β control the degree of nearness and similarity. The parameters are preset with values learned from data in previous work.

2.3 Pipeline

For our most accurate results, we exploit the advantages that each of the above 2 methods provides. FCNs produce masks that generally describes the location, shape, and area. While this method independently struggles to match an accurate prediction mask, we found that it is able to consistently locate the rough pixel values in which the hemorrhage is contained in. In addition, we find that even subtle hematomas can be located effectively, as seen in the results section. CRFs then take as input the FCN predicted mask and input CT-scan, then improve on the mask by matching it more accurately to the true mask. The FCN is the only model which is trained through labelled data, CRF is only used as a post-processing optimization step. The general pipeline is described in Figure 4.

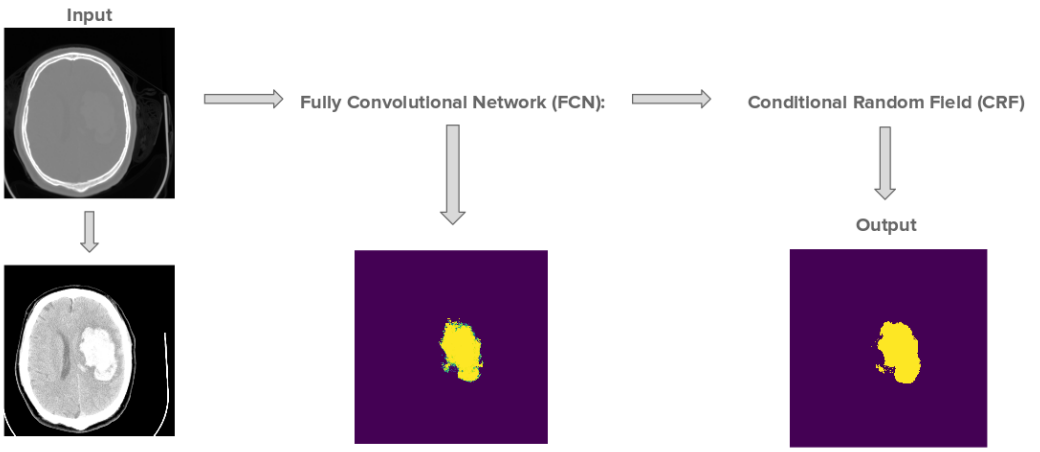

Figure 4: **General pipeline of our program.** The FCN is provided as input the true color CT scan, and provides a rough predicted mask. This predicted mask in addition to the true color scan is applied to the CRF algorithm, which improves results from the FCN.

3 Results

In total, we trained for 200 epochs on about 3600 image scan/hematoma mask pairs. The accuracy is the ratio of correctly classified pixels to incorrect pixels. The loss is calculated using sigmoid

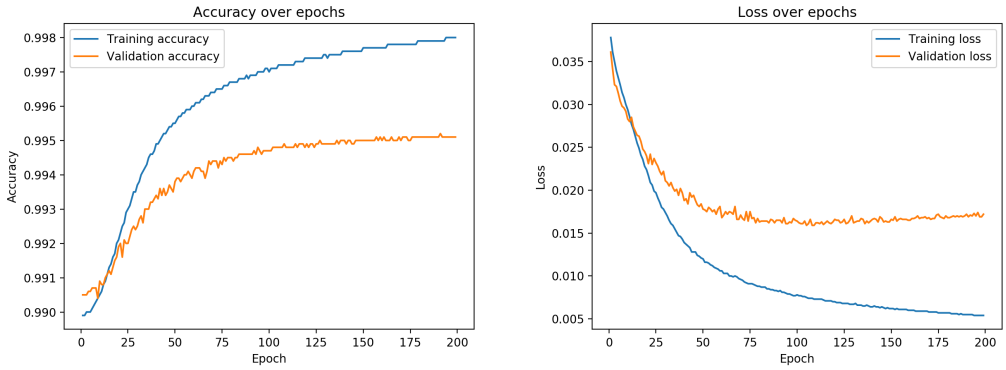

Figure 5: **Pixel accuracy vs epochs and loss vs epochs graphs.** At the first epochs, we find that the accuracy starts high and loss start low. The cause of this is since most of the pixels values are black (or not hematoma) to begin with, the model has a easy time predicting everything as being black to start. It isn't until later on, after about 25 epochs, that we see the model learning that actual hematomas are present.

function on each corresponding pixels from the true/predicted mask. We find that accuracy levelled off after about 100 epochs for the validation set, however continued to rise for the training set. This suggests that the model is overfitted slightly, and more training data or a simpler network with less trainable parameters could improve upon this (See Figure 3 for the current network configuration).

After training the FCN, we run the CRF method to improve upon the results, and used a Jaccard score metric to test for improvement (5). In total, we find on average that the score improves by 20.5% after employing this statistical optimization on the results provided solely by the FCN. Figure 7 shows 3 examples of CT-scans with their corresponding true mask, FCN and CRF predicted masks. Jaccard score is calculated by taking the intersection between the predicted and true mask, divided by their union. The score is between 0 and 1, where 1 describes perfect match, and 0 is no match.

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{5}$$

3.1 Discussion

We find that FCN together with CRF method have produced the most accurate results. Using these 2 methods in succession achieved an average pixel-wise prediction accuracy of 99.5% and Jaccard score of 0.946. In general, we find that hematoma located in the central regions of the brain have a higher accuracy than peripheral hematoma. It has been found that primary intracerebral hemorrhages (PICH) account for about 80-85% of all intracerebral hemorrhages (ICH) [9], which are responsible for central bleeding in the brain. As a result, we expect the training data to be imbalanced to have more central hematoma, and thus the model to be biased towards expecting hemorrhages in these specific regions. In addition, the majority of labels are found within or in close proximity to the subcortical and basal ganglia regions [10], and so the model often struggles to locate hemorrhages found in a different vertical sections than where they are most commonly found at.

3.2 Applications

There are 2 primary applications which our work can be used for. The first is a program which can allow doctors to quickly and effectively locate hematoma in a given CT-scan without having to search themselves. The benefit of this type of program is that it would allow doctors to more easily find the exact location and shape of the blood pool, allowing them to more easily diagnose their patients. The second application is for research purposes. As of now, segmenting hematoma in CT-scans by hand is a laborious process that takes a very long time. Therefore, having a computer to handle this process quickly and accurately will have various uses, particularly to the medical group who supplied us the labelled data.

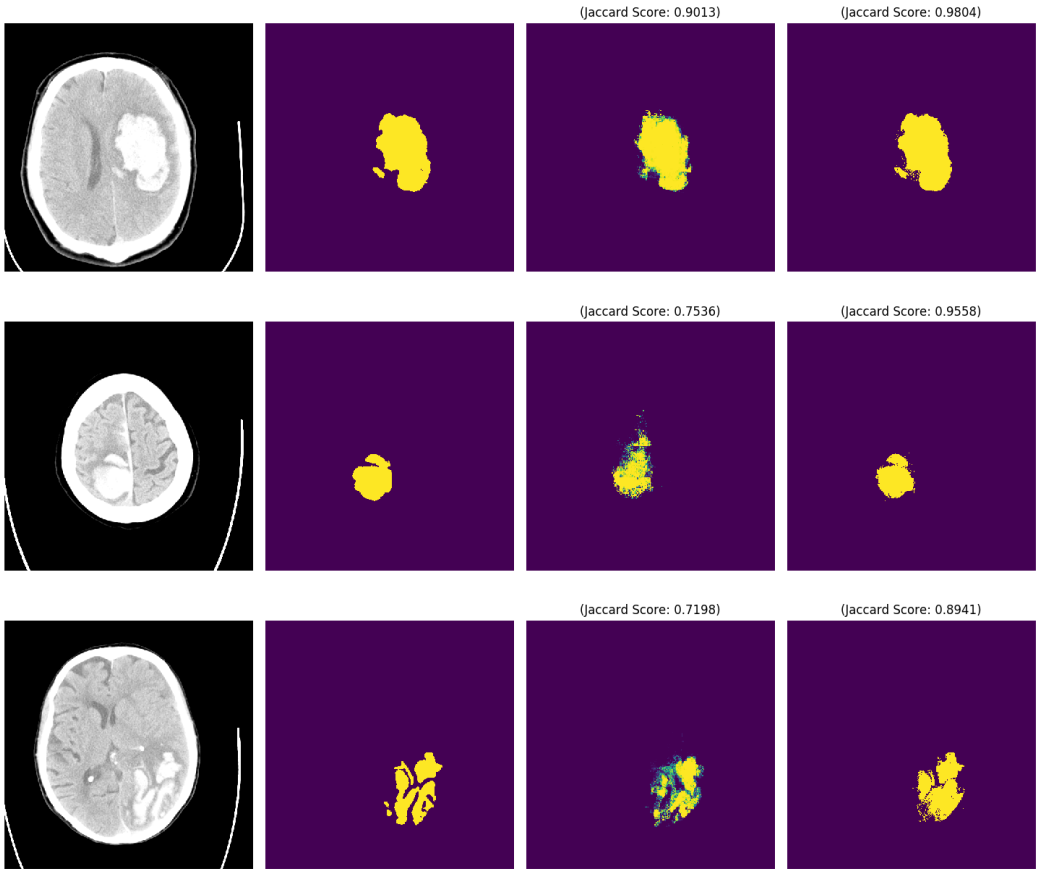

Figure 6: **3 examples demonstrating the results from our program.** The 1st column represents the CT-scan we sampled from. The 2nd corresponds to the predicted label supplied to us to train on. The 3rd and 4th denote the results from the FCN and CRF respectively. Jaccard scores are provided for the latter 2 columns.

As of now, hemorrhage detection and segmentation is a relatively new field which has come to rise upon the resurgence of machine learning. Therefore, there are very few programs that can handle the first application accurately, since medical programs are dependent on a very high degree of accuracy to become viable in clinical use. Therefore, we expect this work to be most useful in research purposes.

3.3 Future Work

CT-scans are inherently 3-dimensional in nature, since each scan consists of a series of constantly separated images from a patient's head [11]. Therefore, we intend to exploit this aspect of CT-scans by training on multiple images at once instead of a single scan/mask pair. In particular, we will feed 3 scans at once, or 1 images with the image directly above and below that given image. This will allow the model to take into account how adjacent layers might affect each other, and will therefore provide us with a rough 3-dimensional model from each scan series of each patient.

As well, we prefer the model to have a higher false positive than false negative rate. In other words, we prefer having the model predict that a hematoma is present when it in fact is not, more often than predicting one is not present, when it is. We will therefore try to increase and optimize the weights of the hematoma pixels, which will allow the model to be more inherently dissuaded from predicting false negative results.

Finally, we have worked closely with T4-Hematoma, who have worked on predicting the expansion of hemorrhages given CT-scans and the hematoma mask. As a first step in their pipeline, they require

that the labelled data be present before training their model. As of now, they have access to many unlabelled CT-scans from around 600 patients, so we can supply our program to them which would act as a precursor for the model they have developed.

4 Conclusion

In conclusion, our goal was to detect the contours of hematoma present in a CT scan of the brain. This was a useful task because it could save medical specialists from the tedious task of segmenting hematoma by hand, and could possibly be used by doctors in the future to quickly and more confidently detect hematoma in the brains of their patients. To achieve this goal we used a Fully Convolutional Network (FCN) and and Conditional Random Field (CRF). The FCN allowed us generate a rough segmentation by mapping each input pixel to a predicted output, and the CRF took as input the predicted mask from the FCN and the CT scan slice and was able to optimize it to better fit the true hematoma mask.

With the 3600 image scan/hematoma mask pairs (88 patients) that we trained the FCN on, we were able to get to about a 99.8% pixel-wise accuracy on the training set and about 99.5% accuracy on the validation set when training for 200 epochs. This accuracy was calculated using the sigmoid cross entropy loss function that was mentioned above, and averaged on each of pixels in the true/predict mask. We also had an average jaccard score between the true mask and the outputted mask of 0.785 for the FCN and 0.946 for the CRF. Combining more data with more model optimizations in the future, we believe these numbers could get much higher than they currently are, and could possibly be used as tool for doctors and other medical specialists to segment/detect hematoma in stroke patients.

References

[1] Debraj Mukherjee and Chirag G Patil. Epidemiology and the global burden of stroke. *World neurosurgery*, 76(6):S85–S90, 2011.

[2] Biljana Kojic, Adnan Burina, Renata Hodzic, Zejneba Pasic, and Osman Sinanovic. Risk factors impact on the long-term survival after hemorrhagic stroke. *Medical Archives*, 63(4):203, 2009.

[3] National Cancer Institute. *Computed Tomography (CT) Scans and Cancer*. `https://www.cancer.gov/about-cancer/diagnosis-staging/ct-scans-fact-sheet`.

[4] Sasank Chilamkurthy, Rohit Ghosh, Swetha Tanamala, Mustafa Biviji, Norbert G Campeau, Vasantha Kumar Venugopal, Vidur Mahajan, Pooja Rao, and Prashant Warier. Deep learning algorithms for detection of critical findings in head ct scans: a retrospective study. *The Lancet*, 392(10162):2388–2396, 2018.

[5] Yabo Fu, Thomas R Mazur, Xue Wu, Shi Liu, Xiao Chang, Yonggang Lu, H Harold Li, Hyun Kim, Michael C Roach, Lauren Henke, et al. A novel mri segmentation method using cnn-based correction network for mri-guided adaptive radiotherapy. *Medical physics*, 45(11):5129–5137, 2018.

[6] Gene Saragnese. Accipio ix. `https://maxq.ai/`.

[7] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.

[8] Vladlen Koltun et al. Efficient inference in fully connected crfs with gaussian edge potentials. *Adv. Neural Inf. Process. Syst*, 4, 2011.

[9] DynaMed Plus. *Intracerebral hemorrhage*. `https://www.dynamed.com/topics/dmp~AN~T115590`.

[10] Daniel J. Bell. *Basal ganglia hemorrhage*. `https://radiopaedia.org/articles/basal-ganglia-haemorrhage-2`.

[11] Melissa Conrad Stöppler. *CT Scan (Computerized Tomography, CAT Scan)*. `https://www.medicinenet.com/cat_scan/article.htm#ct_scan_facts`.