CT SCAN-BASED INJURY DETECTION AND CLASSIFICATION SYSTEM

MACHINE LEARNING & PATTERN RECOGNITION

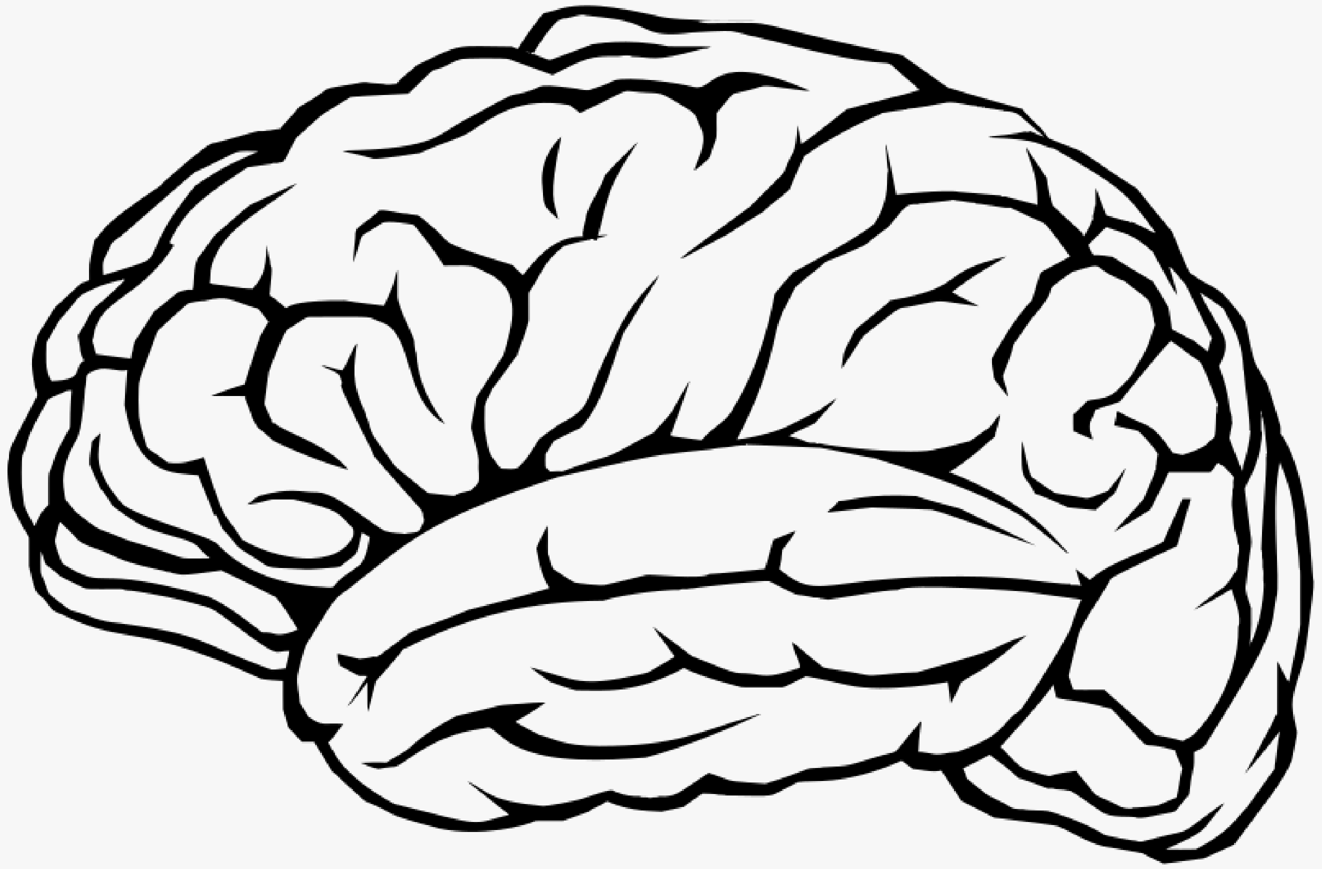
Our Team- Hey CT

ANAGHA VASISTA

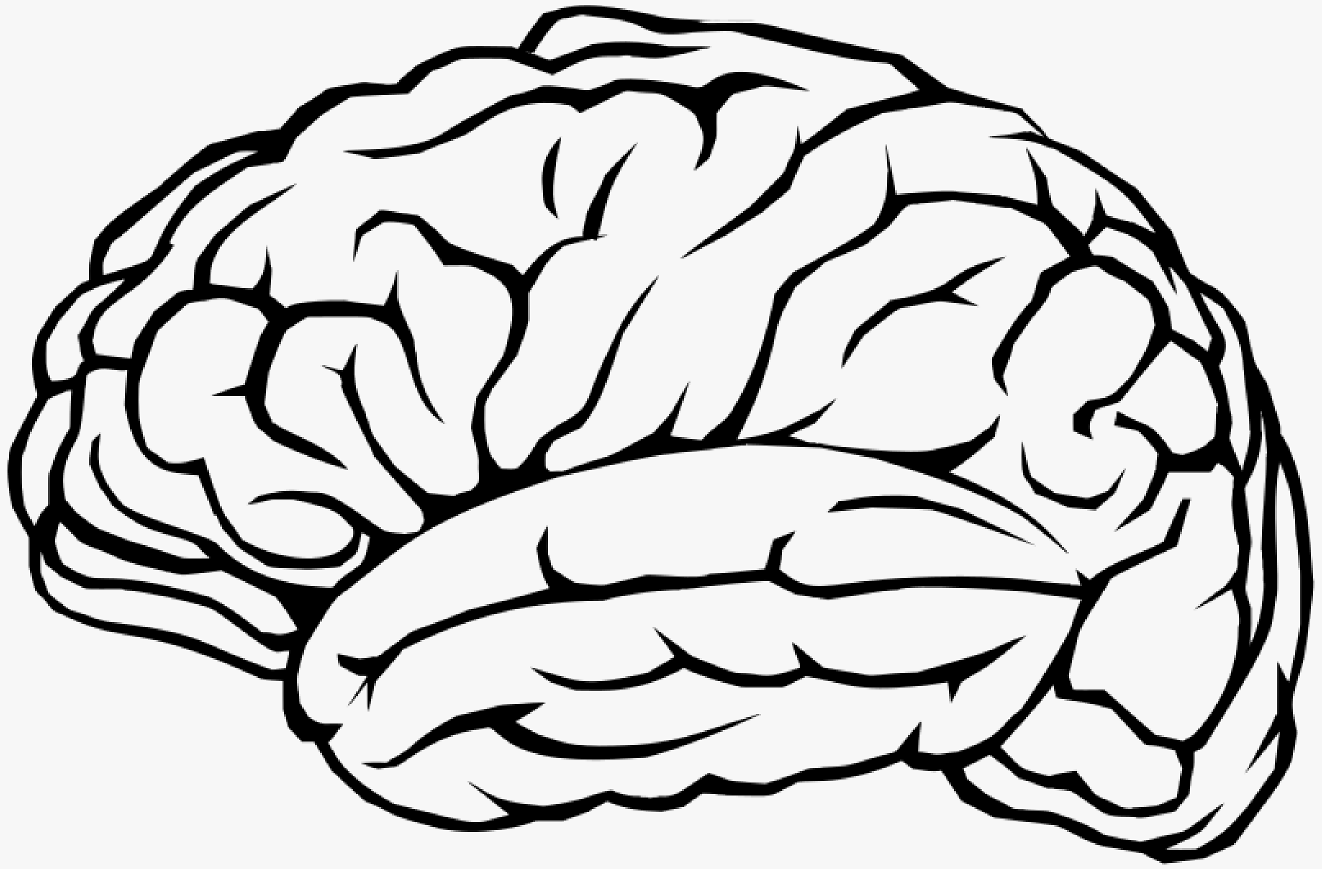
ONA DUBEY

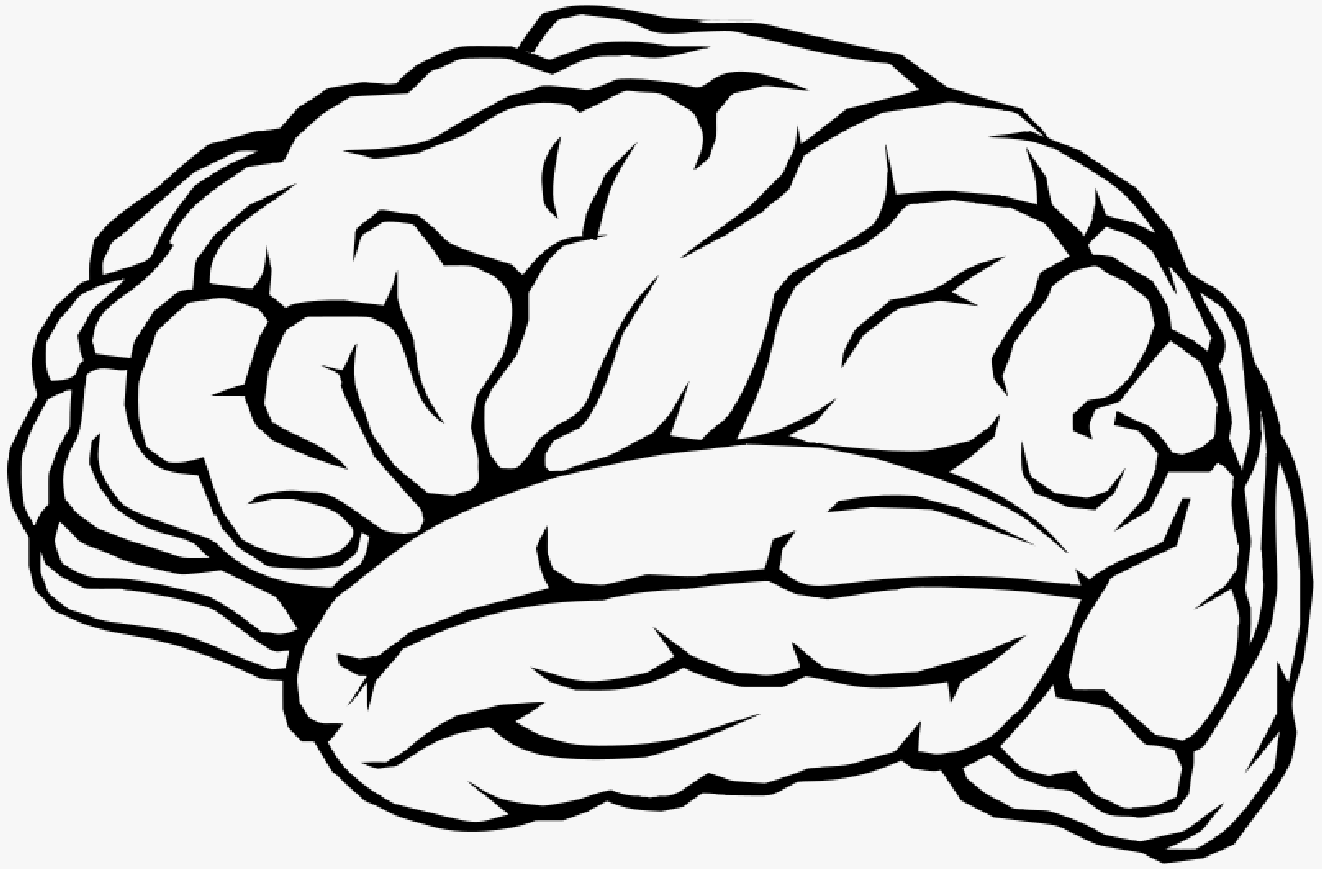
RONIT KADAKIA

Look! A Brain!



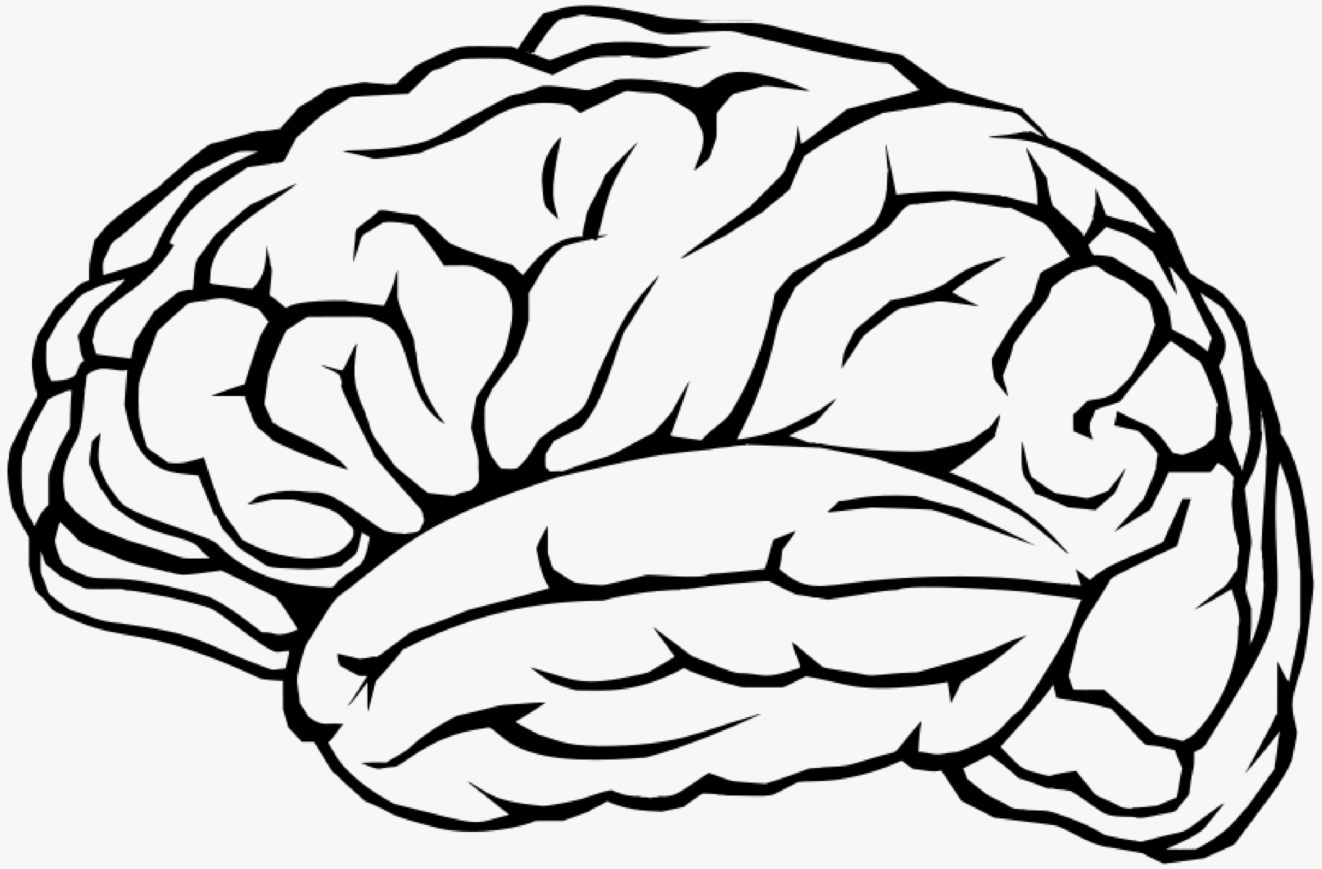
What if it's injured?

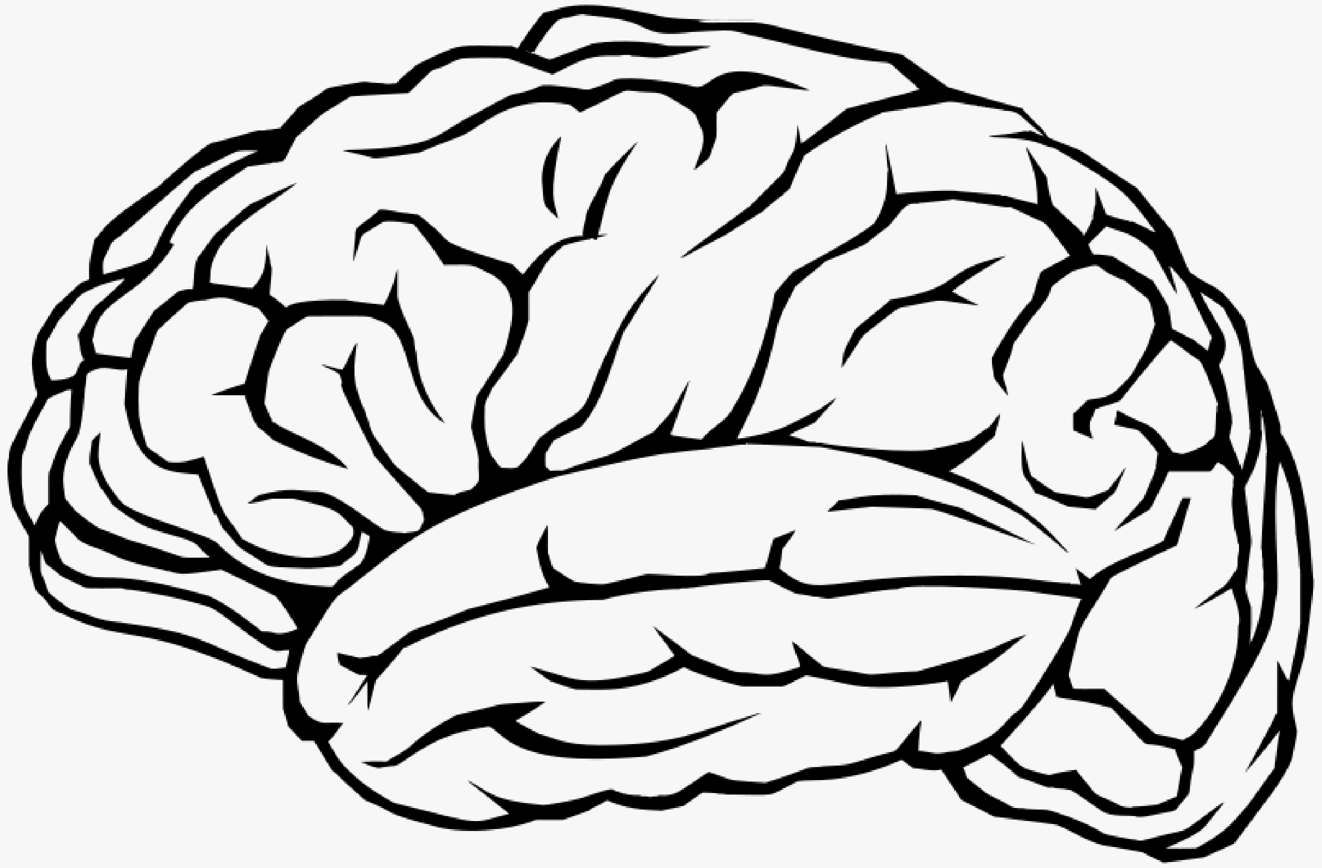


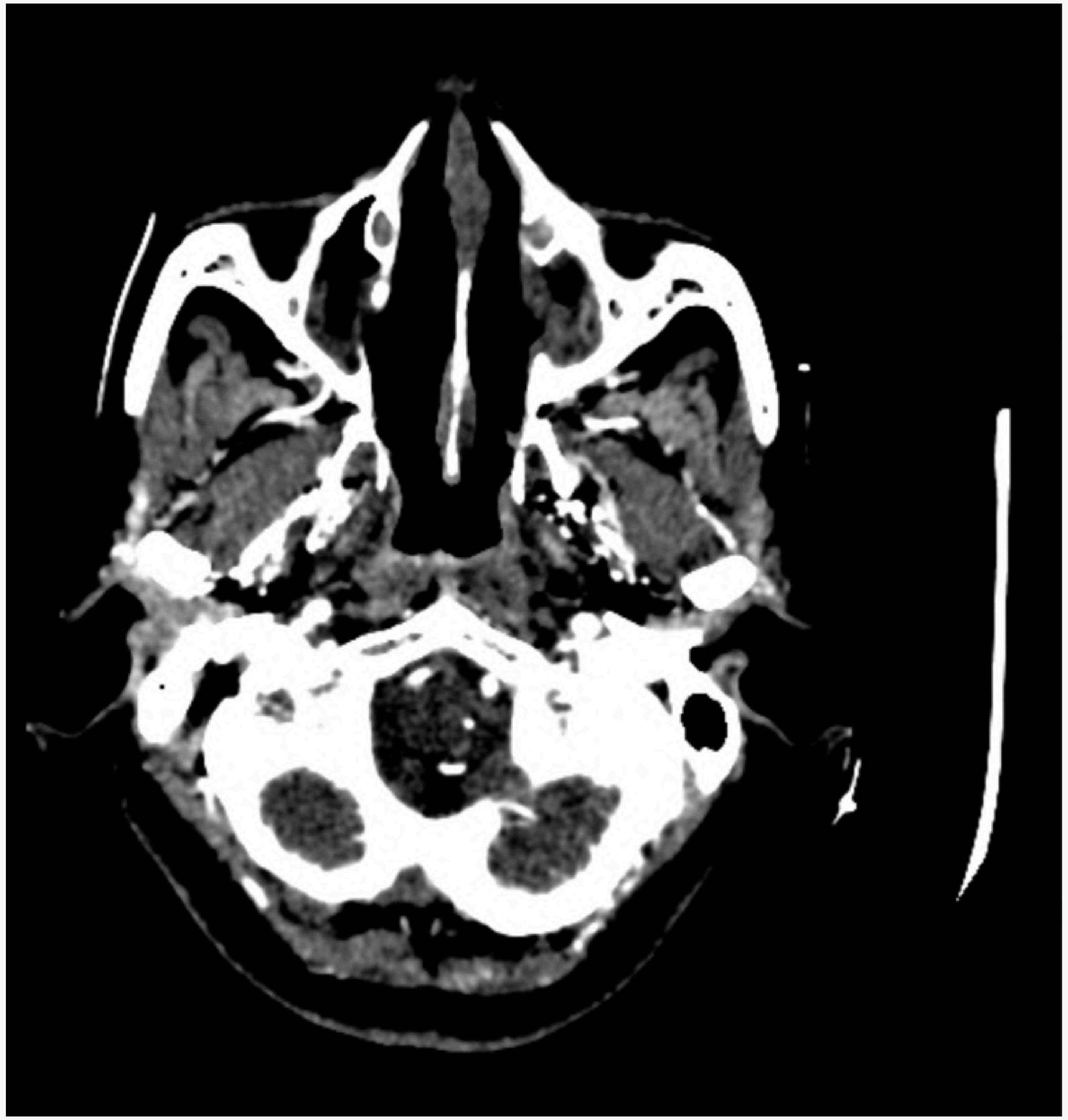


A doctor takes a CT
Scan to figure out
what's wrong with it...

What does that look like?

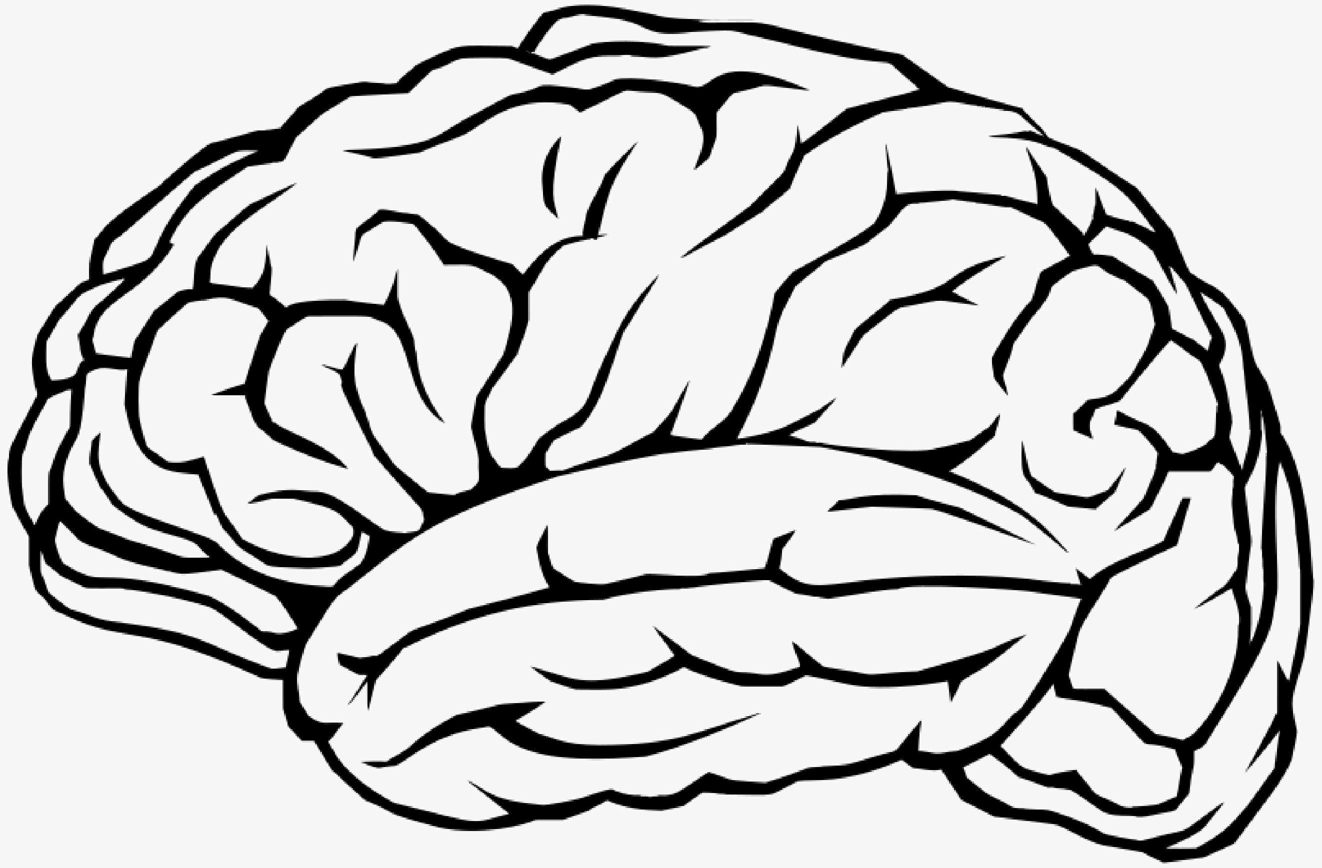


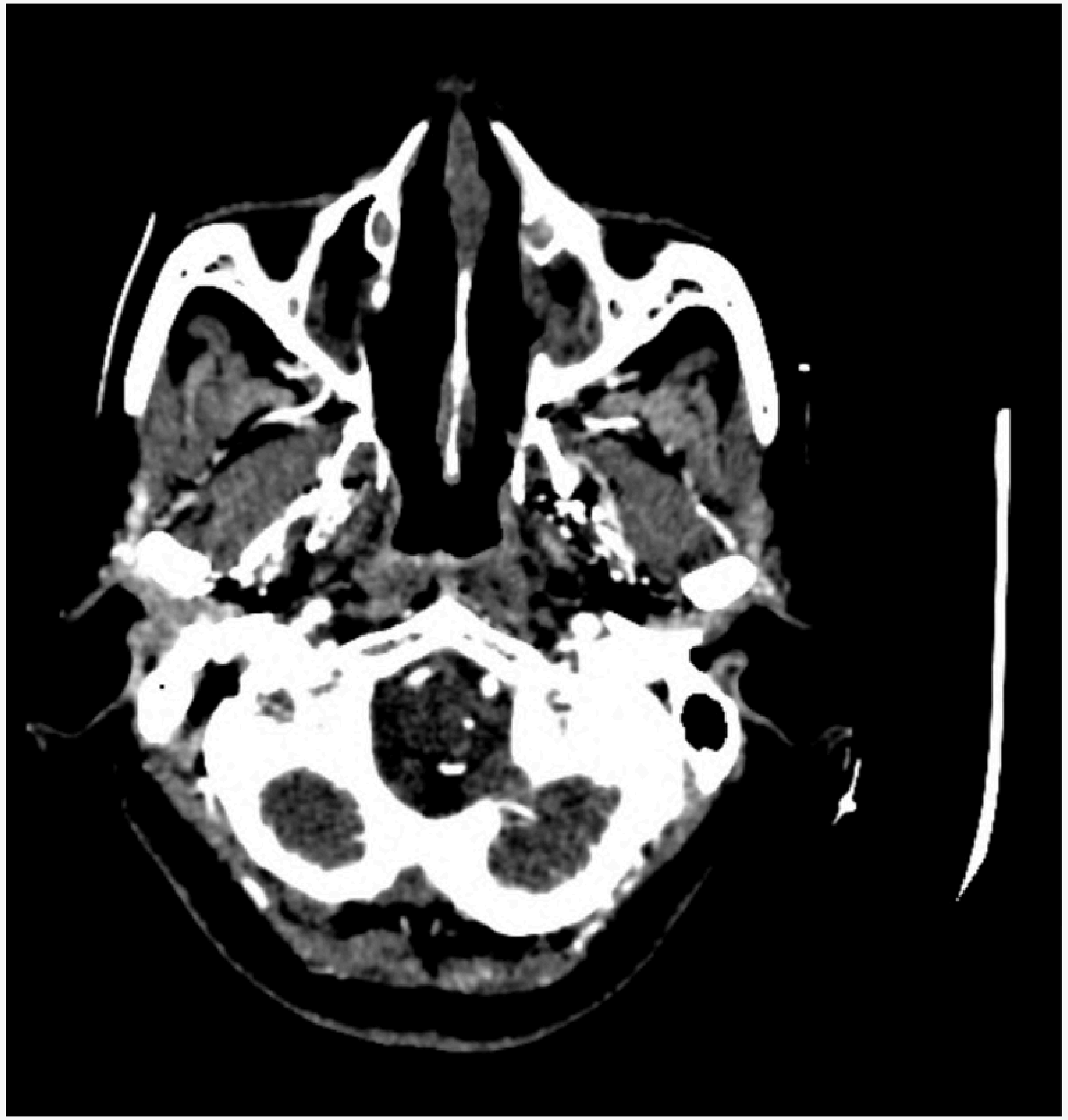


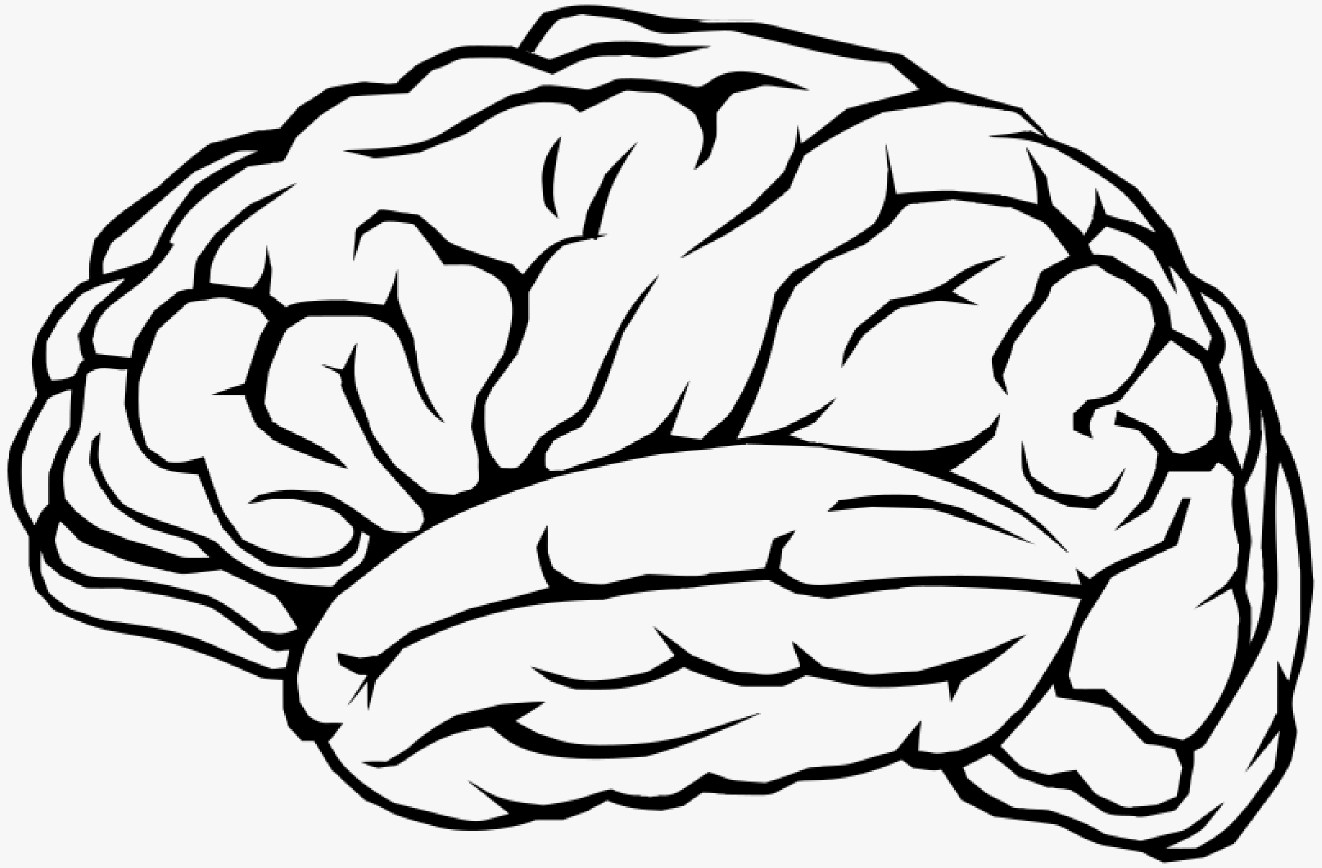


CT Scan

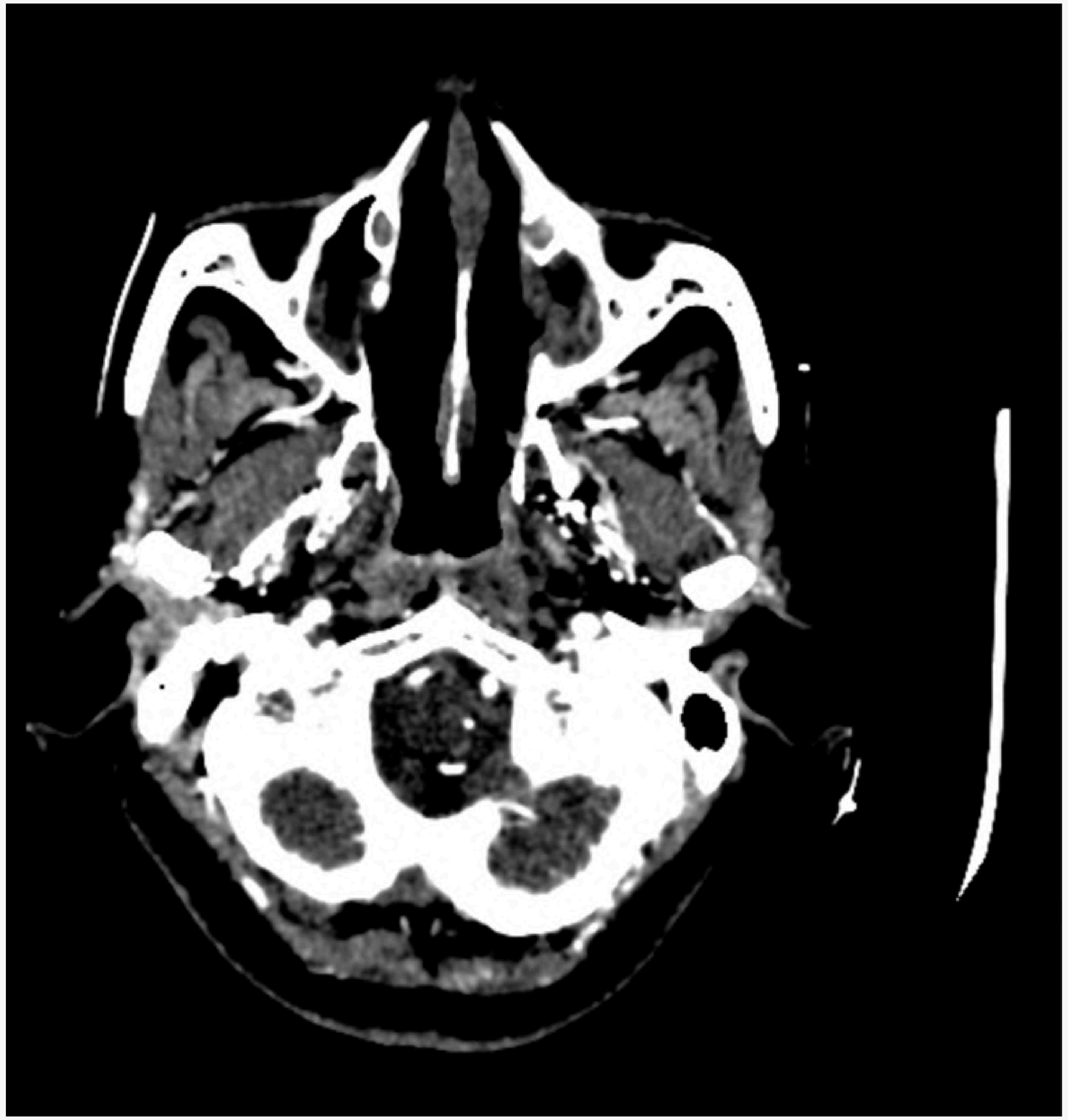
To an untrained eye, it's just a blob of black & white!



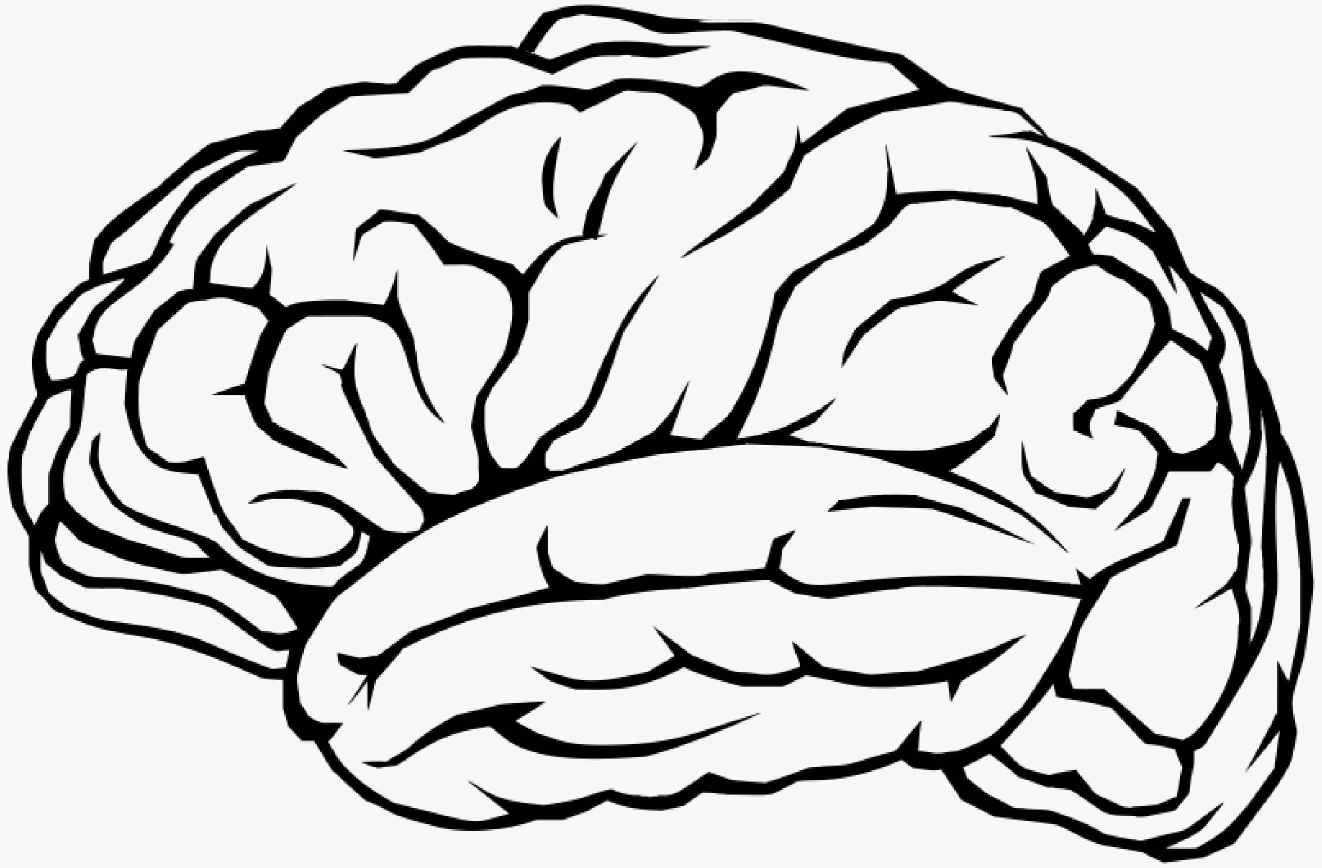


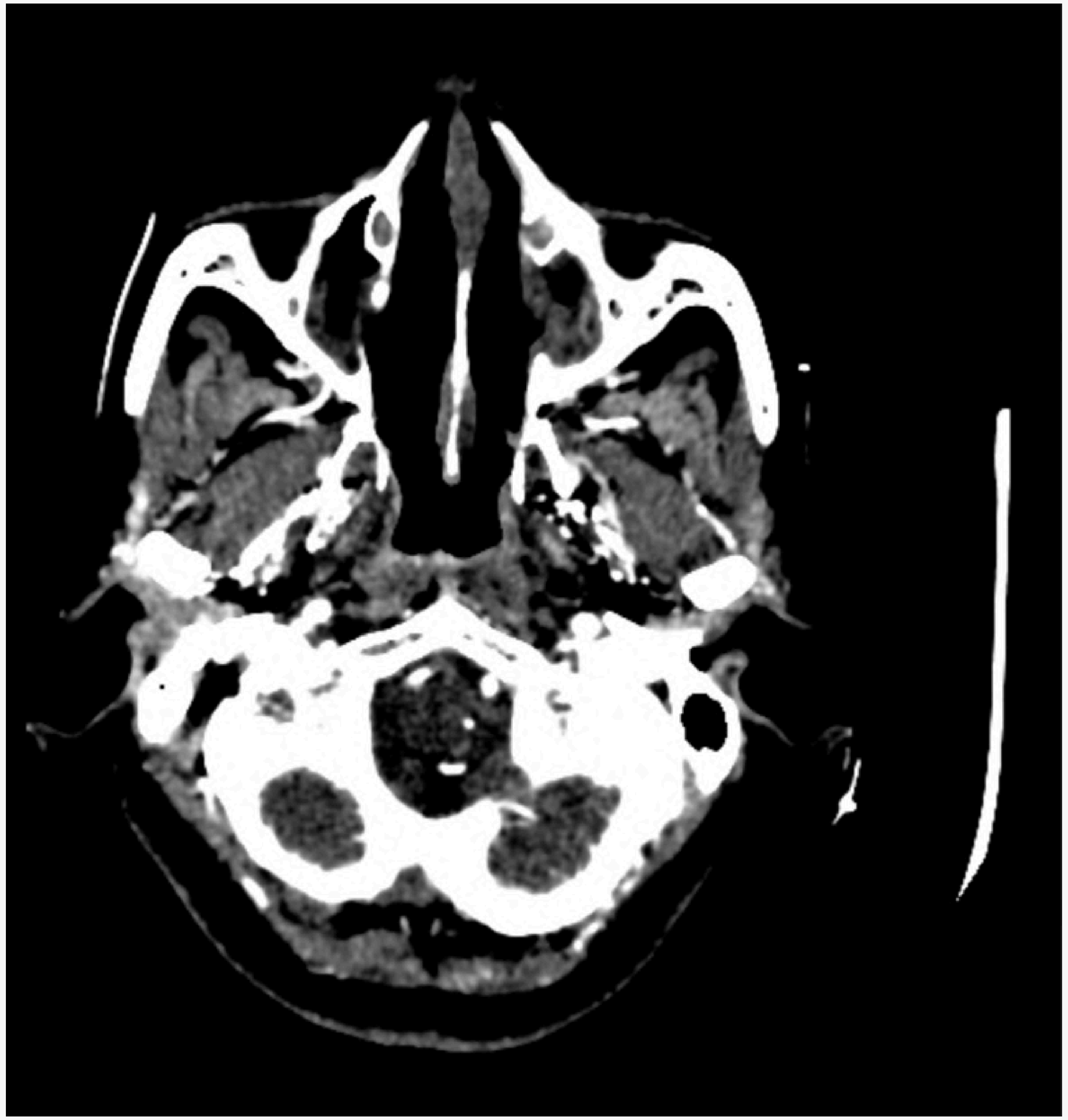


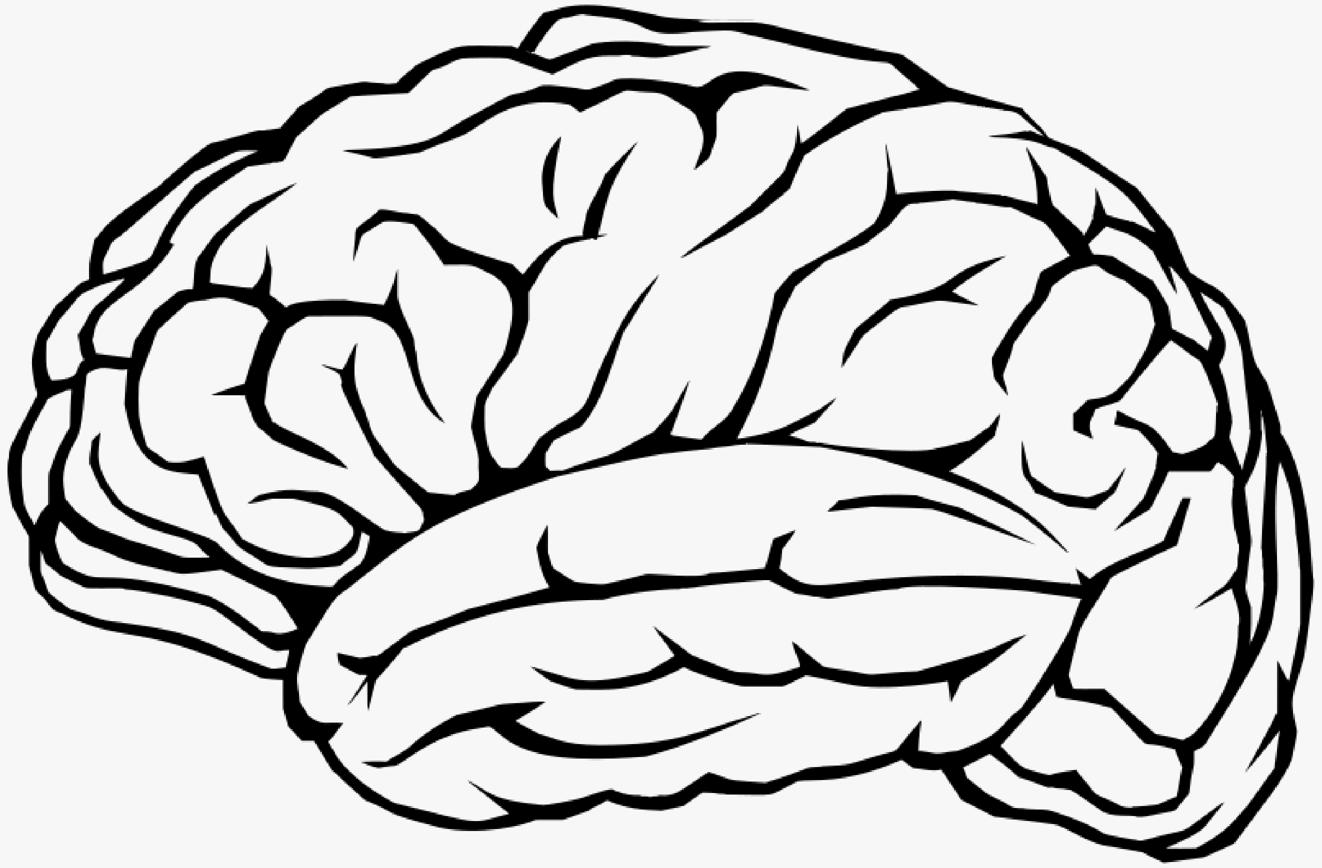
But, a doctor can identify what's wrong with just a look!



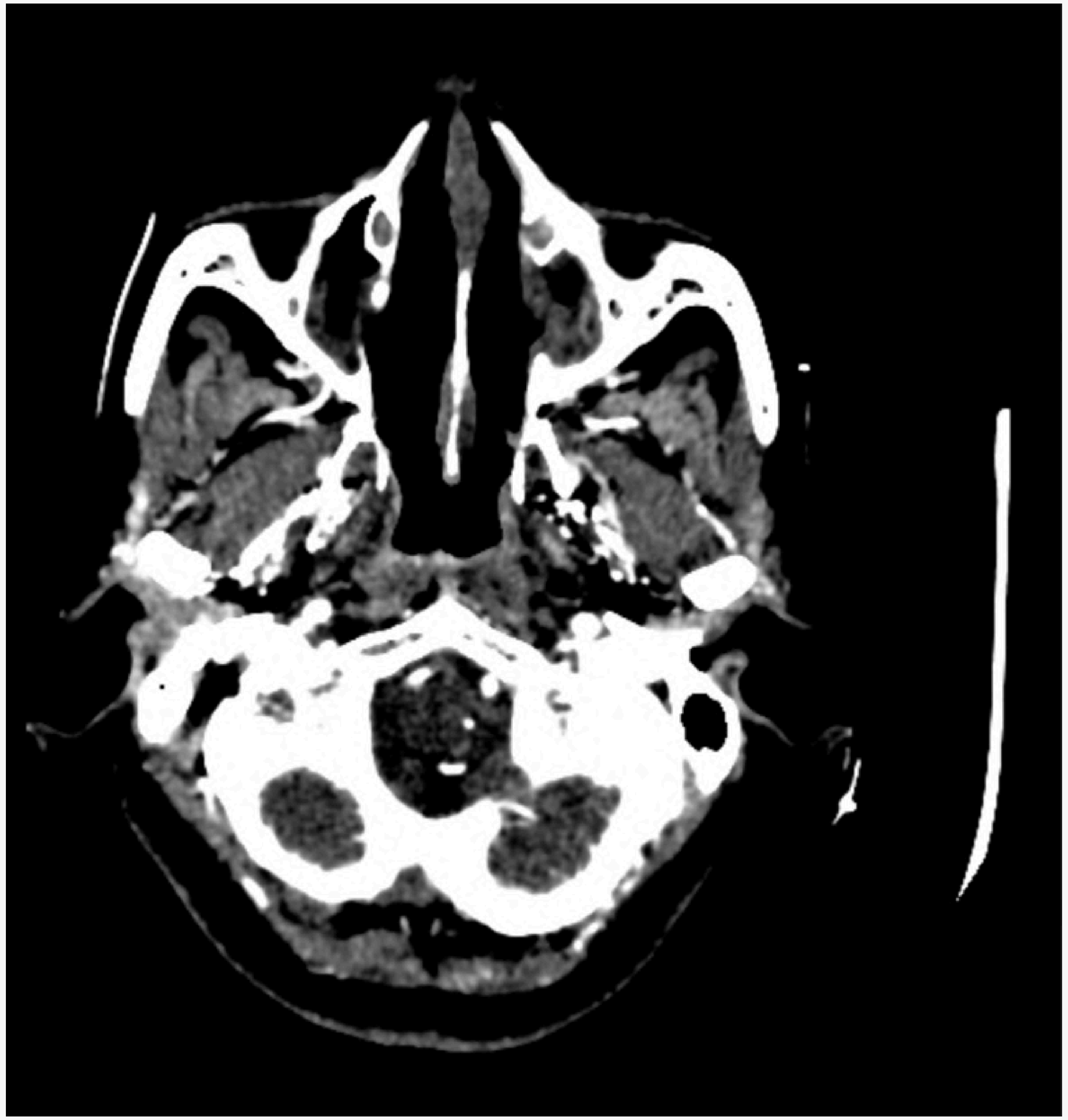
However, even doctors are humans and can make mistakes

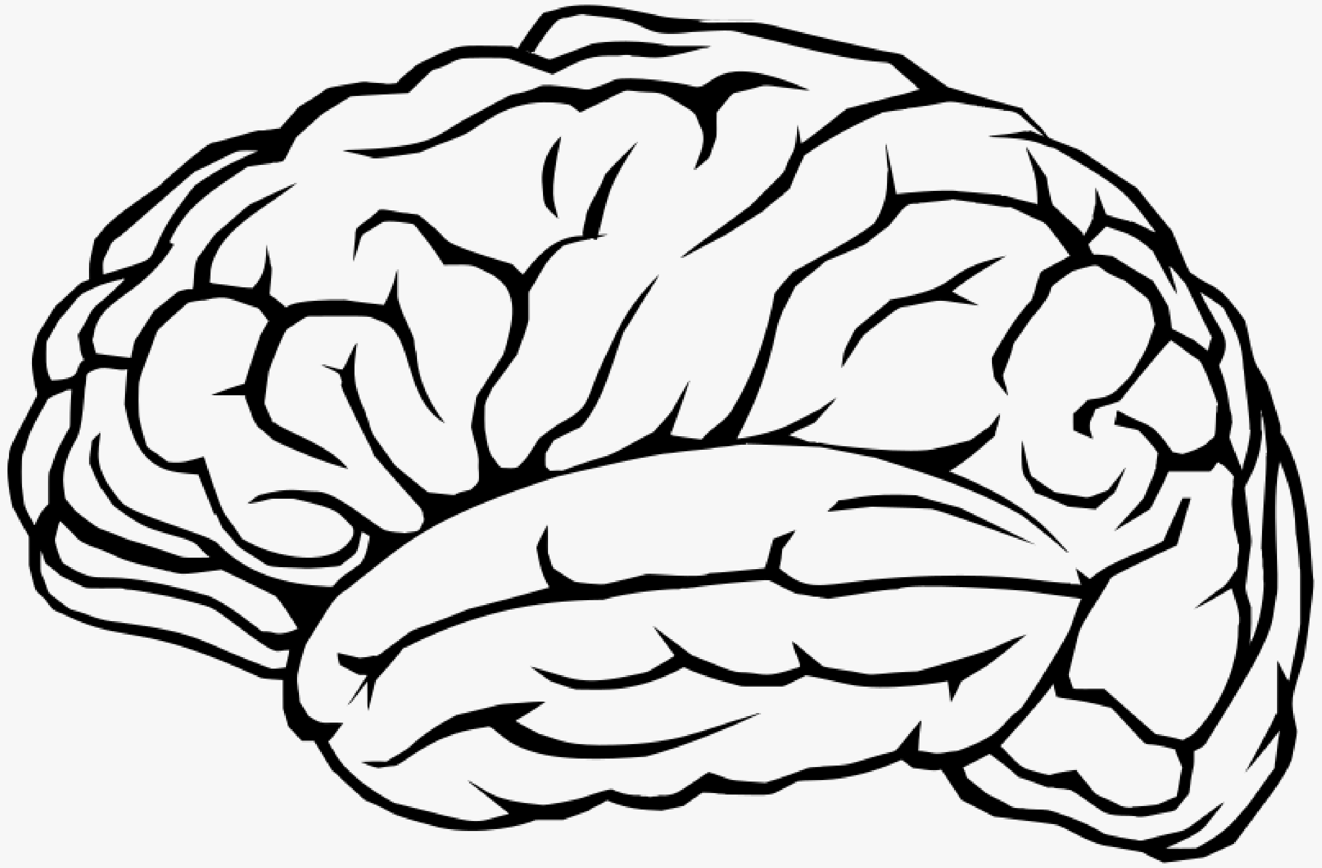


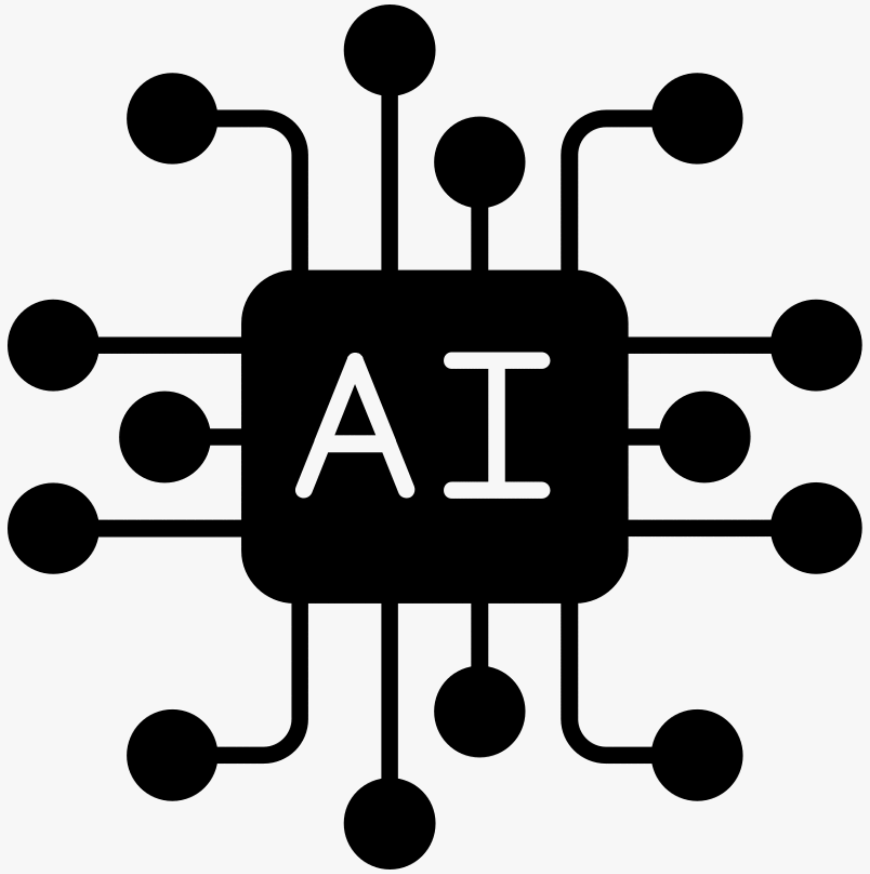


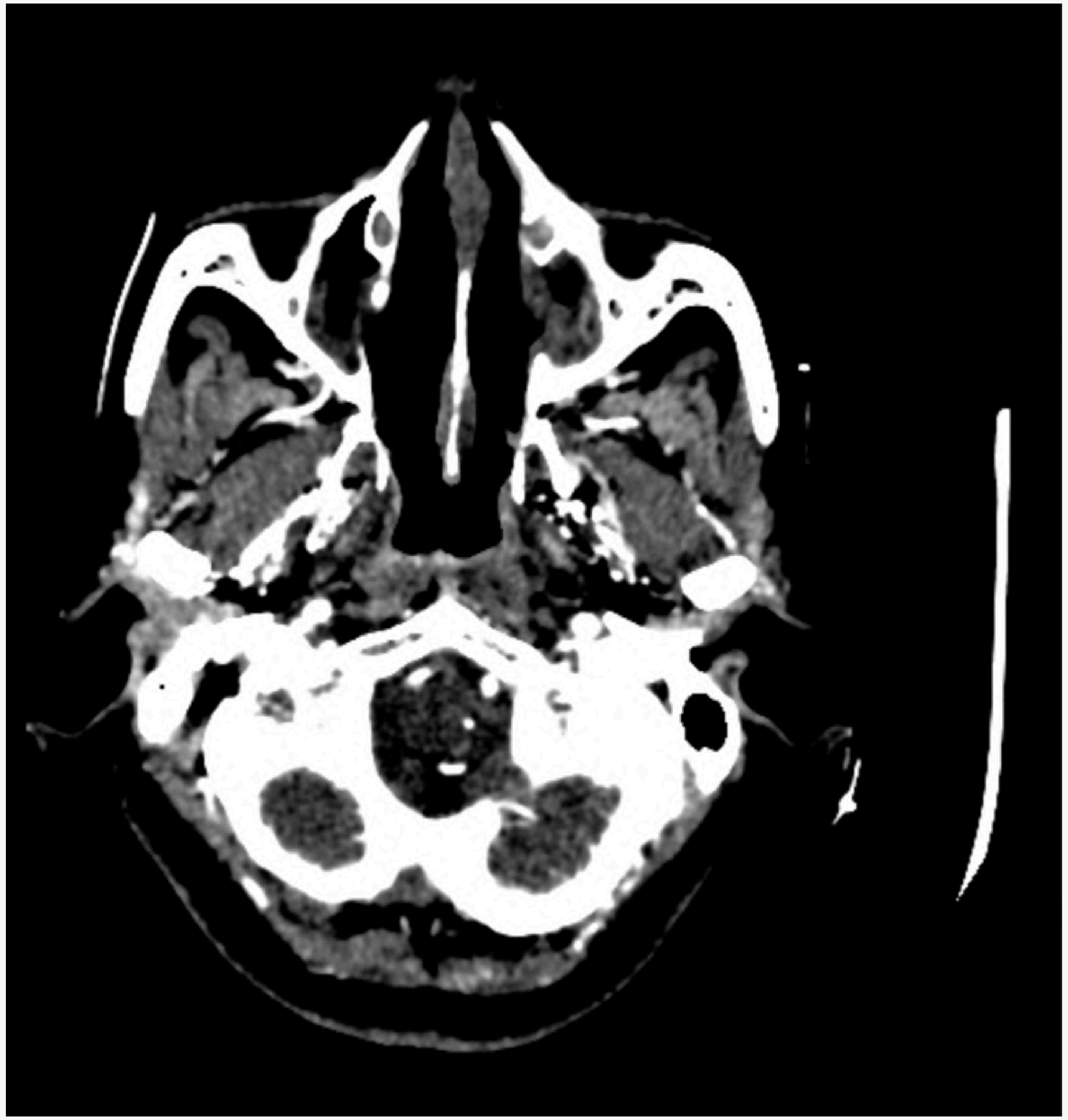


In fact, 13.6% of cases are misdiagnosed even by experienced doctors.



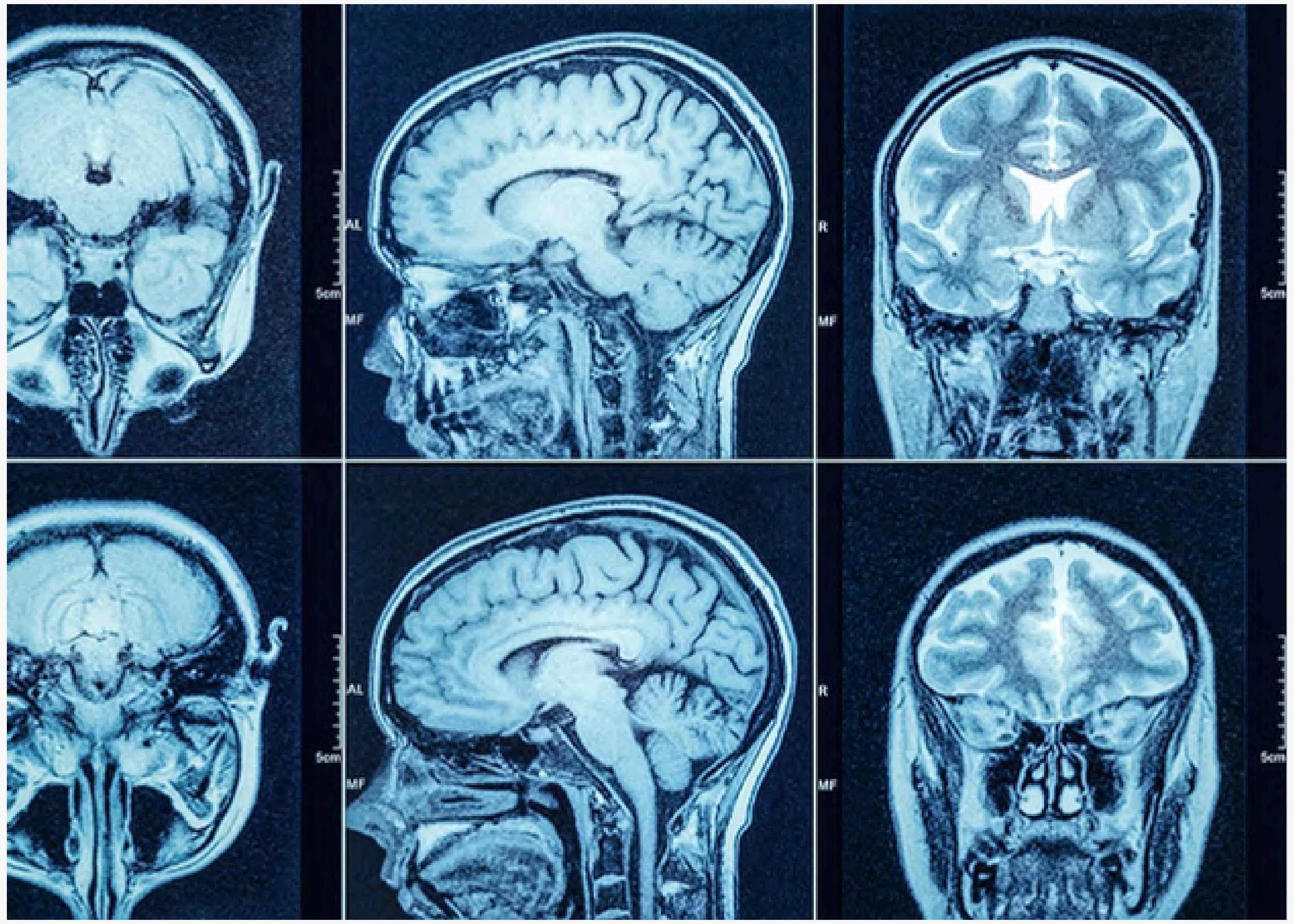






That is where we come in!
Harnessing the power of ML, we aim to aid the doctors identify these injuries faster, and easily.

CT Scans - The Basis For Diagnosis



A **computerized tomography (CT)** scan combines a series of X-ray images taken from different angles around your body and uses computer processing to create **cross-sectional images** (slices) of the bones, blood vessels and soft tissues inside your body.
CT scans can be used to identify disease or injury within various regions of the body.

1. Lvmc. (2021, December 4). 12 reasons you may need a CT scan - Lompoc Valley Medical Center. Lompoc Valley Medical Center. https://www.lompocvmc.com/blogs/2021/december/12-reasons-you-may-need-a-ct-scan/

CT Scans - The Current Problem

A snapshot of diagnostic error and misdiagnosis-related harm rates, highlighting the significant variability in the accuracy of medical diagnoses and the impact these errors can have on patient outcomes.

21.9%

Emergency Department Revisits (with repeat CT or MRI)

13.6%

Median across various conditions: Diagnostic error rate

11.5%

Radiologic Errors (of total diagnostic errors)

5.5%

Misdiagnosis-related harm rate

1. Newman-Toker, D. E., Wang, Z., Zhu, Y., Nassery, N., Tehrani, A. S. S., Schaffer, A. C., Yu-Moe, C. W., Clemens, G., Fanai, M., & Siegal, D. (2020). Rate of diagnostic errors and serious misdiagnosis-related harms for major vascular events, infections, and cancers: toward a national incidence estimate using the "Big Three." Diagnosis, 8(1), 67–84. https://doi.org/10.1515/dx-2019-0104
2. Ahn, Y., Hong, G., Park, K. J., Lee, C. W., Lee, J. H., & Kim, S. (2021). Impact of diagnostic errors on adverse outcomes: learning from emergency department revisits with repeat CT or MRI. Insights Into Imaging, 12(1). https://doi.org/10.1186/s13244-021-01108-0

The Problem Statement

Need for efficient and accurate analysis of head CT scans in hospitals.
By automating the process of generating reports for fractures, bleeding, and other abnormalities, our project aims to address this issue.

Applications

1. Faster Diagnosis:
Automating the analysis reduces time required leading to faster diagnosis
2. Consistent Reporting:
Incorporates multiple opinions,
More consistent and reliable diagnostic information (reducing human error/bias)
3. Resource Optimisation:
Radiologists' time can be redirected from routine to more complex cases requiring human expertise
4. Enhanced Patient Care:
Quicker and more accurate diagnoses
Timely treatment and interventions,

Impact & Significance

1.Improved Healthcare Accessibility:

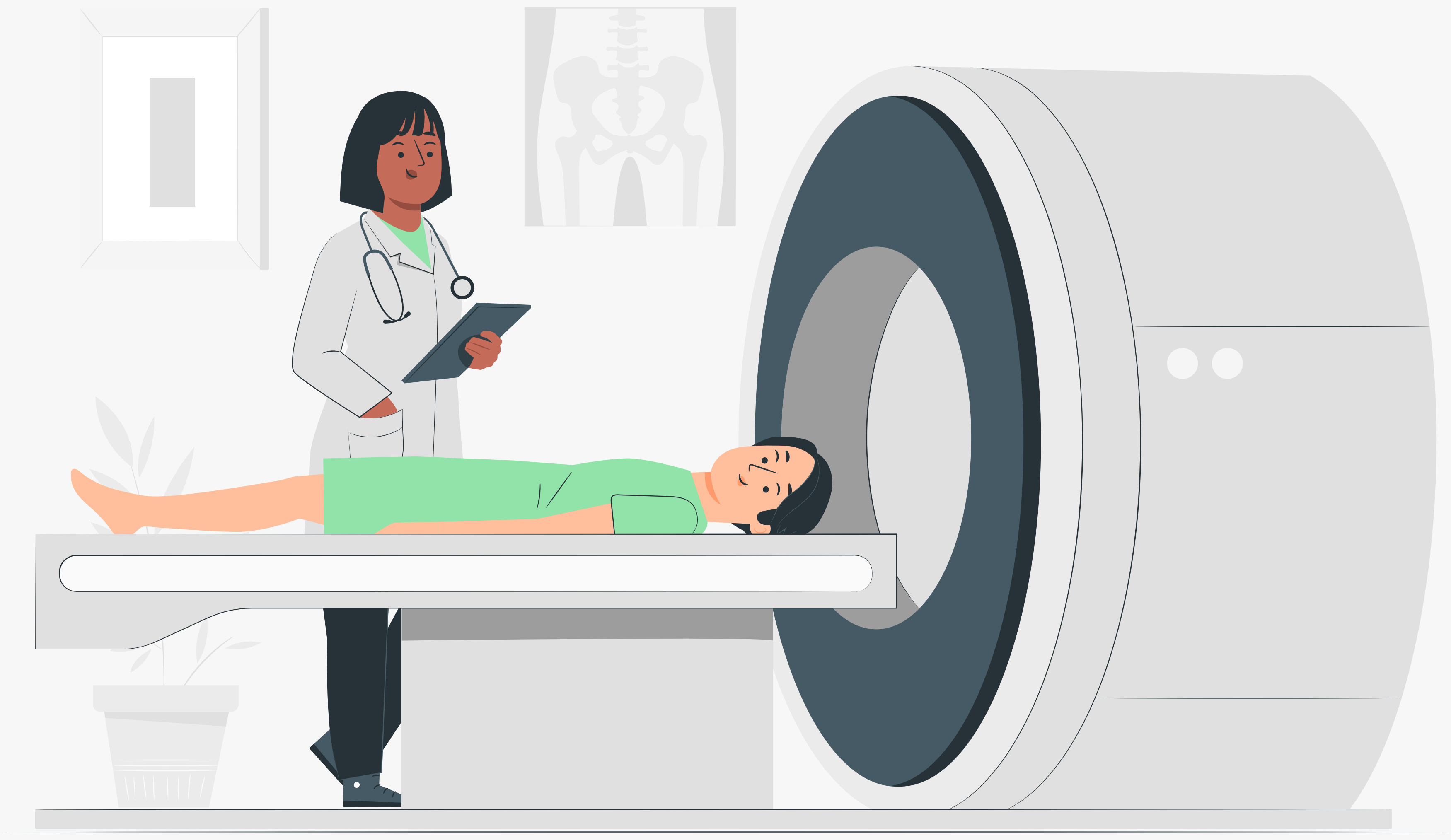
- providing faster diagnostic services.

2.Cost Savings:

- Automating repetitive tasks
- Optimizing resource utilization
- Reducing need for additional staff

3.Standardization of Care:

- Standardize the diagnostic process across healthcare facilities
- Ensuring consistent quality of care



Literature Review

Development and Validation of Deep Learning Algorithms for Detection of Critical Findings in Head CT Scans

- **Preprocessing**:
- Chose non-contrast axial series (Type of CT scan)- no contrast agents, taken parallel to the ground
- Soft reconstruction kernel (to convert raw data from imaging into a visual representation to be interpreted by medical professionals -- to enhance clarity/detail)
- Resampling Slice Thickness-- 5mm
- Resizing Images -- 224 × 224 pixels
- Windowing CT Densities - applied three different window settings to highlight specific features:
  - Brain-- highlights eg. brain tissue
  - Bone-- bony structures
  - Subdural-- skull and potential extra-axial bleeds
- Stacking Windows as Channels -- combined the images from the windows so it has multiple layers/channels

ML Models Used:

Modifications of ResNet18 to detect Hemorrhage & Hematoma

- **Intracranial Hemorrhage (ICH)**
- **Intraparenchymal Hemorrhage (IPH)**
- **Intraventricular Hemorrhage (IVH)**
- **Subdural Hematoma (SDH)**
- **Epidural Hematoma (EDH)**
- **Subarachnoid Hemorrhage (SAH)**

1. Chilamkurthy, S., Ghosh, R., Tanamala, S., Biviji, M., Campeau, N. G., Venugopal, V. K., Mahajan, V., Rao, P., & Warier, P. (2018, March 13). Development and validation of deep learning algorithms for detection of critical findings in head CT scans. arXiv.org. https://arxiv.org/abs/1803.05854

Literature Review

Automated and semi automated detection of findings from head CT scans have been studied by many groups.

- Grewal et al. developed a deep learning approach to automatically detect intracranial hemorrhages. They reported a sensitivity of 0.8864 and a positive predictive value (precision) of 0.8125 on a dataset of 77 brain CT scans read by three radiologists. However, the types of intracranial hemorrhage considered were not mentioned in their report.
- Traditional computer vision techniques like morphological processing were used by Zaki et al. and Yamada et al. to detect and retrieve scans with fracture respectively. Neither of the two studies measured their accuracies on a clinical dataset.
- Convolutional neural networks were used by Gao et al. to classify head CT scans to help diagnose Alzheimer's disease.
- Prevedello et al. evaluated the performance of a deep learning algorithm on a dataset of 50 scans to detect hemorrhage, mass effect, or hydrocephalus (HMH) and suspected acute infarct (SAI).

Our work is novel because we use head CT scans to detect and report on **six types of injuries in 1 model**

1. Grewal, M., Srivastava, M. M., Kumar, P., & Varadarajan, S. (2017, October 13). RADNET: Radiologist Level Accuracy using Deep Learning for HEMORRHAGE detection in CT Scans. arXiv.org. https://arxiv.org/abs/1710.04934
2. Zaki, W. M. D. W., Fauzi, M. F. A., & Besar, R. (2009). A new approach of skull fracture detection in CT brain images. In Lecture notes in computer science (pp. 156–167). https://doi.org/10.1007/978-3-642-05036-7_16
3. Wang, H., Ho, S., Xiao, F., & Chou, J. (2017, March 2). A simple, fast and fully automated approach for midline shift measurement on brain computed tomography. arXiv.org. https://arxiv.org/abs/1703.0079
4. Gao, X., Hui, R., & Tian, Z. (2017). Classification of CT brain images based on deep learning networks. Computer Methods and Programs in Biomedicine, 138, 49–56. https://doi.org/10.1016/j.cmpb.2016.10.007
5. Prevedello, L. M., Erdal, B. S., Ryu, J., Little, K. J., Demirer, M., Qian, S., & White, R. D. (2017). Automated critical test findings identification and online notification system using artificial intelligence in imaging. Radiology, 285(3), 923–931. https://doi.org/10.1148/radiol.2017162664

ResNet-18

- **ResNet (Residual Neural Networks)**
- The key idea behind ResNets is **residual learning**. Instead of trying to learn the exact mapping from input to output, ResNets learn the "residual" or the **difference between the input and the desired output**. This makes it easier for the network to learn complex mappings and prevents degradation in performance as the network gets deeper.
- ResNet-18 has 18 layers (17 convolutional and 1 fully connected). This makes it **faster to train** and less memory intensive while still benefiting from deep feature learning.
- In medical imaging tasks, ResNet architectures, such as **ResNet-18**, **ResNet-50**, etc., are commonly used for their **ability to handle complex features** and learn representations from medical images effectively.

1. Bangar, S. (2022, July 5). Resnet Architecture Explained - Siddhesh Bangar - Medium. Medium. https://medium.com/@siddheshb008/resnet-architecture-explained-47309ea9283d
2. He, K., Zhang, X., Ren, S., & Sun, J. (2015, December 10). Deep residual learning for image recognition. arXiv.org. https://arxiv.org/abs/1512.03385

VGG-16

- **VGG (Visual Geometry Group)**
- VGG-16 is a **deep CNN** with 16 layers (13 convolutional layers and 3 fully connected layers).
- This depth allows it to learn an extensive amount of features from images.
- It uses 3x3 **convolution filters** with stride 1 and maintains spatial resolution with padding, followed by max pooling. This simple, uniform structure makes it easy to understand and implement.
- While VGG-16 is **powerful**, it has a significantly **high computational** requirement.
- It has a very high number of parameters **(138 million)**, making it **computationally expensive** and **memory intensive**.
- This can lead to long training times and may require significant computational resources.

1. Simonyan, K., & Zisserman, A. (2014, September 4). Very deep convolutional networks for Large-Scale image recognition. arXiv.org. https://arxiv.org/abs/1409.1556v6
2. Research on vgg16 convolutional neural network feature classification algorithm based on Transfer Learning. (2021, November 3). IEEE Conference Publication | IEEE Xplore. https://ieeexplore.ieee.org/document/9652277

Our Data

Digital Imaging and Communications in Medicine - **DICOM** data from **qure.ai**

This is a medical imaging format which has a lot of information about the patient, machines, doctor, etc.

Data was collected from Centre for Advanced Research in Imaging, Neurosciences and Genomics (CARING), New Delhi, India.

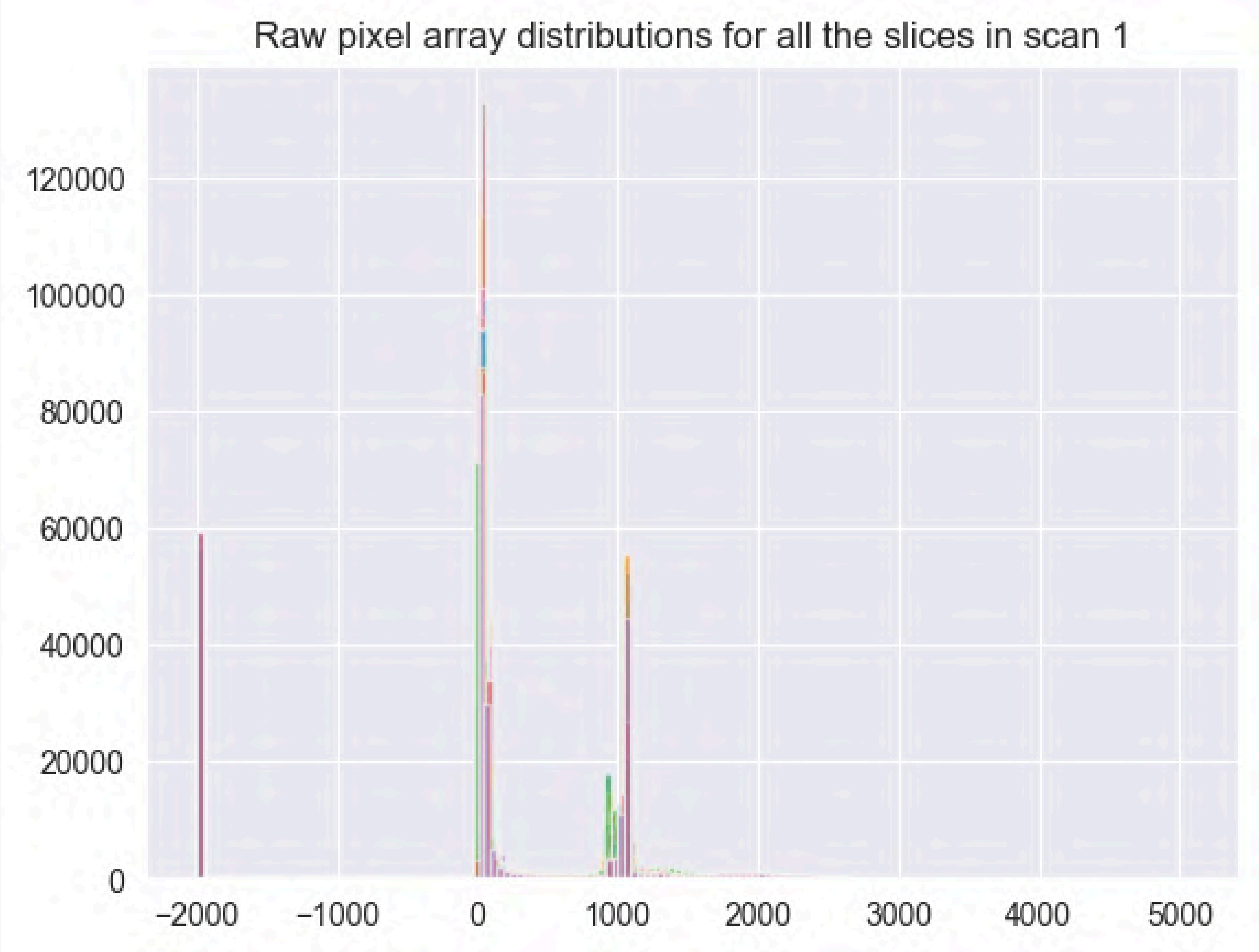
Our dataset has **491 CT scans, 193,317 slices**.

- Each scan has multiple slices, each of which is a .dcm file.
- We use **image information in the dcm file** + the **labels in a csv file.**

Ethical concerns were addressed by removing all identifiers from them.

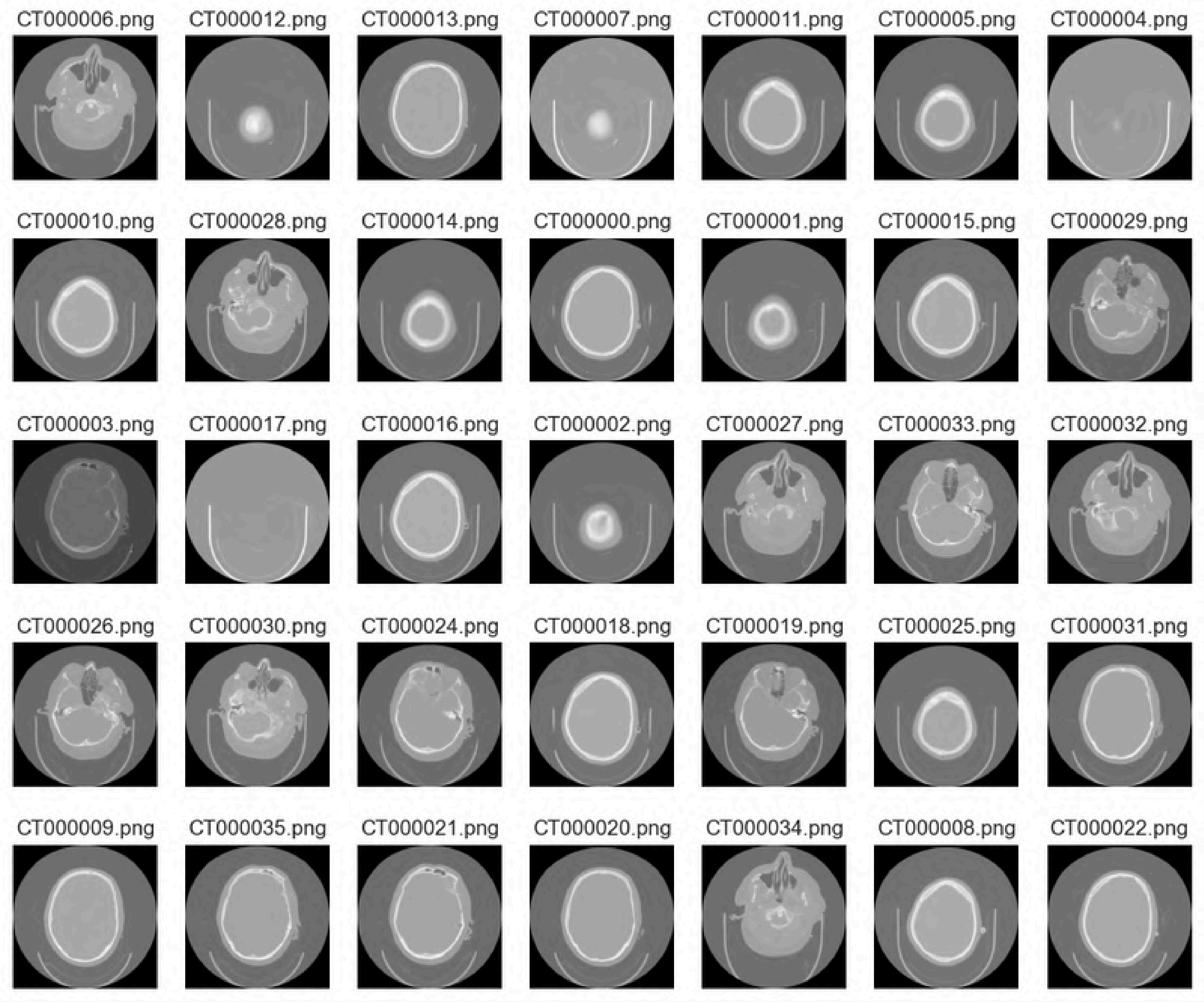
- To **visualise the pixel values**, we plotted the slices in a scan to see the distribution.

- We saw that they were in **HU units**, density ordered
  - -1000 is air
  - 0 is water
  - 1000 is bone.

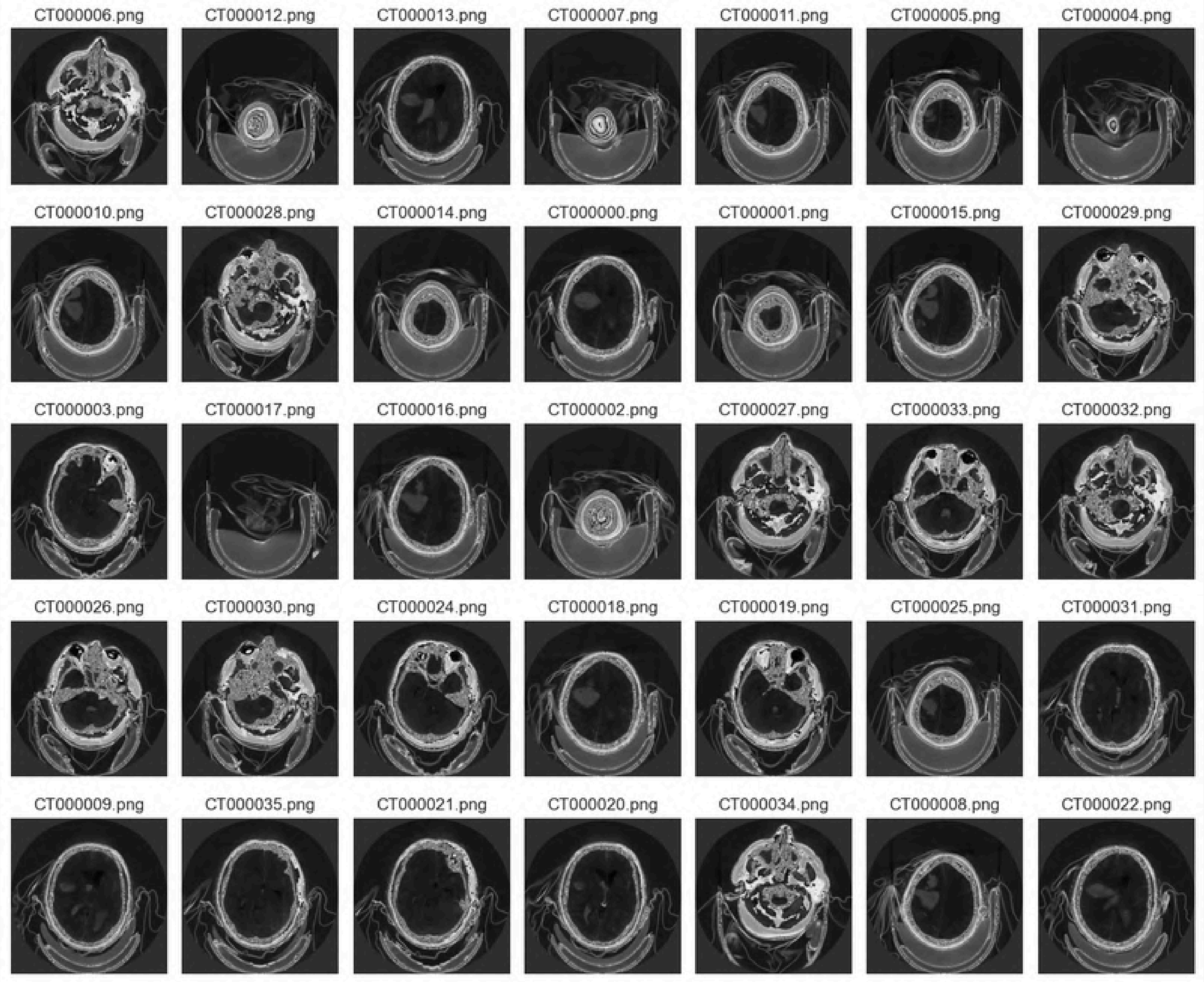


1. Chilamkurthy, S., Ghosh, R., Tanamala, S., Biviji, M., Campeau, N. G., Venugopal, V. K., Mahajan, V., Rao, P., & Warier, P. (2018, March 13). Development and validation of deep learning algorithms for detection of critical findings in head CT scans. arXiv.org. https://arxiv.org/abs/1803.05854

Data Preprocessing - Adjusting Contrast



raw image



preprocessed image

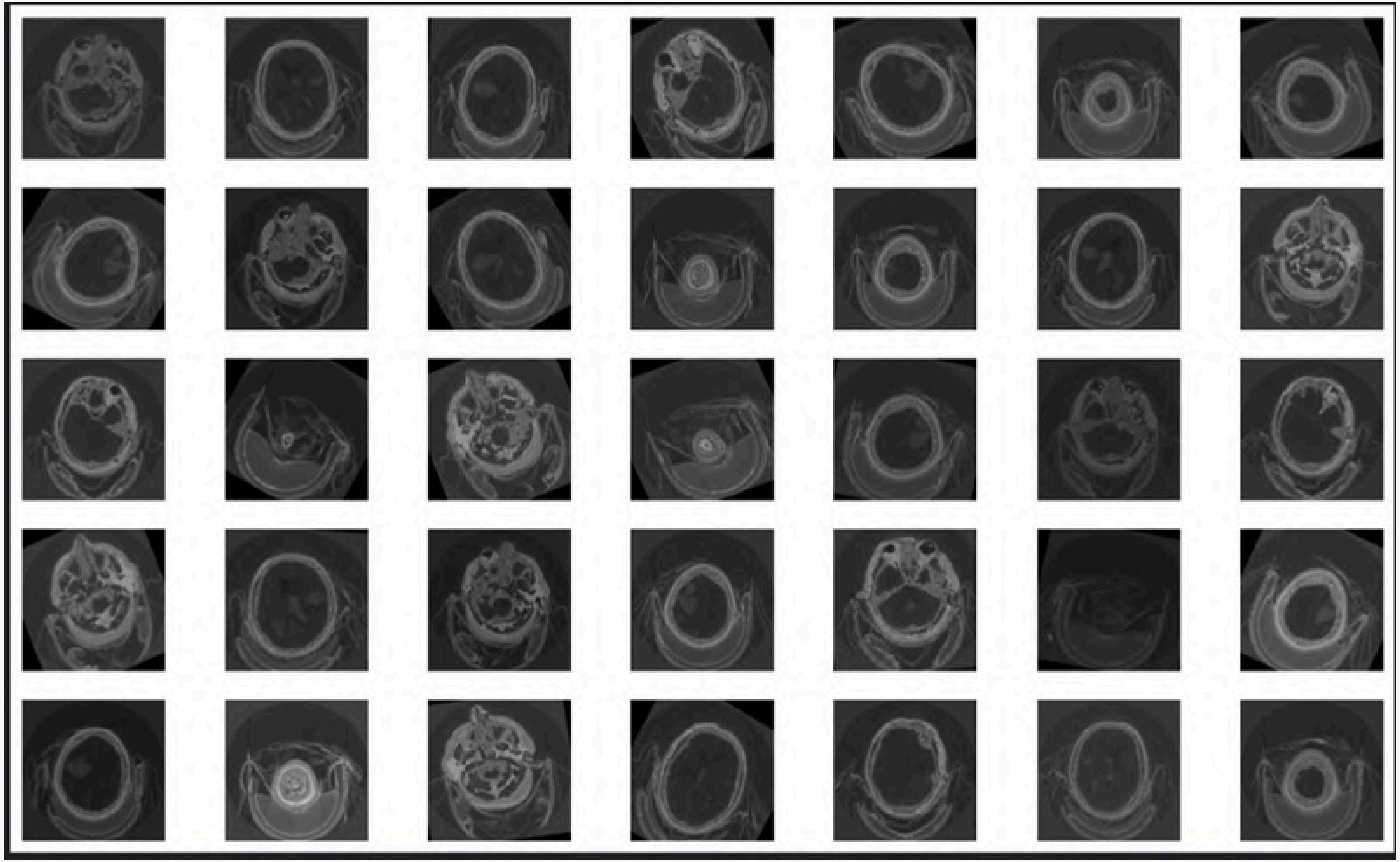
We convert the **dcm file into a numpy matrix** and then we increase its **contrast** to be able to view the abnormalities better. We made them into pngs and stored them so we can feed it to our machien learning model.

Data Preprocessing - Characteristic Bias

1. **Manufacturer's Model Name:** Different CT scanners may produce images with varying quality, contrast, and noise levels, which can affect model performance.
2. **KVP (Kilovoltage Peak)**: Variations in KVP affect image contrast, potentially biasing the model towards certain features.
3. **Window Center** and **Window Width**: These windowing parameters affect the brightness and contrast of the images, which can bias the model if not standardized.

Our dataset has been cleaned (from the source) and it has no characteristic biases, but usually with large datasets from the real world, these biases can exist.

Data Preprocessing - Image Augmentation



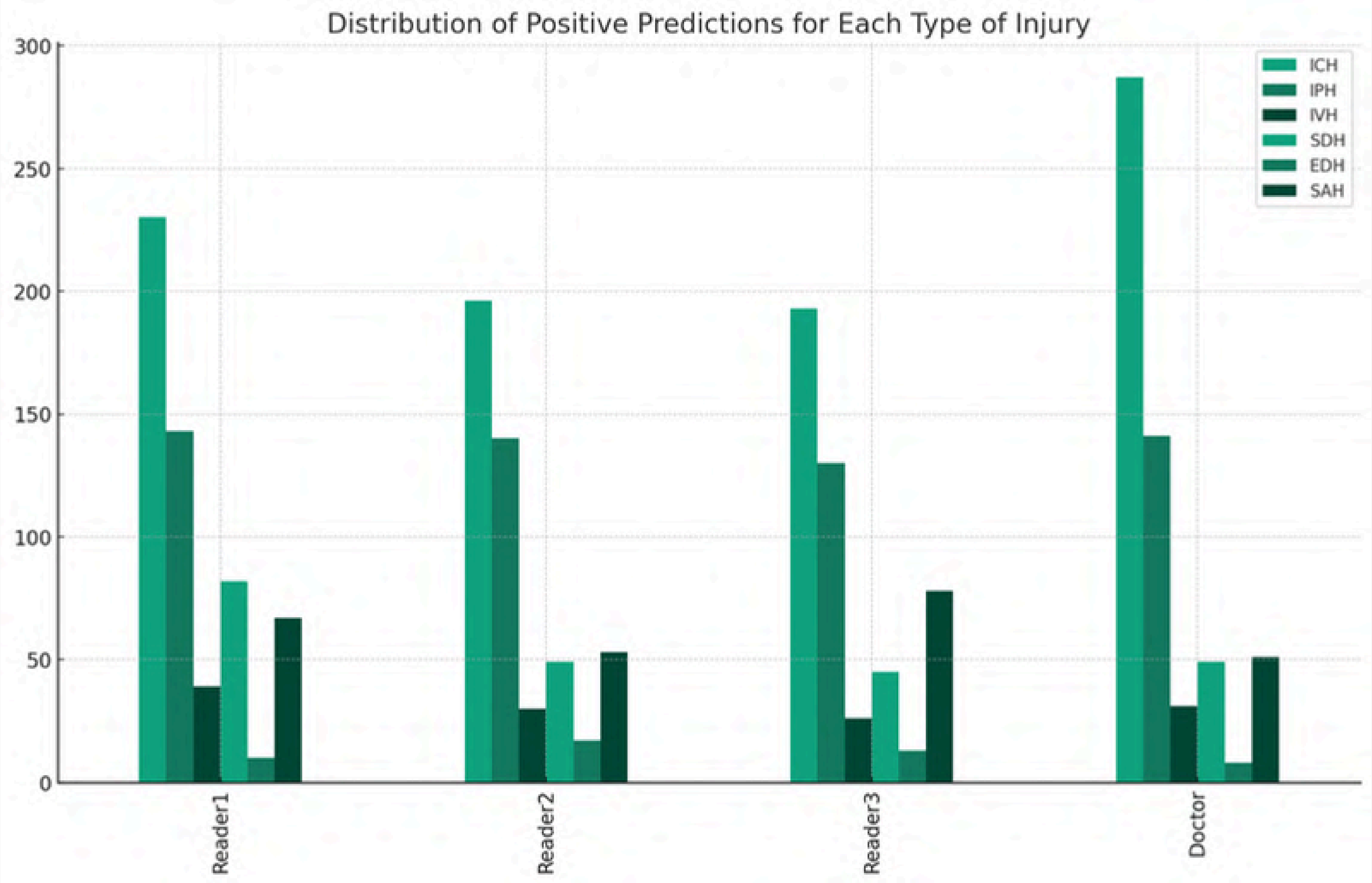
Increases the robustness of the model, and to prevent overfitting by making the data more varied.

Changes made in:

- contrast
- rotation
- flip
- saturation
- normalising the image

(all the pixel value come between 0 and 1)

Data Preprocessing - Reader Bias



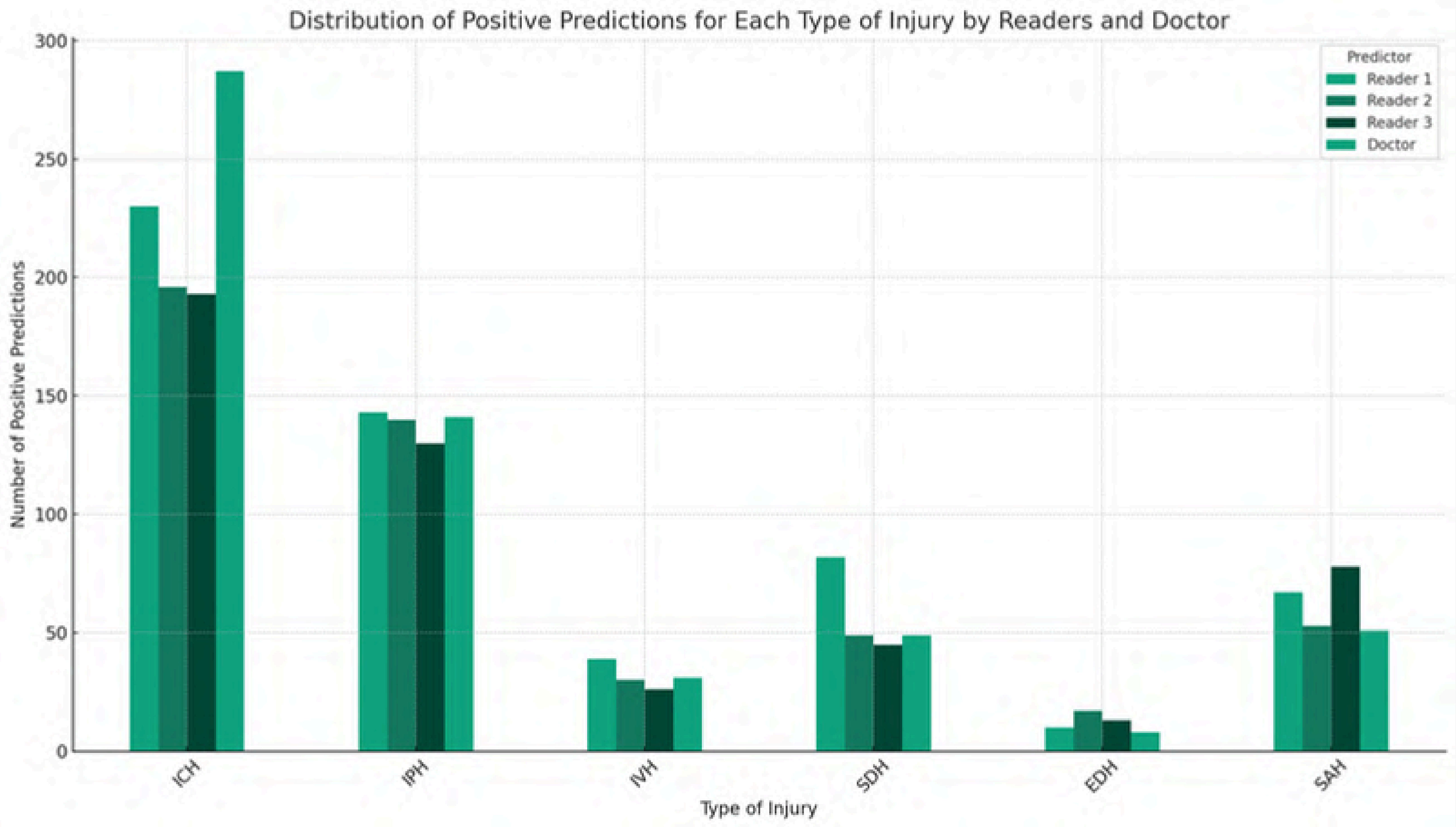
These patterns highlight the **individual tendencies** of each reader **towards** certain types of **injuries**.

Reader 1 appears to be more aggressive in diagnosing, especially for ICH and IPH, while Reader 2 has a specific focus on Midline Shift.
Reader 3 is the most conservative across the board but still pays close attention to ICH and Mass Effect.

The **doctor** is more likely to be accurate due to **increased context** from interactions with the patient.

Intracerebral Hemorrhage (ICH),
Intraparenchymal Hemorrhage (IPH)
Intraventricular Hemorrhage (IVH)
Subdural Hematoma (SDH)
Epidural Hematoma (EDH)
Subarachnoid Hemorrhage (SAH)

Data Preprocessing - Reader Bias



This is the same data, but plotted by grouping the types of injury, for further clarity on the biases.

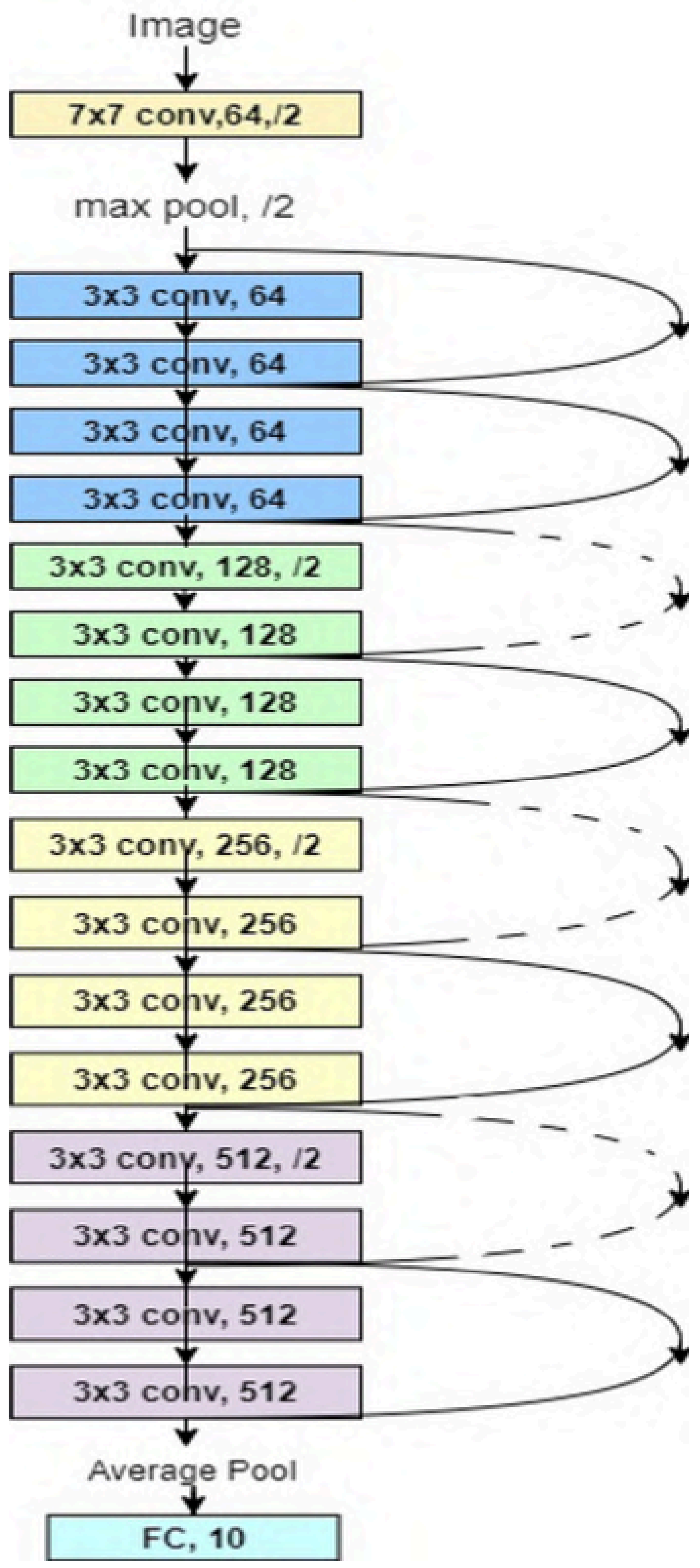
Data Preprocessing - Weighting by Experience

- Weighting the readers by experience, then sum their predictions grouped by injury
    - Reader 1 - 1 (8 years of experience), Reader 2 - 2 (12 years of experience), Reader 3 - 3 (20 years of experience)
- Normalise so the values earlier from 0 to 6 are now 0-1

	BleedLocation-Left	BleedLocation-Right	CalvarialFracture	ChronicBleed	EDH	Fracture	ICH	IPH	IVH
0	0.833333	1.000000	0.0	0.000000	0.000000	0.000000	1.000000	1.000000	0.000000
1	1.000000	0.000000	1.0	0.000000	0.833333	1.000000	1.000000	0.166667	0.000000
2	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
4	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
5	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
6	1.000000	0.166667	0.0	0.000000	0.000000	0.000000	1.000000	1.000000	1.000000
7	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
8	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
9	1.000000	1.000000	0.0	0.500000	0.000000	0.000000	0.666667	0.000000	0.000000
10	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
11	1.000000	1.000000	0.0	0.000000	0.000000	0.000000	1.000000	1.000000	1.000000
12	0.166667	1.000000	0.0	0.000000	0.000000	0.000000	1.000000	1.000000	1.000000
13	0.666667	1.000000	0.0	0.000000	0.000000	0.166667	1.000000	1.000000	0.000000
14	1.000000	0.000000	0.0	0.500000	0.000000	0.000000	1.000000	0.000000	0.000000
15	0.000000	1.000000	0.0	0.000000	0.000000	0.000000	1.000000	1.000000	0.000000
16	0.000000	0.500000	0.0	0.166667	0.000000	0.000000	0.500000	0.500000	0.000000
17	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
18	0.500000	1.000000	0.0	0.000000	0.000000	0.000000	1.000000	1.000000	0.166667
19	1.000000	0.166667	0.0	0.000000	0.000000	0.000000	1.000000	1.000000	0.000000

ResNet-18

- ResNets, or Residual Neural Networks, introduces **residual learning** to address the vanishing gradient problem. ResNets are widely used in medical image analysis because of their ability of **handling more complex features** within images.
- ResNet-18 has **18 layers** (17 convolutional and 1 fully connected). It is faster to train and less memory intensive while still benefiting from deep feature learning.
- Thanks to its residual blocks, ResNet-18 is more **parameter-efficient**, with fewer parameters (around **11.7 million**) compared to VGG-16 (138 million). This efficiency makes it suitable for environments with limited computational resources.
- ResNet-18 adapts well to various data sizes and is effective in avoiding overfitting compared to deeper models, especially when data augmentation and other regularization techniques are employed.
- **The modification of ResNet-18's first convolutional layer to handle single-channel (grayscale) input is crucial for medical imaging tasks, where grayscale images are common.**
- Additionally, ResNet-18's computational efficiency allows it to be trained faster and deployed with **less powerful hardware**, making it a good choice for CT scan analysis.



1. Bangar, S. (2022, July 5). Resnet Architecture Explained - Siddhesh Bangar - Medium. Medium. https://medium.com/@siddheshb008/resnet-architecture-explained-47309ea9283d
2. He, K., Zhang, X., Ren, S., & Sun, J. (2015, December 10). Deep residual learning for image recognition. arXiv.org. https://arxiv.org/abs/1512.03385
3. Architecture Diagram of RESNET-18 [21]. (n.d.). ResearchGate. https://www.researchgate.net/figure/Architecture-Diagram-of-ResNet-18-21_fig2_353655307

Resource Constraints

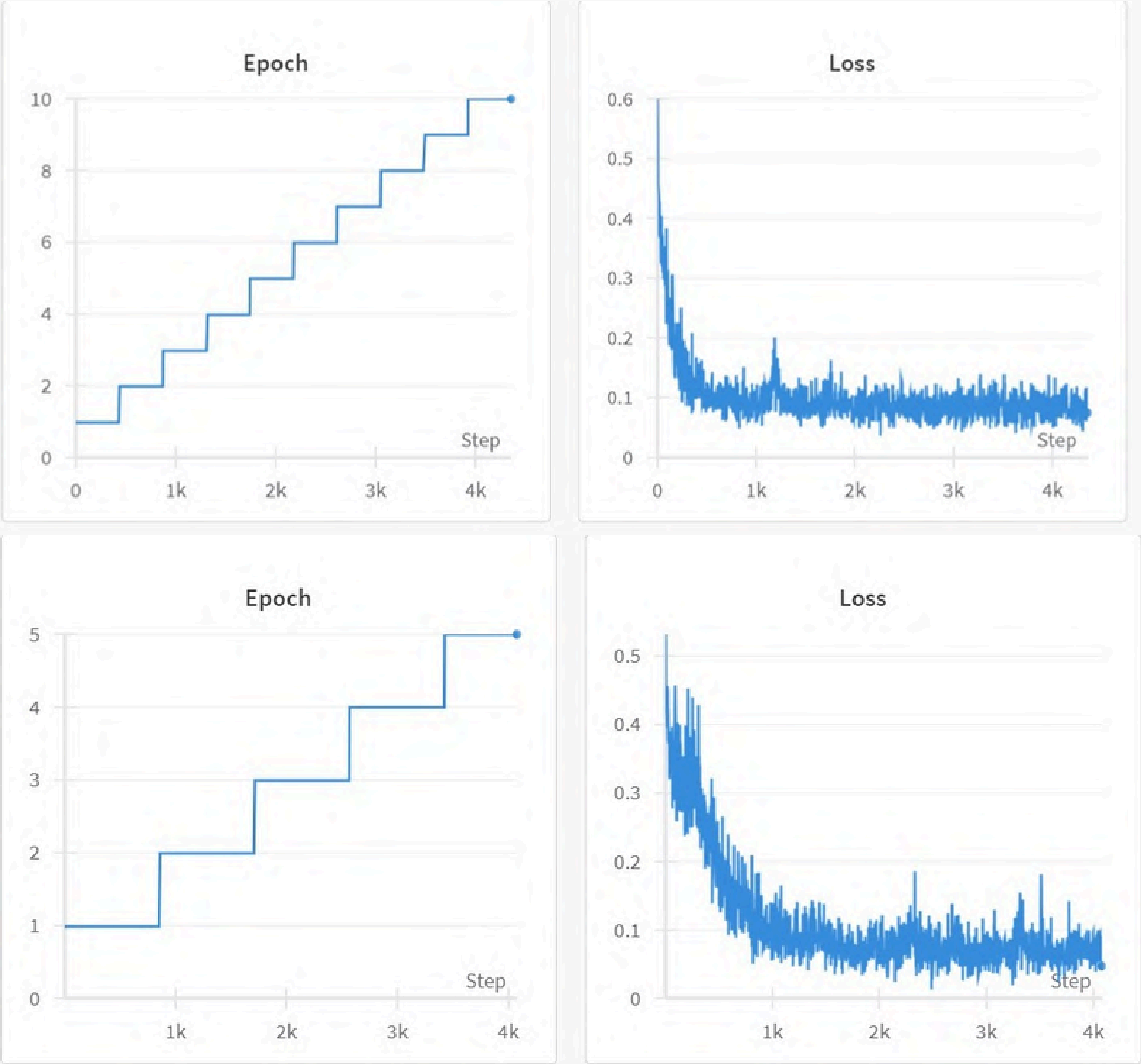
```
58m 9.0s
Epoch [1/10], currBatch: 0/4269, Loss: 0.7090
Epoch [1/10], currBatch: 1/4269, Loss: 0.6092
Epoch [1/10], currBatch: 2/4269, Loss: 0.5639
Epoch [1/10], currBatch: 3/4269, Loss: 0.4205
Epoch [1/10], currBatch: 4/4269, Loss: 0.3500
Epoch [1/10], currBatch: 5/4269, Loss: 0.4358
Epoch [1/10], currBatch: 6/4269, Loss: 0.3778
Epoch [1/10], currBatch: 7/4269, Loss: 0.4239
Epoch [1/10], currBatch: 8/4269, Loss: 0.4500
Epoch [1/10], currBatch: 9/4269, Loss: 0.4580
Epoch [1/10], currBatch: 10/4269, Loss: 0.3965
Epoch [1/10], currBatch: 11/4269, Loss: 0.3086
Epoch [1/10], currBatch: 12/4269, Loss: 0.4083
Epoch [1/10], currBatch: 13/4269, Loss: 0.3038
Epoch [1/10], currBatch: 14/4269, Loss: 0.4702
Epoch [1/10], currBatch: 15/4269, Loss: 0.5051
Epoch [1/10], currBatch: 16/4269, Loss: 0.3969
Epoch [1/10], currBatch: 17/4269, Loss: 0.3935
Epoch [1/10], currBatch: 18/4269, Loss: 0.3504
Epoch [1/10], currBatch: 19/4269, Loss: 0.3474
Epoch [1/10], currBatch: 20/4269, Loss: 0.3707
Epoch [1/10], currBatch: 21/4269, Loss: 0.3390
Epoch [1/10], currBatch: 22/4269, Loss: 0.3671
Epoch [1/10], currBatch: 23/4269, Loss: 0.3629
Epoch [1/10], currBatch: 24/4269, Loss: 0.3694
...
Epoch [1/10], currBatch: 331/4269, Loss: 0.3987
Epoch [1/10], currBatch: 332/4269, Loss: 0.4017
Epoch [1/10], currBatch: 333/4269, Loss: 0.4991
Epoch [1/10], currBatch: 334/4269, Loss: 0.3300
```

Training our 491 scans on a MacBook was a VERY long process, would take multiple days.

Approx. Run Time = 140 hours

Shifting Devices & Approach...

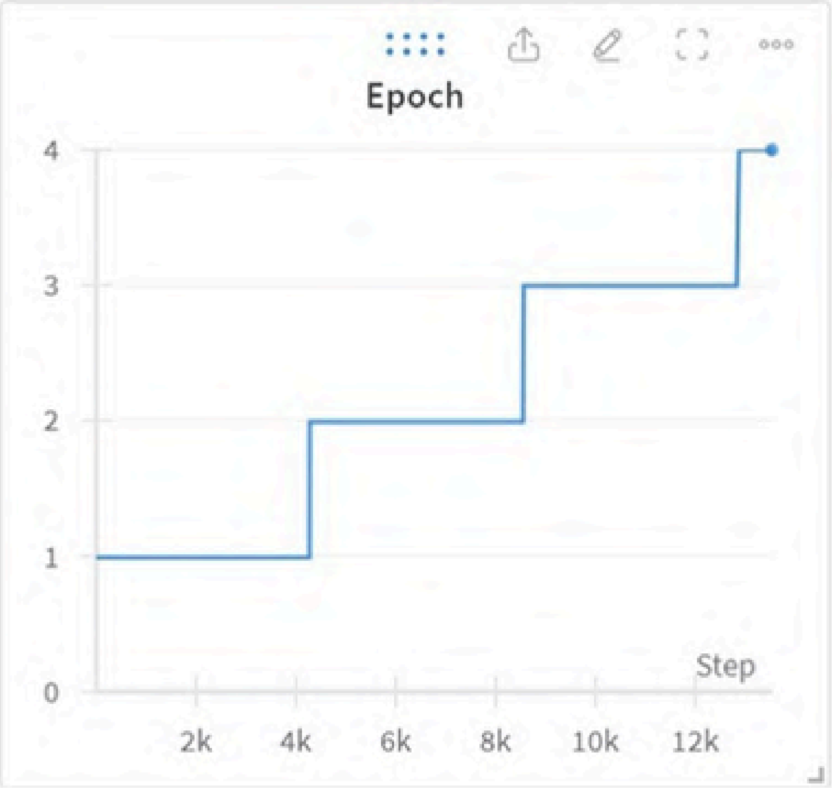
Accuracy on Test Set: 88.72%

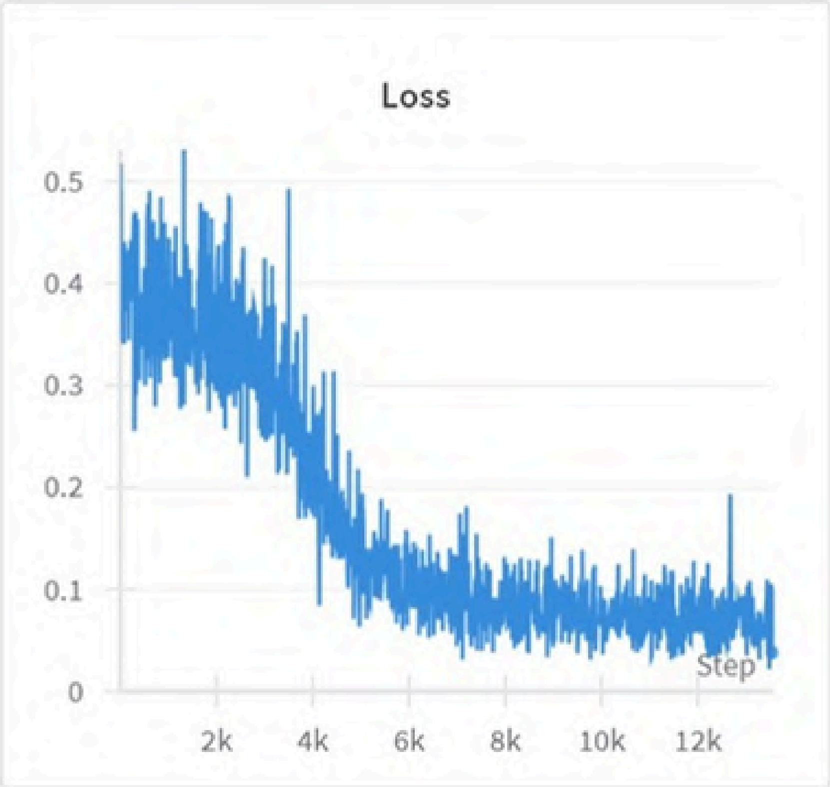


Running 50 & 100 Scans Due To Resource Constraints - They yielded a similar accuracy of 88%

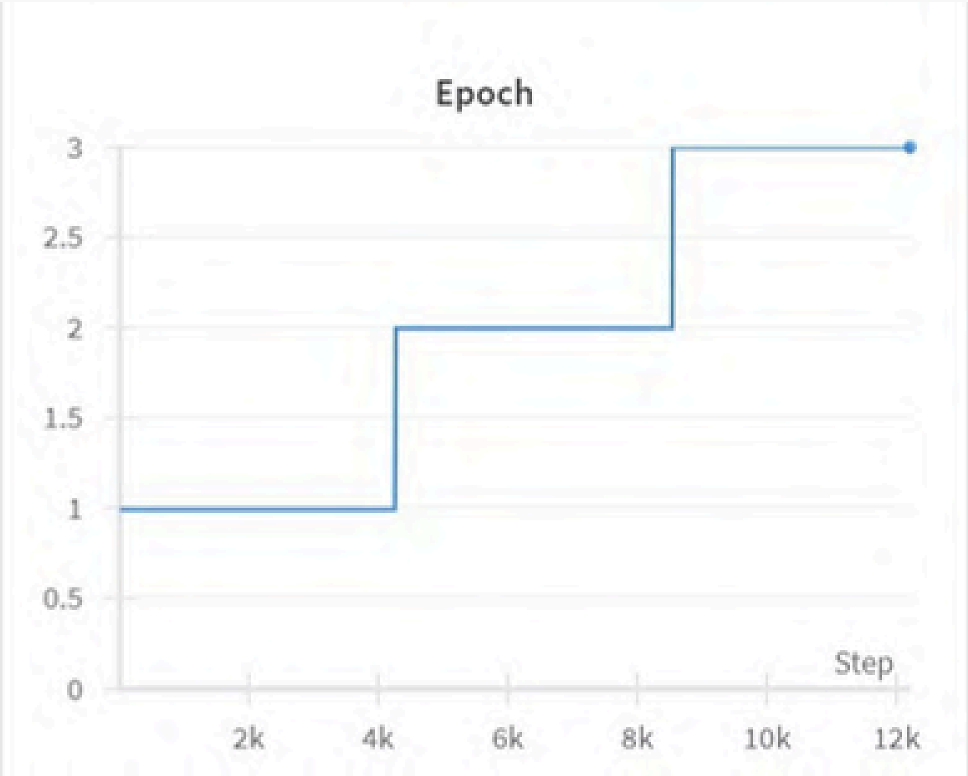
More Problems...

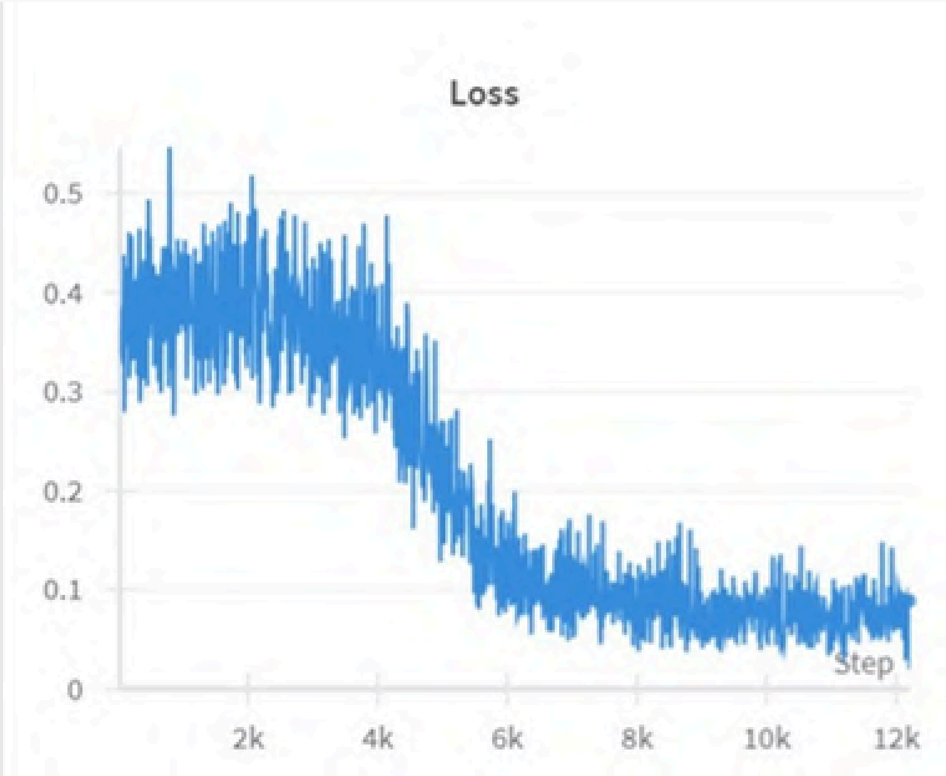
Then, the power went out!





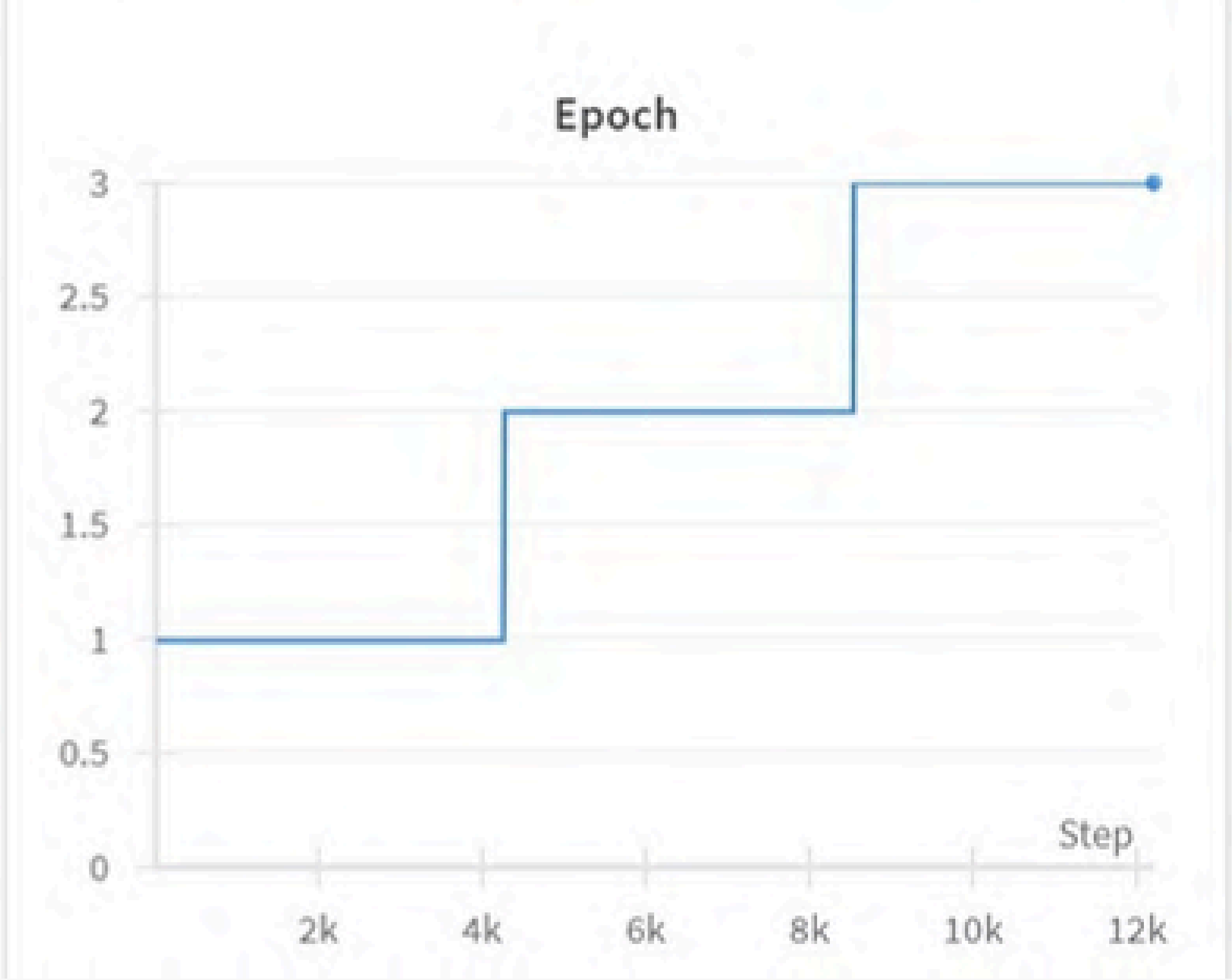
And again!

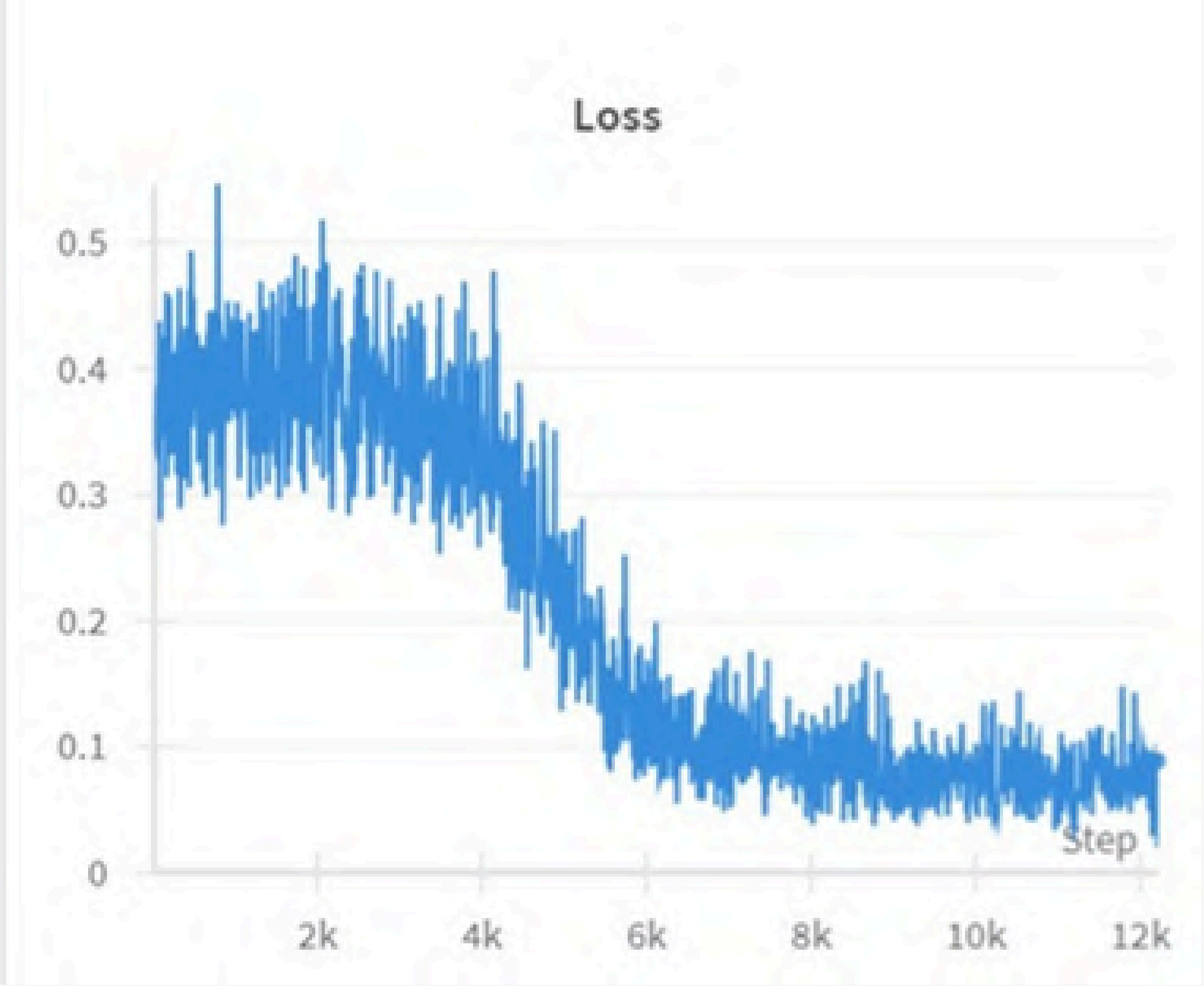




Finally, to prevent restarting after each power failure, we decided to save the weights after each epoch.

Accuracy on Test Set + F1 Scores





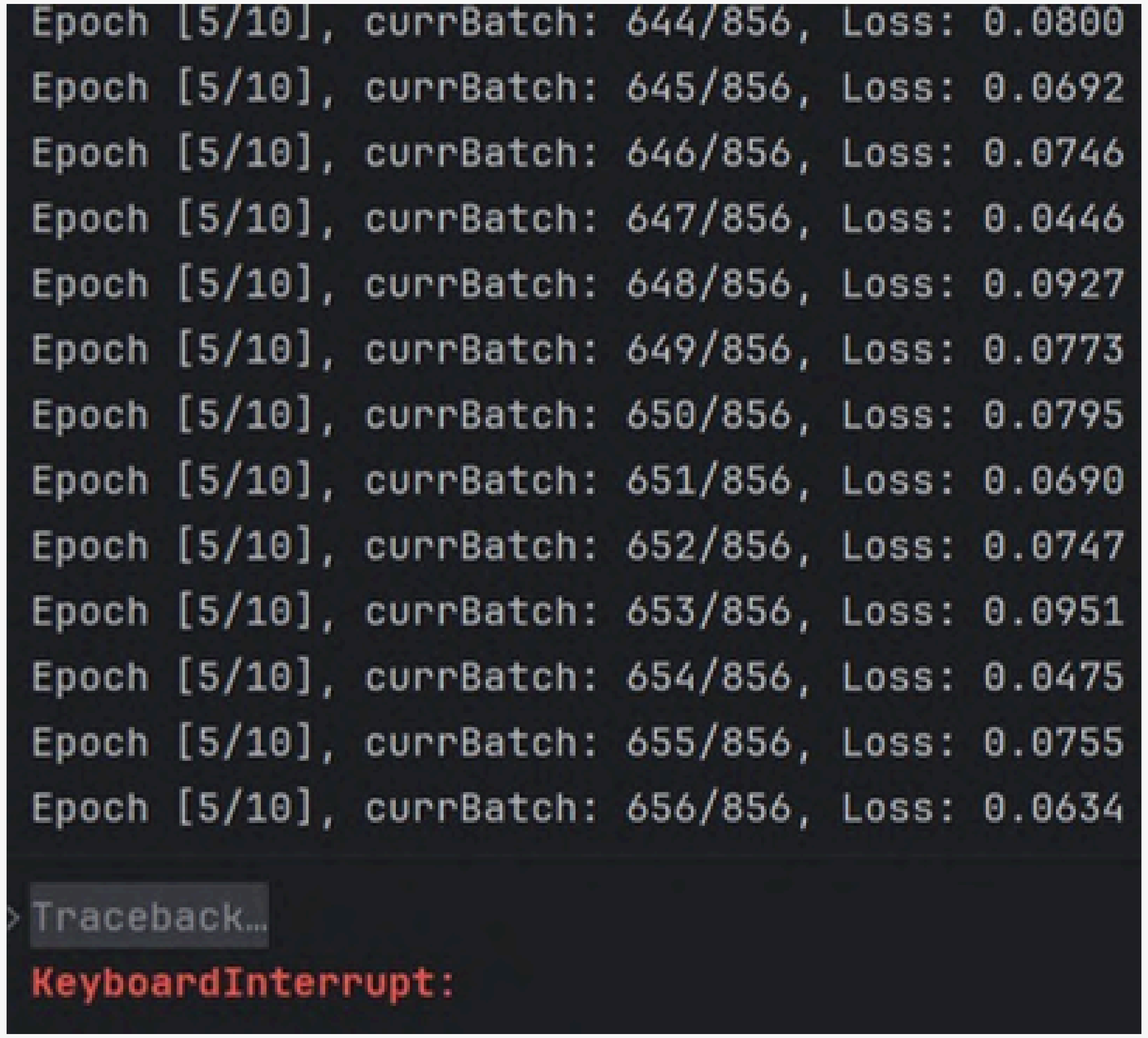

```
Accuracy on Test Set: 89.56%
```

```
F1 Scores per label: [0.96309963 0.96659427 0.93462171 0.8774026  0.93769678
0.91391492]
Average F1 Score: 0.9322216530381261
```

Our model converged! F1 scores:
EDH = 0.963
ICH = 0.966
IPH = 0.934
IVH = 0.877
SAH = 0.937
SDH = 0.913

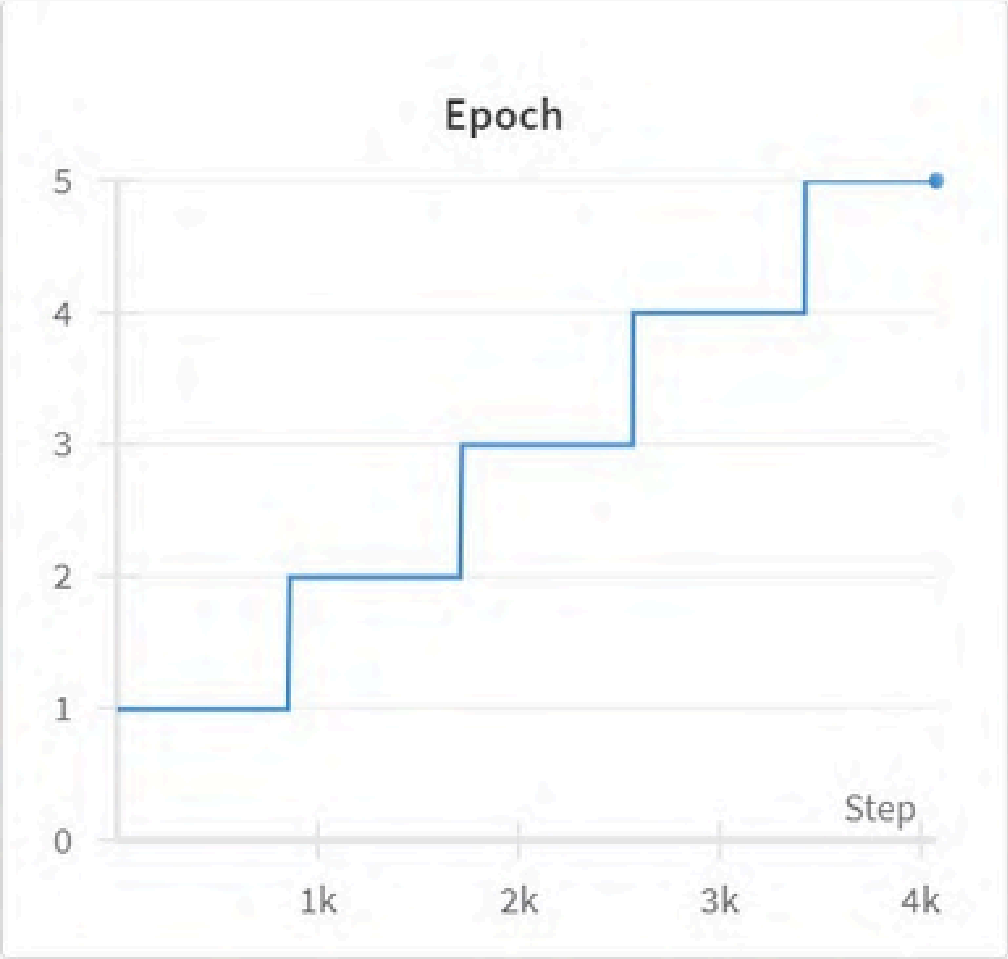
To Prevent Overfitting

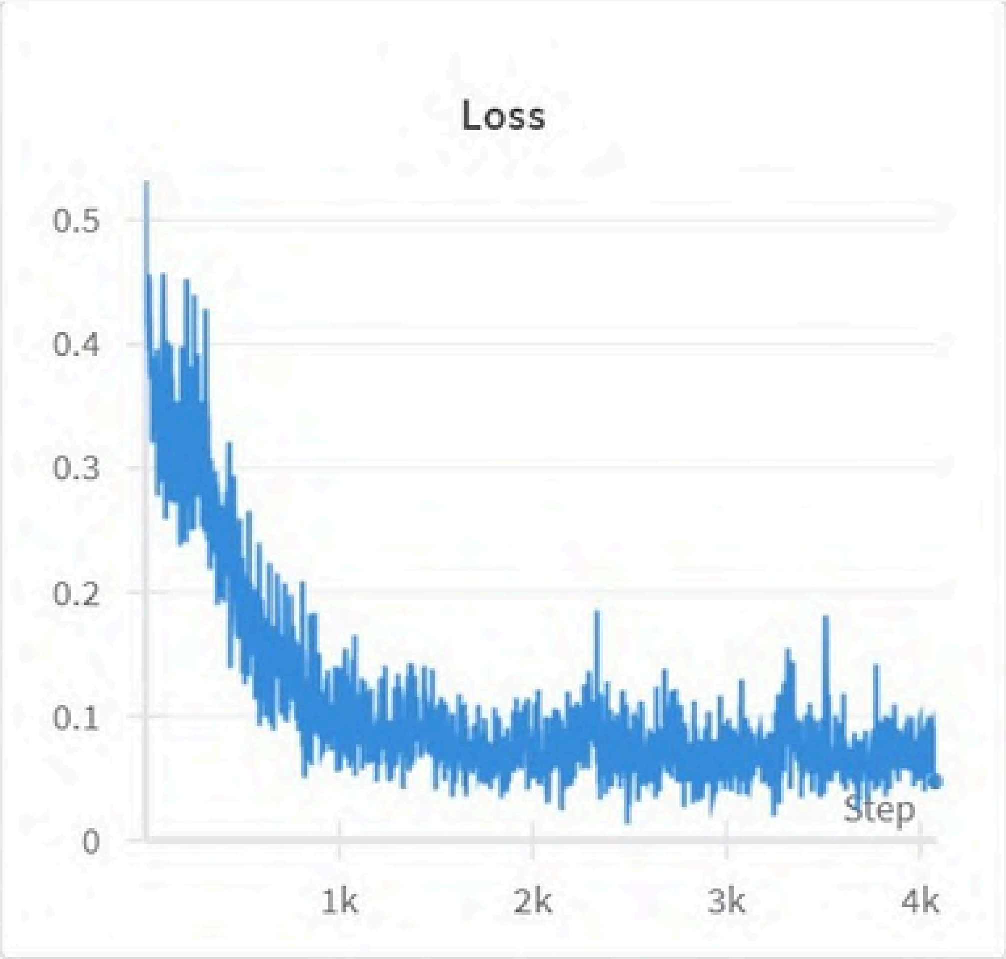
Early Stopping



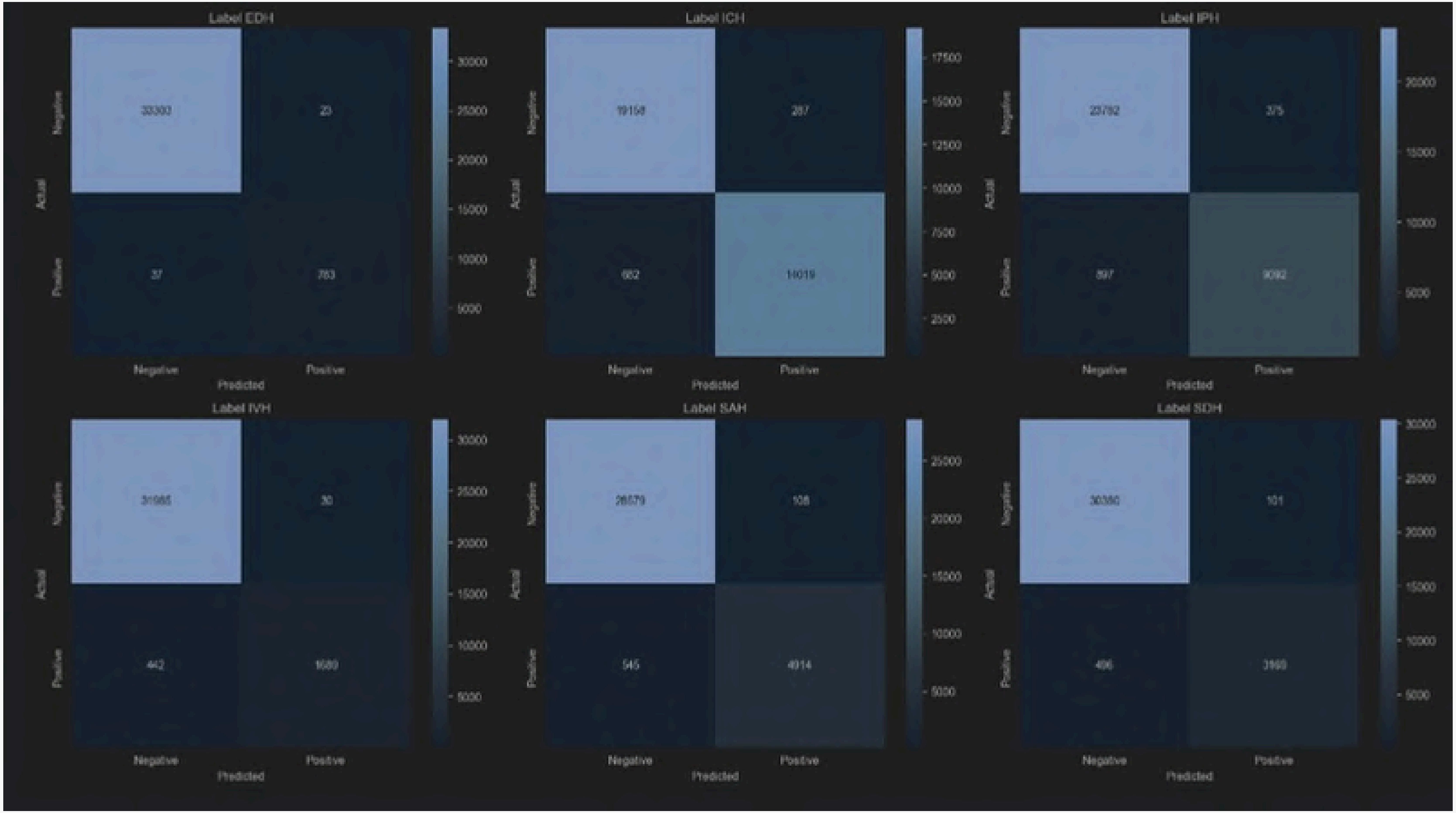
```
Epoch [5/10], currBatch: 644/856, Loss: 0.0800
Epoch [5/10], currBatch: 645/856, Loss: 0.0692
Epoch [5/10], currBatch: 646/856, Loss: 0.0746
Epoch [5/10], currBatch: 647/856, Loss: 0.0446
Epoch [5/10], currBatch: 648/856, Loss: 0.0927
Epoch [5/10], currBatch: 649/856, Loss: 0.0773
Epoch [5/10], currBatch: 650/856, Loss: 0.0795
Epoch [5/10], currBatch: 651/856, Loss: 0.0690
Epoch [5/10], currBatch: 652/856, Loss: 0.0747
Epoch [5/10], currBatch: 653/856, Loss: 0.0951
Epoch [5/10], currBatch: 654/856, Loss: 0.0475
Epoch [5/10], currBatch: 655/856, Loss: 0.0755
Epoch [5/10], currBatch: 656/856, Loss: 0.0634

> Traceback...
KeyboardInterrupt:
```

Logging the losses and epochs to ensure significant change between epochs and general downward trend of losses





Confusion Matrix: Multi-class Multi-label



Balanced Accuracy & Components

Specificity and Sensitivity

Sensitivity = TP/(TP+FN)	Specificity= TN/(TN+FP)	$\texttt{balanced-accuracy} = \frac{1}{2}\left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP}\right)$
EDH = 0.954	EDH = 0.999	EDH = 0.976
ICH = 0.953	ICH = 0.985	ICH = 0.969
IPH = 0.910	IPH = 0.984	IPH = 0.947
IVH = 0.792	IVH = 0.999	IVH = 0.895
SAH = 0.900	SAH = 0.996	SAH = 0.948
SDH = 0.864	SDH = 0.996	SDH = 0.930

Average Balanced Accuracy : 0.944 = **94.4%**

If the balanced accuracy > test accuracy, it suggests that the model is performing well in terms of correctly classifying our data, despite class within the dataset.

Sample Outputs

Scan Number: 14

True:

14	0	1	0	0	0	1

EDH, ICH, IPH, IVH, SAH, SDH

Predicted: `['ICH', 'SDH']`

1/352 slices of Scan 14

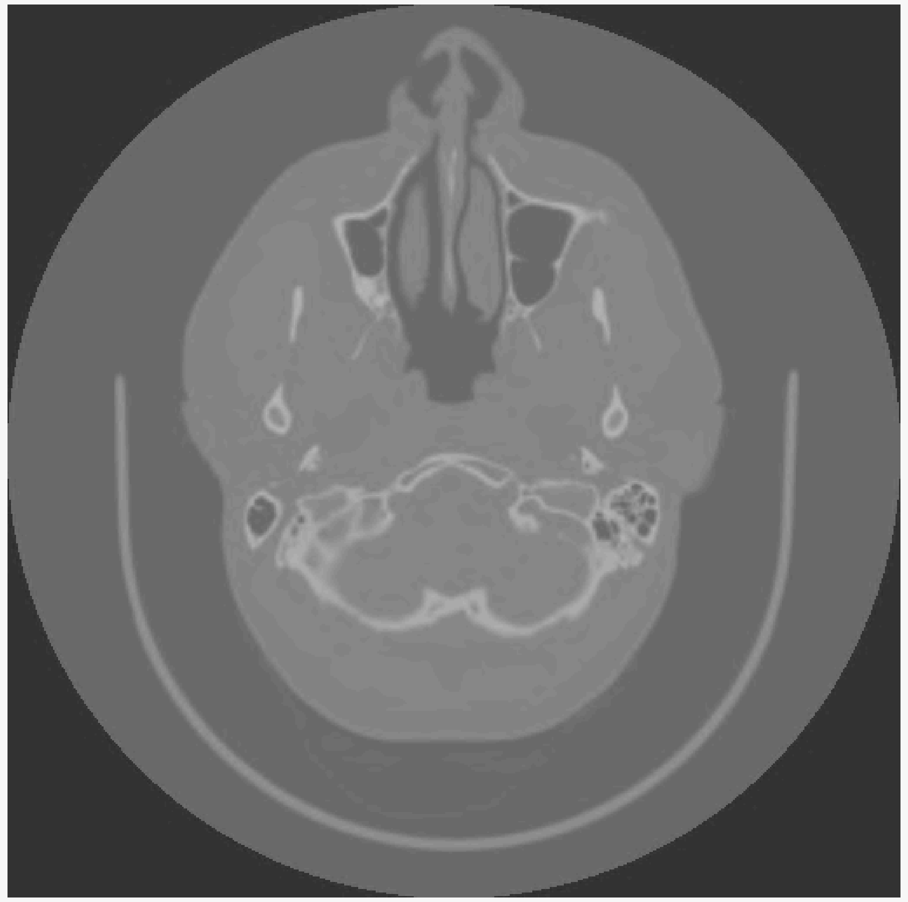
Scan Number: 172

True:

172	0	1	1	0	0	0

EDH, ICH, IPH, IVH, SAH, SDH

Predicted: `['ICH', 'IPH']`



1/1949 slices of Scan 172

Future Report...

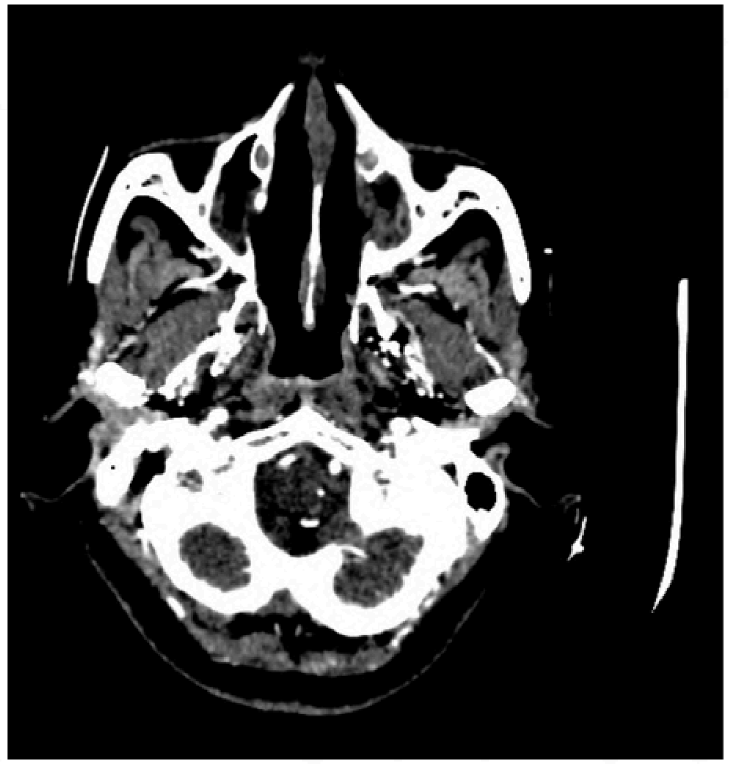
Hey CT!

Patient:	abc xyz	D.O.B :	01/05/1939
Radiologist:		Exam Date:	2017-11-23
Referring physician:	abc xyz		

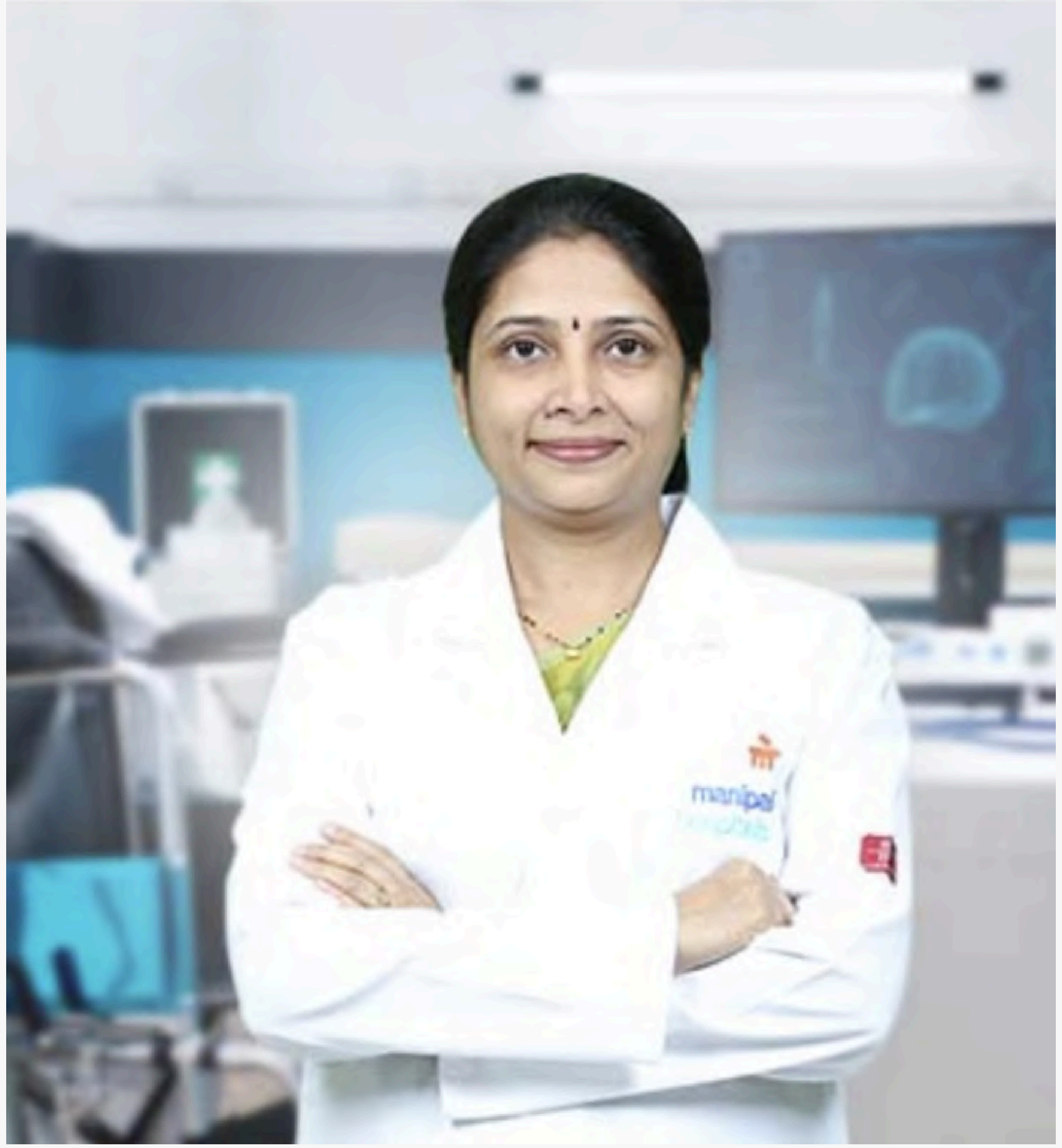
Head CT Report

Injuries Detected!

Intracranial Hemorrhage (ICH)
Intraparenchymal Hemorrhage (IPH)

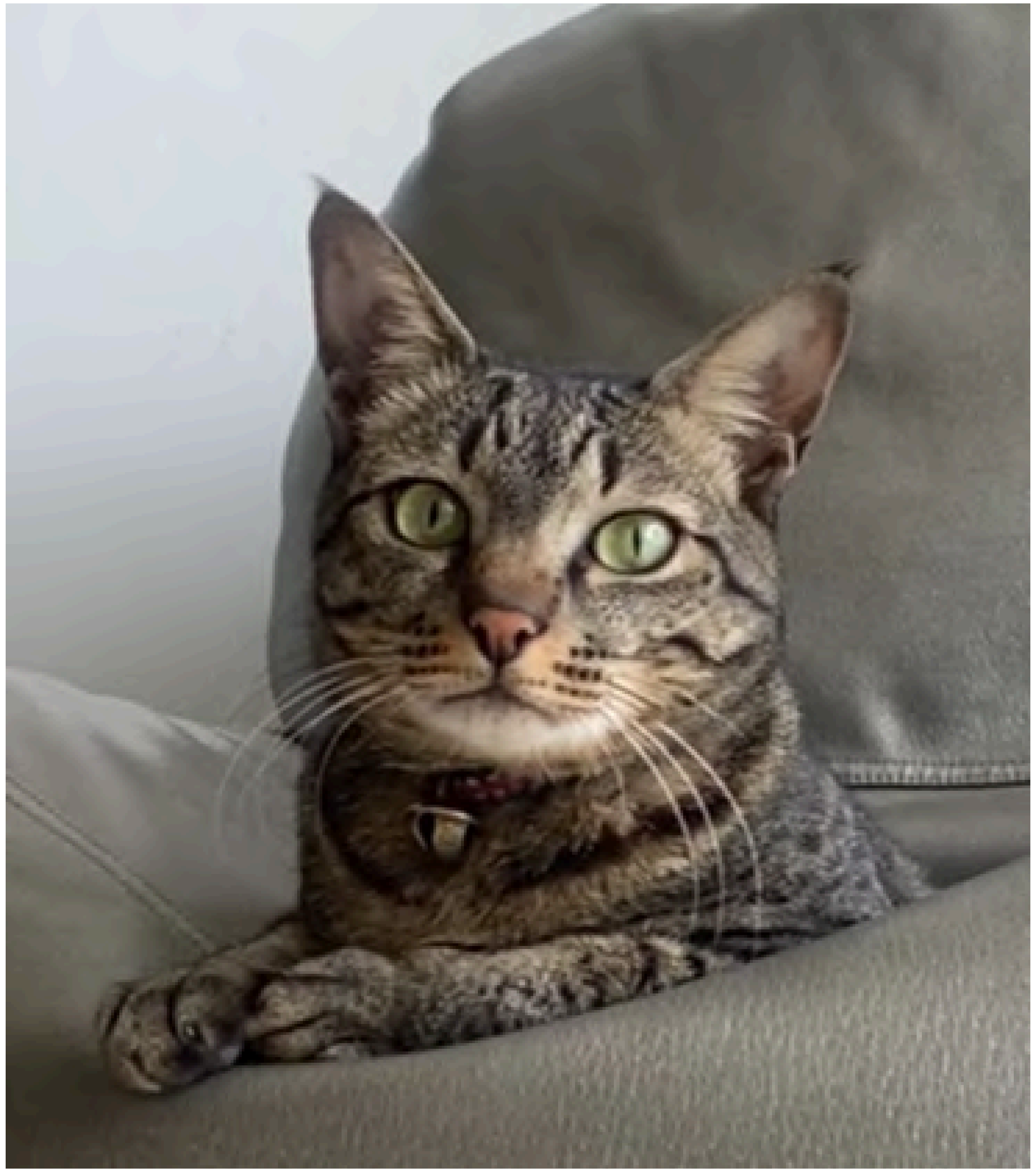


Advisory Board



Dr Anita Nagadi
Lead, Oncological Radiology and Head & Neck Imaging
MHRG, Bangalore

Dr Sumana Pallegar
Consultant Neurosurgeon
Amrita Hospitals



THANK YOU :)

PPT BY ACE & SAEBU