# HUMAN COMPUTER INTERACTION FINAL PROJECT

## **Cervical Spine Fracture Detection**

Ilayda Su Irday - 10702401050

#### • Introduction:

Over 1.5 million spine fractures occur annually in the United States alone resulting in over 17,730 spinal cord injuries annually. The most common site of spine fracture is the cervical spine. There has been a rise in the incidence of spinal fractures in the elderly and in this population, fractures can be more difficult to detect on imaging due to superimposed degenerative disease and osteoporosis. Imaging diagnosis of adult spine fractures is now almost exclusively performed with computed tomography (CT) instead of radiographs (x-rays). Quickly detecting and determining the location of any vertebral fractures is essential to prevent neurologic deterioration and paralysis after trauma.

In this project, I will try to develop CNN models that match the radiologists' performance in detecting and localizing fractures to the seven vertebrae that comprise the cervical spine.

#### Dataset:

**Dataset Description** 

The goal of this competition is to identify fractures in CT scans of the cervical spine (neck) at both the level of a single vertebrae and the entire patient. Quickly detecting and determining the location of any vertebral fractures is essential to prevent neurologic deterioration and paralysis after trauma.

This competition uses a hidden test. When your submitted notebook is scored the actual test data (including a full length sample submission) will be made available to your notebook.

| F        | il | ρ | ς |
|----------|----|---|---|
| <u>.</u> | •  |   | _ |

train.csv:

Metadata for the train test set.

StudyInstanceUID - The study ID. There is one unique study ID for each patient scan.

patient\_overall - One of the target columns. The patient level outcome, i.e. if any of the vertebrae are fractured.

C[1-7] - The other target columns. Whether the given vertebrae is fractured. See this diagram for the real location of each vertbrae in the spine.

#### test.csv:

Metadata for the test set prediction structure. Only the first few rows of the test set are available for download.

row id - The row ID. This will match the same column in the sample submission file.

StudyInstanceUID - The study ID.

prediction\_type - Which one of the eight target columns needs a prediction in this row.

[train/test] images/[StudyInstanceUID]/[slice number].dcm:

The image data, organized with one folder per scan. Expect to see roughly 1,500 scans in the hidden test set.

Each image is in the dicom file format. The DICOM image files are  $\leq 1$  mm slice thickness, axial orientation, and bone kernel. Note that some of the DICOM files are JPEG compressed. You may require additional resources to read the pixel array of these files, such as GDCM and pylibjpeg.

sample submission.csv:

A valid sample submission.

row id - The row ID. See the test.csv for what prediction needs to be filed in that row.

fractured - The target column.

train\_bounding\_boxes.csv Bounding boxes for a subset of the training set.

#### Segmentations:

Pixel level annotations for a subset of the training set. This data is provided in the nifti file format.

A portion of the imaging datasets have been segmented automatically using a 3D UNET model, and radiologists modified and approved the segmentations. The provided segmentation labels have values of 1 to 7 for C1 to C7 (seven cervical vertebrae) and 8 to 19 for T1 to T12 (twelve thoracic vertebrae are located in the center of your upper and middle back), and 0 for everything else. As we focused on the cervical spine, all scans have C1 to C7 labels but not all thoracic labels.

Be aware that the NIFTI files consist of segmentation in the sagittal plane, while the DICOM files are in the axial plane.

#### Datasets can be reached in:

https://www.kaggle.com/competitions/rsna-2022-cervical-spine-fracture-detection/data https://www.kaggle.com/datasets/ujjwalkar/cervical-spine-segmentation (additional data set)

#### Method:

Model Architecture: The model architecture in the provided code is based on transfer learning using the ResNet50 neural network. ResNet50 is a well-known architecture originally designed for image classification tasks. The architecture consists of convolutional layers followed by residual blocks, which help alleviate the vanishing gradient problem and enable training of very deep networks. In this code, the pre-trained ResNet50 model is used as the base model, followed by additional layers for classification.

The model architecture can be summarized as follows:

The base model: ResNet50 with pre-trained weights on ImageNet is loaded as the base model.

Flatten layer: The output from the base model is flattened to a 1D tensor.

Dense layer: A fully connected layer with 256 units and ReLU activation function is added.

Output layer: A dense layer with a single unit and sigmoid activation function is added for binary classification.

Model Functionality: The model takes cervical spine images as input and predicts the presence or absence of a specific clinical condition related to the spine. The ResNet50 base model extracts relevant features from the input images, which are then passed through the additional layers for classification. The sigmoid activation function in the output layer produces a probability score indicating the likelihood of the condition being present in each input image.

Choice of Model: The ResNet50 model is a popular choice for many computer vision tasks due to its effectiveness in capturing features from images and its ability to handle deep networks. Transfer learning with pre-trained models like ResNet50 saves significant computational resources and training time. By leveraging the knowledge learned from a large dataset (ImageNet), the model can generalize well to new tasks, even with limited labeled data. Therefore, choosing ResNet50 as a base model for cervical spine segmentation is a reasonable approach.

Model Training Strategy: The model training strategy in the provided code is supervised learning. The model is trained using labeled data, where each input image is associated with a corresponding binary label indicating the presence or absence of the clinical condition. The rationale behind supervised learning is that the model can learn to generalize patterns and make predictions based on the labeled examples. Since the task is binary classification, supervised learning is appropriate.

Loss Function and Optimizer: The loss function used is Binary Crossentropy, which is commonly used for binary classification problems. It measures the dissimilarity between the predicted probabilities and the true labels. The choice of Binary Crossentropy is appropriate because it encourages the model to optimize the predicted probabilities to match the true labels.

The optimizer used is Adam with a learning rate of 0.0005. Adam is an adaptive optimization algorithm that adjusts the learning rate during training to update the model parameters effectively. It combines the benefits of both AdaGrad and RMSProp algorithms. The learning rate of 0.0005 is a commonly used value, although the optimal learning rate can vary depending on the specific dataset and problem domain.

Overall, the choice of Binary Crossentropy loss function and Adam optimizer is reasonable for the binary classification task, as they are widely used and have proven to be effective in similar scenarios.

#### Experiments:

This code is a combination of data preprocessing, visualization, model training, and evaluation for cervical spine segmentation.

## Loading the dataset:

The code specifies the directory paths for the dataset (datadir), mask images (mask\_dir), and train images (image\_dir).

It creates lists of file paths for images and masks using os.listdir and os.path.join.

#### Visualization functions:

show\_spine(n): Loads and displays an image and its corresponding masks for the nth example in the dataset.

crop\_seg(n): Loads and displays an image with cropped regions of interest (based on masks) for the nth example.

save\_crop\_seg(n): Crops and saves the segmented regions of interest for the nth example.

## Data processing and organization:

The code creates a list called dataset to store the processed data.

The save\_crop\_seg function is used to generate and save cropped masks for all examples in the dataset.

The code creates a pandas DataFrame called data to store information about the processed data.

## Augmentation:

The code does not include explicit augmentation techniques. Augmentation is a common approach to increase the size and diversity of the training data by applying various transformations such as rotation, scaling, flipping, or adding noise to the images. Augmentation can help improve the model's generalization ability and robustness. However, in this case, augmentation techniques are not implemented in the code.

## Dataset usage and data splitting:

The code splits the dataset into a training set and a test set using the train\_test\_split function from the sklearn library. The default test size is set to 0.33, meaning 33% of the data will be used for testing. The training set is used for training the model, and the test set is used for evaluating the model's performance.

## Weight initialization:

In this case, the ResNet50 model is likely initialized with pre-trained weights on ImageNet. This is a common practice in transfer learning, where the weights learned from a large dataset (ImageNet) are used as a starting point for training on a different task (cervical spine segmentation).

## Model building and training:

build\_model(data, tl\_name='resnet', batch\_size=16): This function builds and trains a ResNet-based model for each category of cervical spine segmentation (C1, C2, C3, etc.).

The function splits the data into train and test sets and loads the images and labels accordingly.

It then creates a ResNet50 model and compiles it.

Training is performed using the loaded images and labels, with early stopping to prevent overfitting.

The function evaluates the model on the test set, displays training and validation loss/accuracy, and prints a classification report.

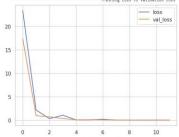
Finally, it plots the training/validation loss and accuracy and visualizes the confusion matrix.

#### Model results:

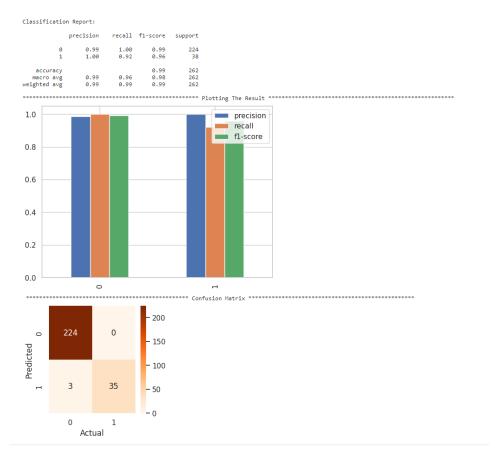
Cervical Spine Fracture Detection in C1 vertbrae Results

| **************                              | SEES Category : C1 Output | file name: spine C1 h5   | **********************              | ****  |  |
|---|---------------------------|--|-------------------------------------|---|--|
| Model: "sequential 1"                       | category : c1, output     | rite mame: spine_ci.ms   |                                     |   |  |
|   |                           |  |                                     |   |  |
| Layer (type)                                | Output Shape              | Param #  |                                     |   |  |
|   | (None, 16, 16, 2048)      | 23581440   |                                     |   |  |
| resnet50 (Functional)                       | (None, 16, 16, 2048)      | 23581440   |                                     |   |  |
| flatten 1 (Flatten)                         | (None, 524288)            | 0  |                                     |   |  |
|   |                           |  |                                     |   |  |
| dense_2 (Dense)                             | (None, 256)               | 134217984  |                                     |   |  |
| dense 3 (Dense)                             | (None, 1)                 | 257  |                                     |   |  |
| dense_3 (bense)                             | (None, 1)                 | 231  |                                     |   |  |
|   |                           |  |                                     |   |  |
| Total params: 157,799,681                   |                           |  |                                     |   |  |
| rainable params: 157,746                    |                           |  |                                     |   |  |
| Non-trainable params: 53,                   | 120                       |  |                                     |   |  |
|   |                           |  |                                     |   |  |
| None  |                           | ining Data *********   |                                     | ****  |  |
| Epoch 1/50                                  | 110                       | Ining Date   |                                     |   |  |
|   | 69s 1s/ste                | p - loss: 23.3485 - accur  | acy: 0.8374 - val loss: 17.3556 - v | val accuracy: 0.8473  |  |
| Epoch 2/50                                  | The second of             |  | A common Barrana                    |   |  |
|   | 31s 911ms/                | step - loss: 2.1616 - acc  | uracy: 0.8847 - val_loss: 0.9648 -  | val_accuracy: 0.8931  |  |
| Epoch 3/50                                  |                           |  |                                     |   |  |
| 54/34 [==================================== |                           | step - loss: 0.3310 - acc  | uracy: 0.9509 - val_loss: 0.7150 -  | Val_accuracy: 0.9237  |  |
|   | 1 - 30s 878ms/            | sten - loss: 1.8621 - acc  | uracy: 0.9452 - val loss: 0.3703 -  | val accuracy: 8.9198  |  |
| Epoch 5/50                                  |                           | scep - 1055. 1.0021 - ac   | Hacy. 0.5452 - Val_1085. 0.5705 -   | Val_accoracy. 0.5150  |  |
| 4/34 [                                      | ********* - 325 953ms/    | step - loss: 0.0646 - acc  | uracy: 0.9849 - val loss: 0.0723 -  | val accuracy: 0.9656  |  |
| poch 6/50                                   |                           |  |                                     |   |  |
|   | 31s 902ms/                | step - loss: 0.0501 - ac   | uracy: 0.9887 - val_loss: 0.0925 -  | val_accuracy: 0.9733  |  |
| poch 7/50                                   | 1 20 000                  |  |                                     |   |  |
| poch 8/50                                   |                           | step - 1055: 0.1863 - acc  | uracy: 0.9735 - val_loss: 0.0684 -  | Val_accuracy: 0.9618  |  |
|   | 1 - 31- 904ms/            | sten - loss: 0 0204 - acc  | uracy: 0.9981 - val loss: 0.0141 -  | val accuracy: 0 9924  |  |
| poch 9/50                                   |                           |  |                                     | ***_**********************************                      |  |
|   | 1 - 32s 952ms/            | step - loss: 7.2477e-84 -  | accuracy: 1.0000 - val loss: 0.02   | 76 - val accuracy: 0.9885                                   |  |
| Epoch 10/50                                 | * 1000000000              |  |                                     | 20 100 <del>=</del> 100 200 200 200 200 200 200 200 200 200 |  |
|   | 30s 879ms/                | step - loss: 4.8174e-06 -  | accuracy: 1.0000 - val_loss: 0.030  | 03 - val_accuracy: 0.9885                                   |  |
| Ipoch 11/50                                 |                           | na Cilia de la composição |                                     |   |  |
|   |                           | step - 10ss: 0.0052 - ac   | uracy: 0.9962 - val_loss: 0.0285 -  | val_accuracy: 0.9962  |  |
| Epoch 12/50                                 |                           |  |                                     |   |  |

| 1259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 | 2259 |



1.00 accuracy vs Validation Accuracy vs Valid



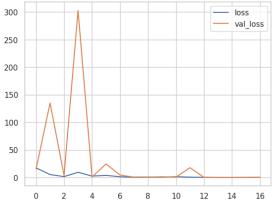
Cervical Spine Fracture Detection in C2 vertbrae Results

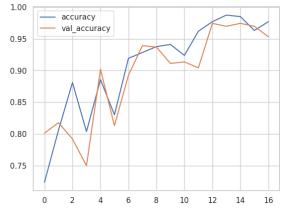
Model: "sequential 2"

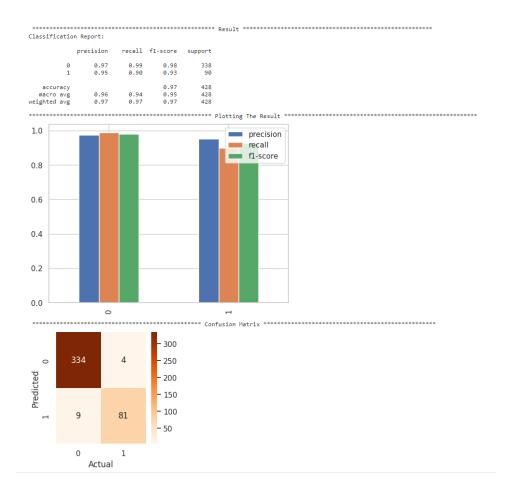
| Layer (type)          | Output | Shape         | Param #   |
|-----------------------|--------|---------------|-----------|
| resnet50 (Functional) | (None, | 16, 16, 2048) | 23581440  |
| flatten_2 (Flatten)   | (None, | 524288)       | 0         |
| dense_4 (Dense)       | (None, | 256)          | 134217984 |
| dense_5 (Dense)       | (None, | 1)            | 257       |

Total params: 157,799,681 Trainable params: 157,746,561 Non-trainable params: 53,120

None Training Data... Epoch 55/55 20/55 [==== Epoch 10/50 55/55 [==== Epoch 11/50 Epoch 55/55 [===== 12/50 Epoch 55/55 ================================ ] - 49s 899ms/step - loss: 0.5615 - accuracy: 0.9619 - val\_loss: 17.6770 - val\_accuracy: 0.9042 13/50 Epoch 55/55 14/50 



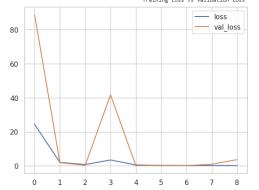




### Cervical Spine Fracture Detection in C3 vertbrae Results

```
Model: "sequential_3"
Layer (type)
              Output Shape
                            Param #
resnet50 (Functional)
              (None, 16, 16, 2048)
                           23581440
flatten_3 (Flatten)
              (None, 524288)
dense 6 (Dense)
              (None, 256)
                           134217984
dense_7 (Dense)
              (None, 1)
                           257
Total params: 157,799,681
Trainable params: 157,746,561
Non-trainable params: 53,120
None Training Data...
Epoch 1/50
67/67 [====
Epoch 2/50
67/67 [====
Epoch 3/50
67/67 [====
Epoch 4/50
67/67 [====
Epoch 6/50
67/67 [====
      :=========] - 74s 582ms/step - loss: 24.3452 - accuracy: 0.7205 - val_loss: 88.8398 - val_accuracy: 0.7159
         ========] - 33s 501ms/step - loss: 3.4064 - accuracy: 0.8612 - val_loss: 41.7688 - val_accuracy: 0.7727
         67/67 [====
Epoch 7/50
```







Classification Report:

 precision
 recall
 f1-score
 support

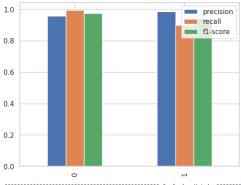
 0
 0.96
 0.99
 0.98
 184

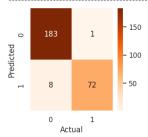
 1
 0.99
 0.90
 0.94
 88

 accuracy
 0.97
 0.95
 0.95
 264

 weighted avg
 0.97
 0.95
 0.96
 264

 owighted avg
 0.97
 0.97
 0.97
 0.97





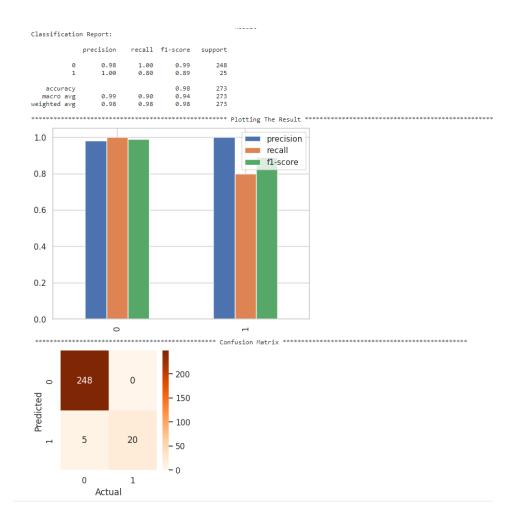
## Cervical Spine Fracture Detection in C4 vertbrae Results

Output Shape Param # resnet50 (Functional) (None, 16, 16, 2048) 23581440 flatten\_4 (Flatten) (None, 524288) dense\_8 (Dense) (None, 256) 134217984 dense\_9 (Dense) (None, 1) 257 Total params: 157,799,681 Trainable params: 157,746,561 Non-trainable params: 53,120 None Training Data... | Training Da 35 - loss --- val\_loss 30 25 20 15 10 5 0 4 5 0 3 1 1.00 accuracy val\_accuracy 0.98 0.96 0.94 0.92 0.90 0.88

0

1

2



Cervical Spine Fracture Detection in C5 vertbrae Results

Model: "sequential 5"

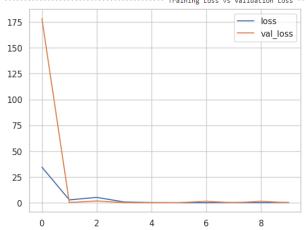
| Layer (type)          | Output Shape        | Param #     |
|-----------------------|---------------------|-------------|
| resnet50 (Functional) | (None, 16, 16, 2048 | 3) 23581440 |
| flatten_5 (Flatten)   | (None, 524288)      | 0           |
| dense_10 (Dense)      | (None, 256)         | 134217984   |

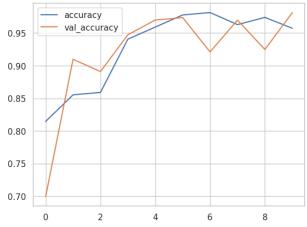
dense\_10 (Dense) dense\_11 (Dense) (None, 1)

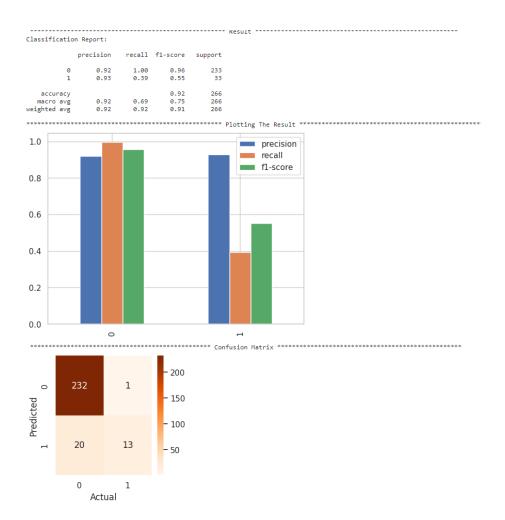
Total params: 157,799,681 Trainable params: 157,746,561 Non-trainable params: 53,120

None Training Data...

cpocn 6/30 68/68 [==========] - 34s 507ms/step - loss: 0.1111 - accuracy: 0.9777 - val\_loss: 0.1817 - val\_accuracy: 0.9737 Epoch 7/50 - 34s 495ms/step - loss: 0.1818 - accuracy: 0.9740 - val\_loss: 1.5373 - val\_accuracy: 0.9248
Epoch 10/50 ### Comparison of Comparison o





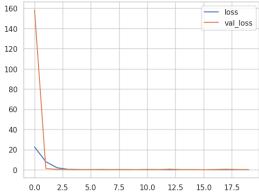


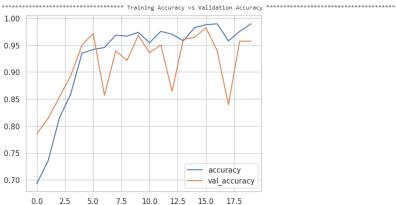
Cervical Spine Fracture Detection in C6 vertbrae Results

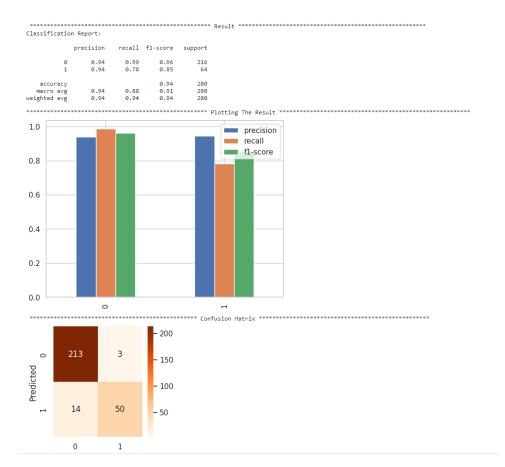
Model: "sequential 6"

| Layer (type)          | Output Shape         | Param #   |
|-----------------------|----------------------|-----------|
| resnet50 (Functional) | (None, 16, 16, 2048) | 23581440  |
| flatten_6 (Flatten)   | (None, 524288)       | 0         |
| dense_12 (Dense)      | (None, 256)          | 134217984 |
| dense_13 (Dense)      | (None, 1)            | 257       |
|                       |                      |           |

Trainable params: 157,799,681
Trainable params: 157,746,561
Non-trainable params: 53,120

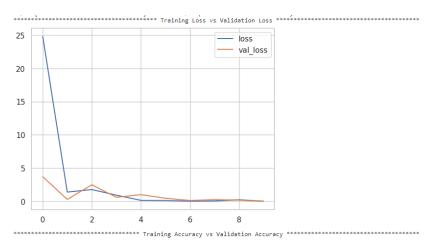


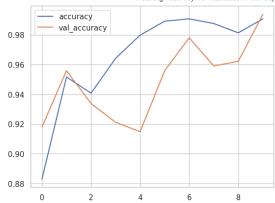


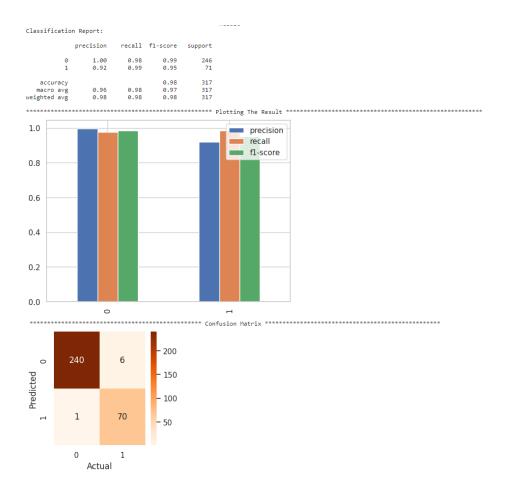


#### Cervical Spine Fracture Detection in C7 vertbrae Results

```
Model: "sequential_7"
Layer (type)
      Output Shape
resnet50 (Functional)
      (None, 16, 16, 2048)
           23581440
flatten_7 (Flatten)
      (None, 524288)
dense_14 (Dense)
      (None, 256)
            134217984
dense_15 (Dense)
Total params: 157,799,681
Trainable params: 157,746,561
Non-trainable params: 53,120
None Training Data...
161/161 [===
```







## • Conclusion and Discussion:

Model is generalize succesfuly.

Additional Steps: Data Augmentation can be applicable for better generalized models. Also, instead of binary class classification, multiclass classification can be used.