Problem Approach

In this problem, we are given CT scan data of Covid & non-Covid patients. These images are stored separately respective folders. Following steps are done to solve this problem-

1. All the images resized to a same size (224x224) and saved in different folder.
2. Now, we have around 1200 images for each covid & non-covid. Then split the dataset in three parts – train, validation, test in ratio (0.7,0.2,0.1).
3. Then on train data, data augmentation done with variables- rescale=1./255, rotation_range=45, width_shift_range=0.2, height_shift_range=0.2, shear_range=0.2 and zoom_range=0.2. And train, val dataset generator created.
4. Added modelcheckpoints with monitoring parameter 'accuracy' and added early stopping with patience value 10.
5. Resnet50 model created using sequential, and layers added to it. Dropout value of 0.5 defined.
6. Model compiled for binary_crossentropy loss, .0001 LR, and evaluation metrics = Accuracy, precision, recall, and AUC.
7. Then model fit on train data for 100 epochs and step = total data/batch_size.
8. Model's history of metrics & loss is saved and plotted for comparison.
9. Final model saved for future reference.