Artefact:

Computer-assisted breast cancer detection is a useful and widely used technique that aids pathologists in making clinical diagnoses and running their businesses more efficiently. In existing research, breast cancer identification is based mostly on a single imaging characteristic. In this work, we propose the use of Resnet 50, which is a state-of-the-art technique, for these types of detecting issues. Our model is trained on 1440 histopathological images which are already verified by pathologists along with the count of the cells present in the image. To begin, we extract a complementary characteristic known as cell group, which is used to compute the density of tumour cells, one of the cell group's categories. For training and evaluating our Resnet50 model, we try to divide all of the pictures in the NuCLS dataset into two groups. The data pre-processing and augmentation are carried out in accordance with standard processing protocols, which we will go over in detail below. We use various combinations of learning rates, trainable layers, optimizers, and other elements to create and construct our models' multiple times. Finally, we compare the results of all the models and iterations. Our analyses reveal that the Resnet 50 model, which has all layers trainable using the learning rate obtained by performing hyperparameter tuning, outperforms baselines on normalised histopathological validation picture data. Our method encourages the use of hyperparameter adjustment to combat overfitting.

Pre-requisites:

- Windows (Tested on Windows 11 latest).
- NVIDIA GPU (Tested on Nvidia GeForce RTX 2060 x 14).
- Distribution: Anaconda Navigator (Tested on version 4.11.0).
- Python (3.6.2), numpy (1.9.3), opencv-python (4.2.0), json (2.0.9), scipy (1.4.1), tensorflow (2.3.0), tensorflow.keras (2.4.0), matplotlib (3.3.2), skimage (0.17.2), scikit-learn (0.23.2).

Installation Guide for Anaconda (using Windows):

- Installation Guide

Installation guide for libraries:

```
pip install <name_of_the_library>==<version>
(or)
python -m pip install <name_of_the_library>==<version>
(or)
conda install <name_of_the_library>==<version>
```

Commands to run python codes:

```
python <name_of_the_file>.py
(or)
python3 <name_of_the_file>.py
```

Directory Set-up:

```
→ Working_Directory/

        ├──codes/

              ├──Data_Extractor.py
              ├──Density_Finder.py
              ├──optimization_resnet50.py
              ├──Res50.py
              ├──Res50_Conv4_5.py
              ├──Res50_Conv5.py
              ├──Res101.py
              ├──Res101_Conv4_5.py
              ├──Res101_Conv5.py
              ├──Res152.py
              ├──Res152_Conv4_5.py
              └──Res152_Conv5.py
        ├──Ep_100
              ├──LR_0_005
                    ├──Code
                        └──Res50
                             ├──Density.npy
                             ├──Res50.py
                             └──test.json
                    ├──Graphs
                        └──Res50
                             ├──cosine
                             ├──loss
                             ├──mae
                             ├──mse
                             ├──msle
                             └──rmse
                    ├──Log_Files
                        └──Res50
                             └──log0.xlsx
                    └──Models
                        └──Res50
                             └──best_model<epoch>-<val_loss>.h5
                             └──best_model<epoch>-<val_loss>.h5
                             └──best_model<epoch>-<val_loss>.h5
                             └──  …
              ├──LR_0_000006
                    ├──Code
                        └──Res50
                             ├──Density.npy
                             ├──Res50.py
                             └──test.json
                    ├──Graphs
                        └──Res50
                             ├──cosine
                             ├──loss
                             ├──mae
                             ├──mse
                             ├──msle
                             └──rmse
                    ├──Log_Files
                        └──Res50
                             └──log0.xlsx
                    └──Models
```

```
                └──Res50
                     └──best_model<epoch>-<val_loss>.h5
                     └──best_model<epoch>-<val_loss>.h5
                     └──best_model<epoch>-<val_loss>.h5
                     └── …
      ├──LR_0_000006_Dropout
            ├──Code
                └──Res50
                     ├──Density.npy
                     ├──Res50.py
                     └──test.json
            ├──Graphs
                └──Res50
                     ├──cosine
                     ├──loss
                     ├──mae
                     ├──mse
                     ├──msle
                     └──rmse
            ├──Log_Files
                └──Res50
                     └──log0.xlsx
            └──Models
                └──Res50
                     └──best_model<epoch>-<val_loss>.h5
                     └──best_model<epoch>-<val_loss>.h5
                     └──best_model<epoch>-<val_loss>.h5
                     └── …
      ├──LR_0_000006_Dropout_05
            ├──Code
                └──Res50
                     ├──Density.npy
                     ├──Res50.py
                     └──test.json
            ├──Graphs
                └──Res50
                     ├──cosine
                     ├──loss
                     ├──mae
                     ├──mse
                     ├──msle
                     └──rmse
            ├──Log_Files
                └──Res50
                     └──log0.xlsx
            └──Models
                └──Res50
                     └──best_model<epoch>-<val_loss>.h5
                     └──best_model<epoch>-<val_loss>.h5
                     └──best_model<epoch>-<val_loss>.h5
                     └── …
├──LR_0_00001
      ├──Graphs
            ├──Res50
                ├──loss
                └── mae
            ├──Res50_Conv4_5
                ├──loss
                └── mae
```

```
      ├──Res50_Conv_5
      │   ├──loss
      │   └── mae
      ├──Res101
      │   ├──loss
      │   └── mae
      ├──Res101_Conv_4_5
      │   ├──loss
      │   └── mae
      ├──Res101_Conv_5
      │   ├──loss
      │   └── mae
      ├──Res152
      │   ├──loss
      │   └── mae
      ├──Res152_Conv4_5
      │   ├──loss
      │   └── mae
      └──Res152_Conv_5
          ├──loss
          └── mae
├──Log_Files
      ├──Res50
      │   └──log0
      ├──Res50_Conv4_5
      │   └──log0
      ├──Res50_Conv_5
      │   └──log0
      ├──Res101
      │   └──log0
      ├──Res101_Conv_4_5
      │   └──log0
      ├──Res101_Conv_5
      │   └──log0
      ├──Res152
      │   └──log0
      ├──Res152_Conv4_5
      │   └──log0
      └──Res152_Conv_5
          └──log0
└──Models
      ├──Res50
      │   └──best_model<epoch>-<val_loss>.h5
      │   └──best_model<epoch>-<val_loss>.h5
      │   └── …
      ├──Res50_Conv4_5
      │   └──best_model<epoch>-<val_loss>.h5
      │   └──best_model<epoch>-<val_loss>.h5
      │   └── …
      ├──Res50_Conv_5
      │   └──best_model<epoch>-<val_loss>.h5
      │   └──best_model<epoch>-<val_loss>.h5
      │   └── …
      ├──Res101
      │   └──best_model<epoch>-<val_loss>.h5
      │   └──best_model<epoch>-<val_loss>.h5
      │   └── …
      ├──Res101_Conv_4_5
      │   └──best_model<epoch>-<val_loss>.h5
```

```
                └──best_model<epoch>-<val_loss>.h5
                └── …
            ├──Res101_Conv_5
                └──best_model<epoch>-<val_loss>.h5
                └──best_model<epoch>-<val_loss>.h5
                └── …
            ├──Res152
                └──best_model<epoch>-<val_loss>.h5
                └──best_model<epoch>-<val_loss>.h5
                └── …
            ├──Res152_Conv4_5
                └──best_model<epoch>-<val_loss>.h5
                └──best_model<epoch>-<val_loss>.h5
                └── …
            └──Res152_Conv_5
                └──best_model<epoch>-<val_loss>.h5
                └──best_model<epoch>-<val_loss>.h5
                └── …
├──LR_0_0001
      ├──Graphs
            ├──Res50
                ├──loss
                └── mae
            ├──Res50_Conv4_5
                ├──loss
                └── mae
            ├──Res50_Conv_5
                ├──loss
                └── mae
            ├──Res101
                ├──loss
                └── mae
            ├──Res101_Conv_4_5
                ├──loss
                └── mae
            ├──Res101_Conv_5
                ├──loss
                └── mae
            ├──Res152
                ├──loss
                └── mae
            ├──Res152_Conv4_5
                ├──loss
                └── mae
            └──Res152_Conv_5
                ├──loss
                └── mae
      ├──Log_Files
            ├──Res50
                └──log0
            ├──Res50_Conv4_5
                └──log0
            ├──Res50_Conv_5
                └──log0
            ├──Res101
                └──log0
            ├──Res101_Conv_4_5
                └──log0
            ├──Res101_Conv_5
```

```
                └──log0
            ├──Res152
                └──log0
            ├──Res152_Conv4_5
                └──log0
            └──Res152_Conv_5
                └──log0
        └──Models
            ├──Res50
                └──best_model<epoch>-<val_loss>.h5
                └──best_model<epoch>-<val_loss>.h5
                └── …
            ├──Res50_Conv4_5
                └──best_model<epoch>-<val_loss>.h5
                └──best_model<epoch>-<val_loss>.h5
                └── …
            ├──Res50_Conv_5
                └──best_model<epoch>-<val_loss>.h5
                └──best_model<epoch>-<val_loss>.h5
                └── …
            ├──Res101
                └──best_model<epoch>-<val_loss>.h5
                └──best_model<epoch>-<val_loss>.h5
                └── …
            ├──Res101_Conv_4_5
                └──best_model<epoch>-<val_loss>.h5
                └──best_model<epoch>-<val_loss>.h5
                └── …
            ├──Res101_Conv_5
                └──best_model<epoch>-<val_loss>.h5
                └──best_model<epoch>-<val_loss>.h5
                └── …
            ├──Res152
                └──best_model<epoch>-<val_loss>.h5
                └──best_model<epoch>-<val_loss>.h5
                └── …
            ├──Res152_Conv4_5
                └──best_model<epoch>-<val_loss>.h5
                └──best_model<epoch>-<val_loss>.h5
                └── …
            └──Res152_Conv_5
                └──best_model<epoch>-<val_loss>.h5
                └──best_model<epoch>-<val_loss>.h5
                └── …
├──LR_0_001
        ├──Graphs
            ├──Res50
                ├──loss
                └── mae
            ├──Res50_Conv4_5
                ├──loss
                └── mae
            ├──Res50_Conv_5
                ├──loss
                └── mae
            ├──Res101
                ├──loss
                └── mae
            ├──Res101_Conv_4_5
```

```
        ├──loss
        └── mae
    ├──Res101_Conv_5
        ├──loss
        └── mae
    ├──Res152
        ├──loss
        └── mae
    ├──Res152_Conv4_5
        ├──loss
        └── mae
    └──Res152_Conv_5
        ├──loss
        └── mae
├──Log_Files
    ├──Res50
        └──log0
    ├──Res50_Conv4_5
        └──log0
    ├──Res50_Conv_5
        └──log0
    ├──Res101
        └──log0
    ├──Res101_Conv_4_5
        └──log0
    ├──Res101_Conv_5
        └──log0
    ├──Res152
        └──log0
    ├──Res152_Conv4_5
        └──log0
    └──Res152_Conv_5
        └──log0
└──Models
    ├──Res50
        └──best_model<epoch>-<val_loss>.h5
        └──best_model<epoch>-<val_loss>.h5
        └── …
    ├──Res50_Conv4_5
        └──best_model<epoch>-<val_loss>.h5
        └──best_model<epoch>-<val_loss>.h5
        └── …
    ├──Res50_Conv_5
        └──best_model<epoch>-<val_loss>.h5
        └──best_model<epoch>-<val_loss>.h5
        └── …
    ├──Res101
        └──best_model<epoch>-<val_loss>.h5
        └──best_model<epoch>-<val_loss>.h5
        └── …
    ├──Res101_Conv_4_5
        └──best_model<epoch>-<val_loss>.h5
        └──best_model<epoch>-<val_loss>.h5
        └── …
    ├──Res101_Conv_5
        └──best_model<epoch>-<val_loss>.h5
        └──best_model<epoch>-<val_loss>.h5
        └── …
    ├──Res152
```

```
            └──best_model<epoch>-<val_loss>.h5
            └──best_model<epoch>-<val_loss>.h5
            └── …
        ├──Res152_Conv4_5
            └──best_model<epoch>-<val_loss>.h5
            └──best_model<epoch>-<val_loss>.h5
            └── …
        └──Res152_Conv_5
            └──best_model<epoch>-<val_loss>.h5
            └──best_model<epoch>-<val_loss>.h5
            └── …
├──LR_0_01
    ├──Graphs
        ├──Res50
            ├──loss
            └── mae
        ├──Res50_Conv4_5
            ├──loss
            └── mae
        ├──Res50_Conv_5
            ├──loss
            └── mae
        ├──Res101
            ├──loss
            └── mae
        ├──Res101_Conv_4_5
            ├──loss
            └── mae
        ├──Res101_Conv_5
            ├──loss
            └── mae
        ├──Res152
            ├──loss
            └── mae
        ├──Res152_Conv4_5
            ├──loss
            └── mae
        └──Res152_Conv_5
            ├──loss
            └── mae
    ├──Log_Files
        ├──Res50
            └──log0
        ├──Res50_Conv4_5
            └──log0
        ├──Res50_Conv_5
            └──log0
        ├──Res101
            └──log0
        ├──Res101_Conv_4_5
            └──log0
        ├──Res101_Conv_5
            └──log0
        ├──Res152
            └──log0
        ├──Res152_Conv4_5
            └──log0
        └──Res152_Conv_5
            └──log0
```

```
    └──Models
          ├──Res50
              └──best_model<epoch>-<val_loss>.h5
              └──best_model<epoch>-<val_loss>.h5
              └── …
          ├──Res50_Conv4_5
              └──best_model<epoch>-<val_loss>.h5
              └──best_model<epoch>-<val_loss>.h5
              └── …
          ├──Res50_Conv_5
              └──best_model<epoch>-<val_loss>.h5
              └──best_model<epoch>-<val_loss>.h5
              └── …
          ├──Res101
              └──best_model<epoch>-<val_loss>.h5
              └──best_model<epoch>-<val_loss>.h5
              └── …
          ├──Res101_Conv_4_5
              └──best_model<epoch>-<val_loss>.h5
              └──best_model<epoch>-<val_loss>.h5
              └── …
          ├──Res101_Conv_5
              └──best_model<epoch>-<val_loss>.h5
              └──best_model<epoch>-<val_loss>.h5
              └── …
          ├──Res152
              └──best_model<epoch>-<val_loss>.h5
              └──best_model<epoch>-<val_loss>.h5
              └── …
          ├──Res152_Conv4_5
              └──best_model<epoch>-<val_loss>.h5
              └──best_model<epoch>-<val_loss>.h5
              └── …
          └──Res152_Conv_5
              └──best_model<epoch>-<val_loss>.h5
              └──best_model<epoch>-<val_loss>.h5
              └── …
├──my_dir
    └──ResNet50_Hyperparameter_Optimization
          └──trial0
          └──trial1
          └── …
└──Dataset
    └──PsAreTruth_E-20211020T171415Z-001-F3
          └──PsAreTruth_E
              ├──contours
                    ├──<filename>.csv
                    ├──<filename>.csv
                    └── …
              ├──mask
                    ├──<filename>.png
                    ├──<filename>.png
                    └── …
              ├──rgbs
                    ├──<filename>.png
                    ├──<filename>.png
                    └── …
              └──vis
                    ├──<filename>.png
```

```
├──&lt;filename&gt;.png
└── …
```

Working directory is the root directory of our project. In the Ep_100 folder models which ever trained for 100 Epochs can be found along with their codes, graphs, log files, and best saved models. The LR_0_00001 folder consists of the models and supported files which ever iterations carried out with 1e-4 learning rate. similarly, the folder LR_0_0001 consists of the models and supported files which ever iterations carried out with 1e-3 learning rate. The same carries with the folder LR_0_001 and LR_0_01. Moreover, the folder my_dir consists of the files which ever belongs to hyperparameter optimizations. On top of that, the codes folder carries the script files for the LR_0_00001, LR_0_0001, LR_0_001, LR_0_01. Same codes belong to all folders, but we have to manually change the learning rates and destinations.

Code files Explanation:

Data_Extractor: This file extracts the information from the .csv files and save the information is test.json file.

Density_Finder: This file is not called in the code, but the lines used in this code is directly used in every code. For instance, Figure 1, represents the coding lines used in every codes.

Res50.py: This file is the complete model building file of the model Resnet 50 which trains all its layers.

Res50_Conv_4_5.py: This file is the complete model building file of the model Resnet 50 which does not train the layer 4 and 5.

Res50_Conv_5.py: This file is the complete model building file of the model Resnet 50 which does not train the layer 5.

Res101.py: This file is the complete model building file of the model Resnet 50 which trains all its layers.

Res101_Conv_4_5.py: This file is the complete model building file of the model Resnet 50 which does not train the layer 4 and 5.

Res101_Conv_5.py: This file is the complete model building file of the model Resnet 50 which does not train the layer 5.

Res152.py: This file is the complete model building file of the model Resnet 50 which trains all its layers.

Res152_Conv_4_5.py: This file is the complete model building file of the model Resnet 50 which does not train the layer 4 and 5.

Res152_Conv_5.py: This file is the complete model building file of the model Resnet 50 which does not train the layer 5.

Extracting density:

```
########################
dictlist = []
def read_json(file):
  data = json.load(file)
  for key, value in data.items():
    val = list(value.values())
    tumour = val[0]
    total = val[1]
    result = tumour / total
    dictlist.append(result)
    np.save('Density', dictlist)

with open("test.json") as f:
    read_json(f)

labels = np.load('Density.npy')
#labels = labels.tolist() #y as array dtype = np.float32

print('labels length: ', len(labels))
tr_label = labels[:1000]
vl_label = labels[1000:]
print('train labels length: ', len(tr_label))
print('val labels length: ', len(vl_label))
```

Figure 1

The read_json function reads test.json file which has the extracted information of the .csv files. Then, a simple manipulation takes place for finding the density of tumour with the count of the cells present. Furthermore, the extracted information will be saved as a numpy array in the file called Density.npy.