Computer Vision

Assignment 4 - Image Segmentation

February 1, 2022

Ahmed Ashraf	5803
Mohamed Aiman	6017
Abdelrahman Habib	6019
Osama Sherif	6012

Contents

1 Data preparation

The data was resized, to make it smaller so that it will be easy too upload and train our model. The compressed size for each image in the data set size is (256 x 256 x 3):

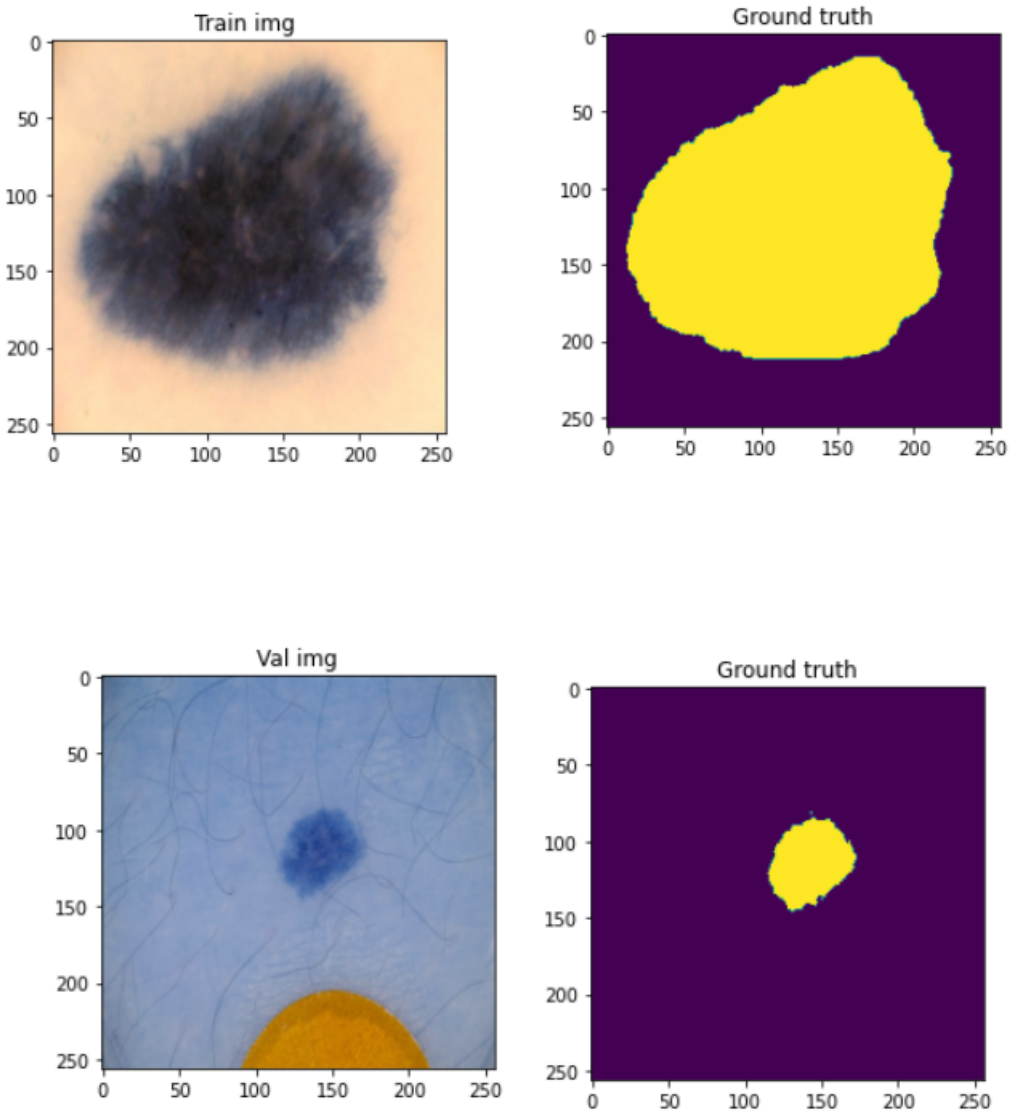


Figure 1: samples form the Train and val imges with there corresponding Ground truth

2 Training process

Training Details

- The models was trained with different optimizers:
  - SGD.
  - RMSprop.
  - $ADAM$.
- The $Jaccard - Loss$ was implemented to be the loss function for the model.
- Evalutaion metric :
  - IoU score
  - Jacard score.

3 UNet

3.1 Network Architecture

```
Model: "model_3"
__________________________________________________________________________________________________
Layer (type)                    Output Shape         Param #     Connected to
==================================================================================================
image (InputLayer)              [(None, 256, 256, 3) 0
__________________________________________________________________________________________________
lambda (Lambda)                 (None, 256, 256, 3)  0           image[0][0]
__________________________________________________________________________________________________
conv2d_15 (Conv2D)              (None, 256, 256, 16) 448         lambda[0][0]
__________________________________________________________________________________________________
conv2d_16 (Conv2D)              (None, 256, 256, 16) 2320        conv2d_15[0][0]
__________________________________________________________________________________________________
max_pooling2d_5 (MaxPooling2D)  (None, 128, 128, 16) 0           conv2d_16[0][0]
__________________________________________________________________________________________________
conv2d_17 (Conv2D)              (None, 128, 128, 32) 4640        max_pooling2d_5[0][0]
__________________________________________________________________________________________________
conv2d_18 (Conv2D)              (None, 128, 128, 32) 9248        conv2d_17[0][0]
__________________________________________________________________________________________________
max_pooling2d_6 (MaxPooling2D)  (None, 64, 64, 32)   0           conv2d_18[0][0]
__________________________________________________________________________________________________
conv2d_19 (Conv2D)              (None, 64, 64, 64)   18496       max_pooling2d_6[0][0]
__________________________________________________________________________________________________
conv2d_20 (Conv2D)              (None, 64, 64, 64)   36928       conv2d_19[0][0]
__________________________________________________________________________________________________
max_pooling2d_7 (MaxPooling2D)  (None, 32, 32, 64)   0           conv2d_20[0][0]
__________________________________________________________________________________________________
conv2d_21 (Conv2D)              (None, 32, 32, 128)  73856       max_pooling2d_7[0][0]
__________________________________________________________________________________________________
conv2d_22 (Conv2D)              (None, 32, 32, 128)  147584      conv2d_21[0][0]
__________________________________________________________________________________________________
max_pooling2d_8 (MaxPooling2D)  (None, 16, 16, 128)  0           conv2d_22[0][0]
__________________________________________________________________________________________________
conv2d_23 (Conv2D)              (None, 16, 16, 256)  295168      max_pooling2d_8[0][0]
__________________________________________________________________________________________________
conv2d_24 (Conv2D)              (None, 16, 16, 256)  590080      conv2d_23[0][0]
__________________________________________________________________________________________________
conv2d_transpose_1 (Conv2DTrans (None, 32, 32, 128)  131200      conv2d_24[0][0]
__________________________________________________________________________________________________
concatenate (Concatenate)       (None, 32, 32, 256)  0           conv2d_transpose_1[0][0]
                                                                 conv2d_22[0][0]
__________________________________________________________________________________________________
conv2d_25 (Conv2D)              (None, 32, 32, 256)  590080      concatenate[0][0]
__________________________________________________________________________________________________
conv2d_26 (Conv2D)              (None, 32, 32, 256)  590080      conv2d_25[0][0]
__________________________________________________________________________________________________
conv2d_transpose_2 (Conv2DTrans (None, 64, 64, 64)   65600       conv2d_26[0][0]
__________________________________________________________________________________________________
concatenate_1 (Concatenate)     (None, 64, 64, 128)  0           conv2d_transpose_2[0][0]
                                                                 conv2d_20[0][0]
__________________________________________________________________________________________________
conv2d_27 (Conv2D)              (None, 64, 64, 128)  147584      concatenate_1[0][0]
__________________________________________________________________________________________________
conv2d_28 (Conv2D)              (None, 64, 64, 128)  147584      conv2d_27[0][0]
__________________________________________________________________________________________________
conv2d_transpose_3 (Conv2DTrans (None, 128, 128, 32) 16416       conv2d_28[0][0]
__________________________________________________________________________________________________
concatenate_2 (Concatenate)     (None, 128, 128, 64) 0           conv2d_transpose_3[0][0]
                                                                 conv2d_18[0][0]
__________________________________________________________________________________________________
conv2d_29 (Conv2D)              (None, 128, 128, 64) 36928       concatenate_2[0][0]
__________________________________________________________________________________________________
conv2d_30 (Conv2D)              (None, 128, 128, 64) 36928       conv2d_29[0][0]
__________________________________________________________________________________________________
conv2d_transpose_4 (Conv2DTrans (None, 256, 256, 16) 4112        conv2d_30[0][0]
__________________________________________________________________________________________________
concatenate_3 (Concatenate)     (None, 256, 256, 32) 0           conv2d_transpose_4[0][0]
                                                                 conv2d_16[0][0]
__________________________________________________________________________________________________
conv2d_31 (Conv2D)              (None, 256, 256, 32) 9248        concatenate_3[0][0]
__________________________________________________________________________________________________
conv2d_32 (Conv2D)              (None, 256, 256, 32) 9248        conv2d_31[0][0]
__________________________________________________________________________________________________
conv2d_33 (Conv2D)              (None, 256, 256, 1)  33          conv2d_32[0][0]
==================================================================================================
Total params: 2,963,809
Trainable params: 2,963,809
Non-trainable params: 0
```

3.2 Results

Different LR results

Figure 3: lr=0.1

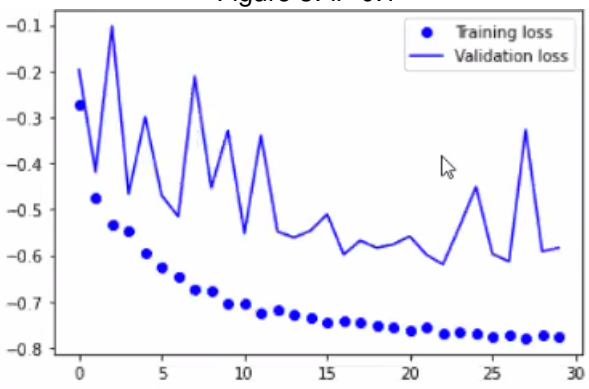


Figure 4: lr=0.01

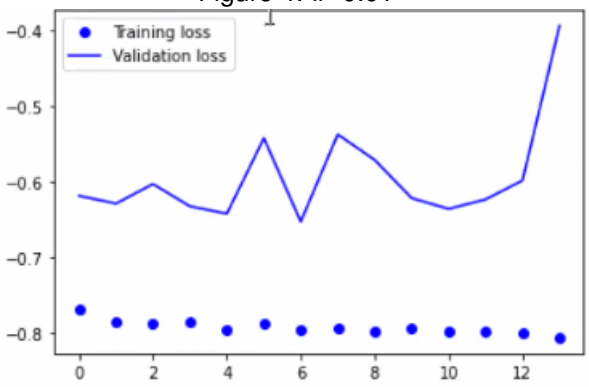


Figure 5: lr=0.001

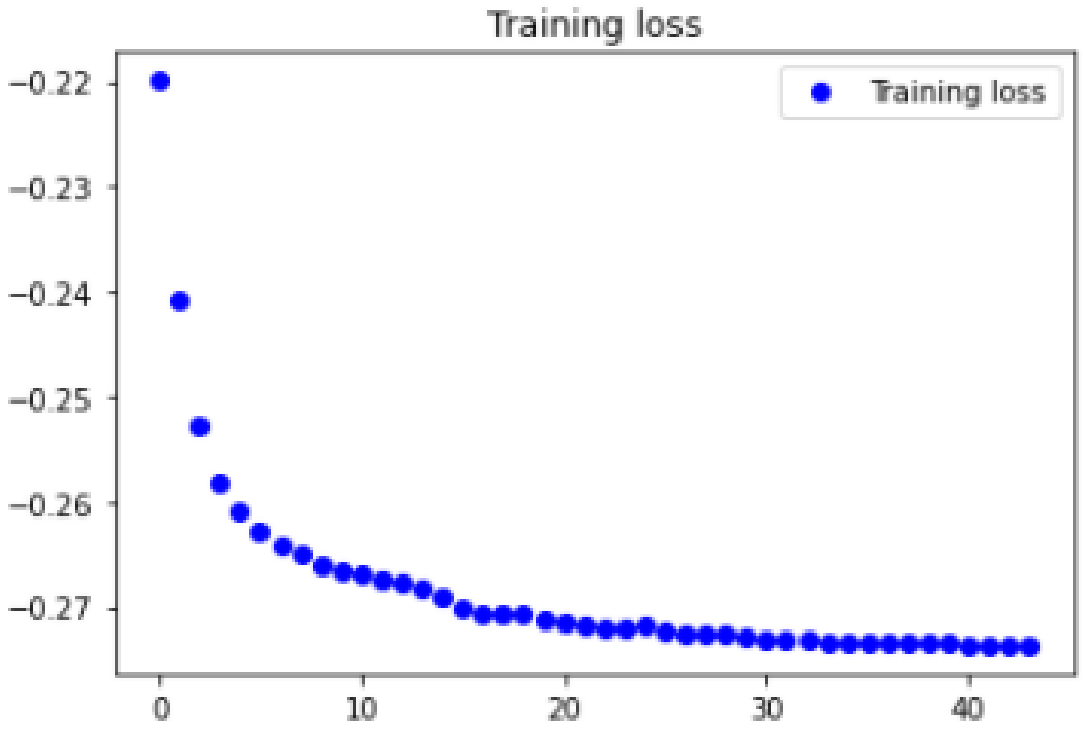
4 Fully Convolution Network (FCN):

4.1 Network Architecture

```
Model: "model"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
image (InputLayer)           [(None, 256, 256, 3)]     0
_________________________________________________________________
conv2d (Conv2D)              (None, 256, 256, 64)      1792
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 256, 256, 64)      36928
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 128, 128, 64)      0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 128, 128, 128)     73856
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 128, 128, 128)     147584
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 64, 64, 128)       0
_________________________________________________________________
conv2d_4 (Conv2D)            (None, 64, 64, 256)       295168
_________________________________________________________________
conv2d_5 (Conv2D)            (None, 64, 64, 256)       590080
_________________________________________________________________
conv2d_6 (Conv2D)            (None, 64, 64, 256)       590080
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 32, 32, 256)       0
_________________________________________________________________
conv2d_7 (Conv2D)            (None, 32, 32, 512)       1180160
_________________________________________________________________
conv2d_8 (Conv2D)            (None, 32, 32, 512)       2359808
_________________________________________________________________
conv2d_9 (Conv2D)            (None, 32, 32, 512)       2359808
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 16, 16, 512)       0
_________________________________________________________________
conv2d_10 (Conv2D)           (None, 16, 16, 512)       2359808
_________________________________________________________________
conv2d_11 (Conv2D)           (None, 16, 16, 512)       2359808
_________________________________________________________________
conv2d_12 (Conv2D)           (None, 16, 16, 512)       2359808
_________________________________________________________________
max_pooling2d_4 (MaxPooling2 (None, 8, 8, 512)         0
_________________________________________________________________
conv2d_13 (Conv2D)           (None, 8, 8, 4096)        2101248
_________________________________________________________________
conv2d_14 (Conv2D)           (None, 8, 8, 1)           4097
_________________________________________________________________
conv2d_transpose (Conv2DTran (None, 256, 256, 1)       50
=================================================================
Total params: 16,820,083
Trainable params: 16,820,083
Non-trainable params: 0
```

4.2 Results

Over fitting a Small Batch of Data



Different LR results

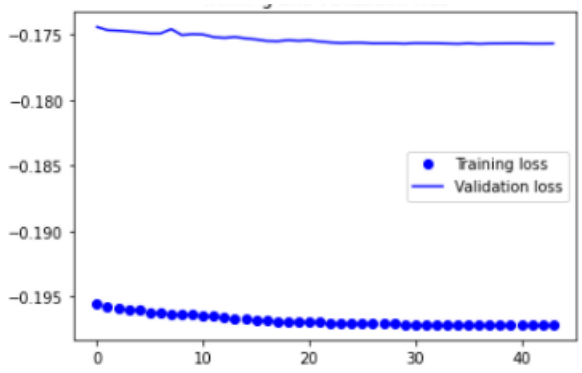


Figure 7: lr=0.1

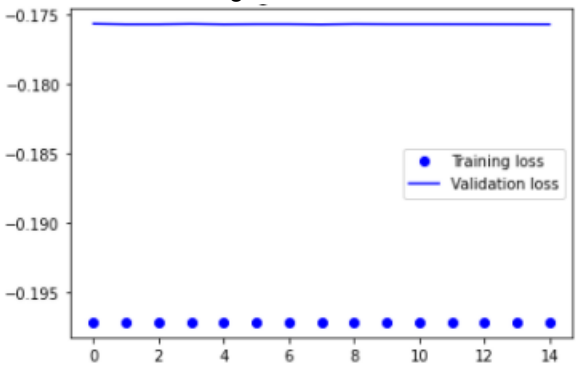


Figure 8: lr=0.01

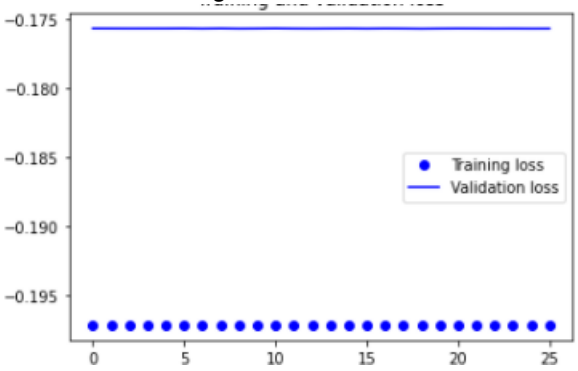


Figure 9: lr=0.001

5 FCN-pretrained model

5.1 Network Architecture

Layer (type)	Output Shape	Param #
image (InputLayer)	[(None, 256, 256, 3)]	0
vgg16 (Functional)	(None, 8, 8, 512)	14714688
global_average_pooling2d (Gl	(None, 512)	0
dense (Dense)	(None, 256)	131328
reshape (Reshape)	(None, 16, 16, 1)	0
conv2d_transpose_5 (Conv2DTr	(None, 32, 32, 16)	160
conv2d_transpose_6 (Conv2DTr	(None, 64, 64, 32)	4640
conv2d_transpose_7 (Conv2DTr	(None, 128, 128, 64)	18496
conv2d_transpose_8 (Conv2DTr	(None, 256, 256, 2)	1154
conv2d_34 (Conv2D)	(None, 256, 256, 1)	3

Total params: 14,870,469
Trainable params: 155,781
Non-trainable params: 14,714,688

Figure 10: U-Net architecture

5.2 Results

Different LR results

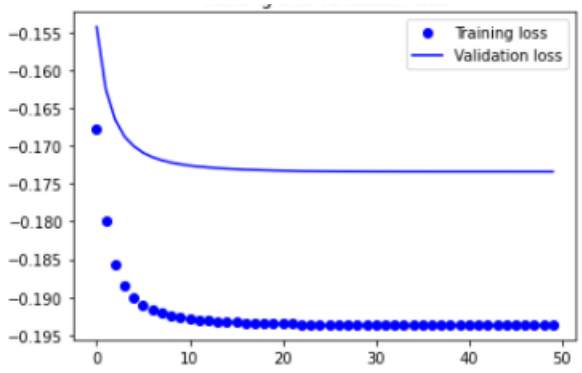


Figure 11: lr=0.1

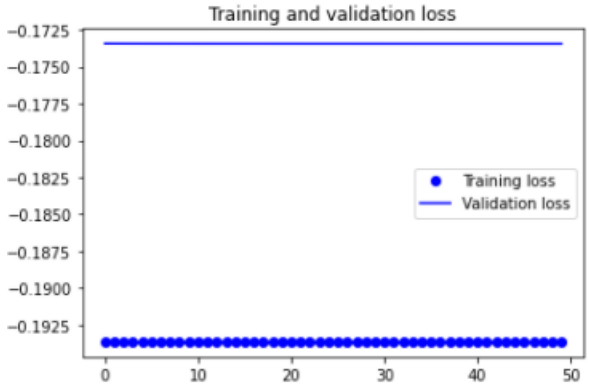


Figure 12: lr=0.01

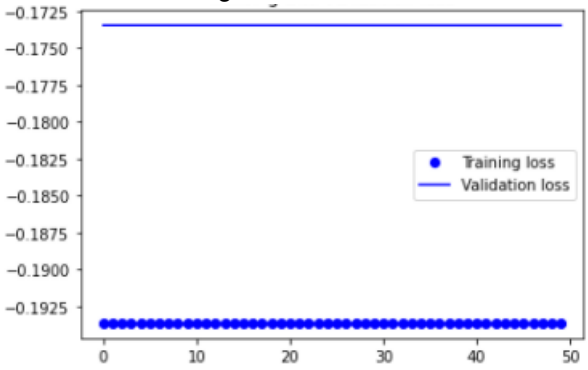
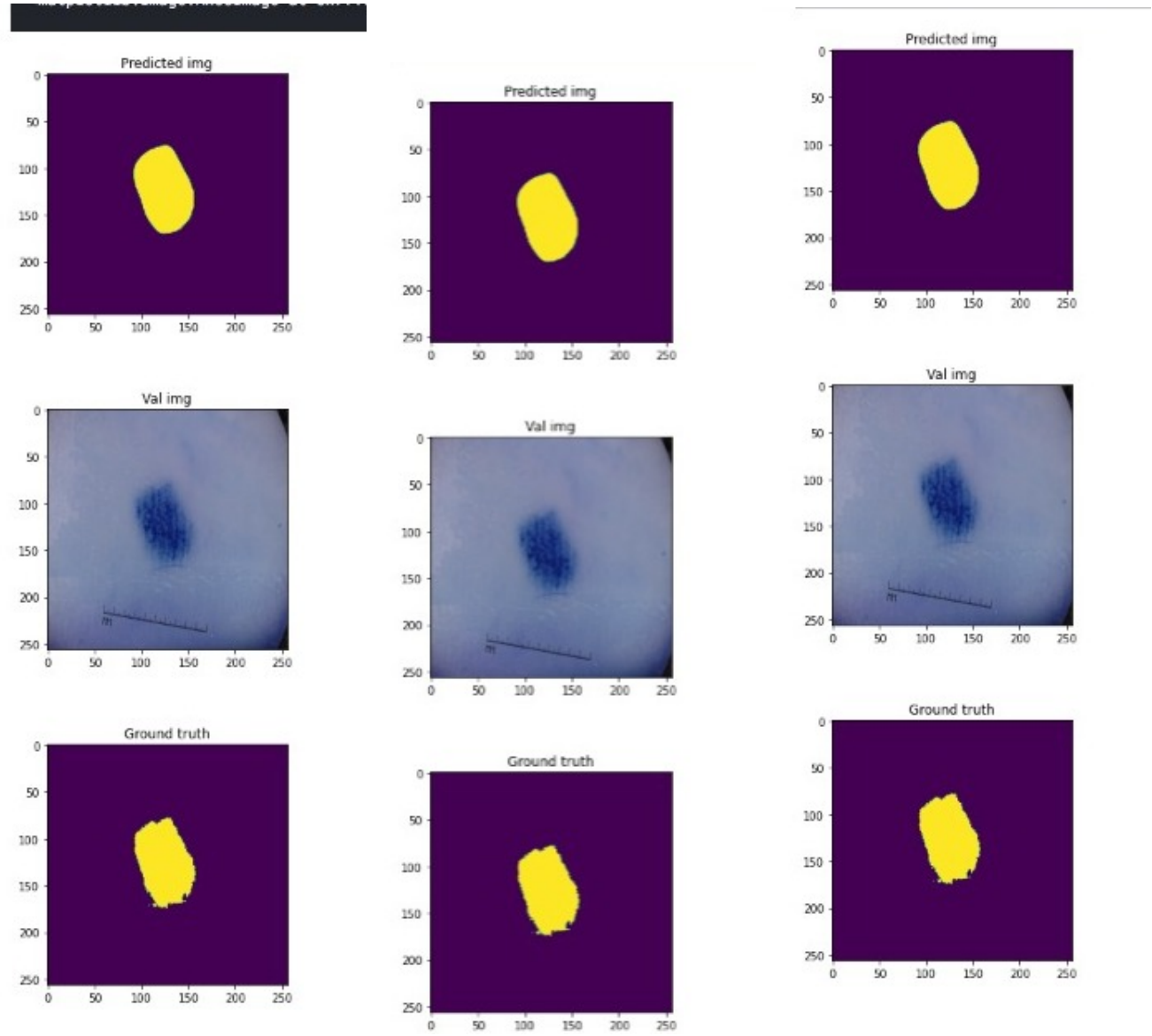


Figure 13: lr=0.001

6 Test cases

Identify some failure and success cases :

6.1 UNet-results



6.2 FCN-results