Bachelor's project

Image-based quality evaluation of otoscopy images

Weekly Report 9

Freja Rindel Peulicke, s185393

October 23, 2021

Literature

What has been done this week

1 Metric performance

1.1 Variance of Laplacian

This metric has been implemented for both jpg and png filetypes.

Outlier
An outlier in the sharp images have been discovered. As can be seen in the image below, a few images are classified as much sharper than the rest of the images with a score of around 6 compared to the rest of the scores around 1.

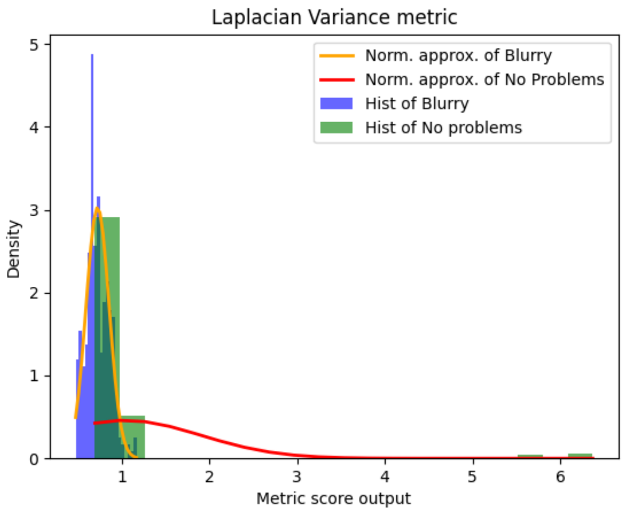
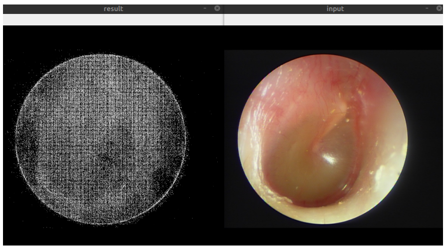
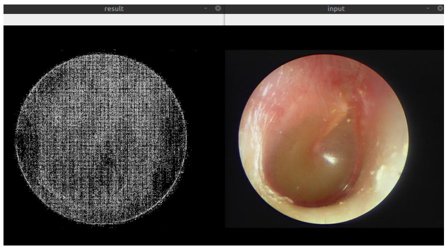


Figure 1: update diagram* A density plot on the training data of filetype jpg. The true blurry images are visualized in the blue histogram and the orange pdf-curve, and the sharp images are vizualised by the green histogram and the red pdf-curve. No gaussian blurred images have been included.

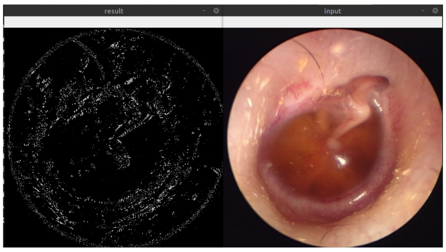
Displaying the image after applying the laplacian filter in the algorithm produces the output displayed in figure 2.



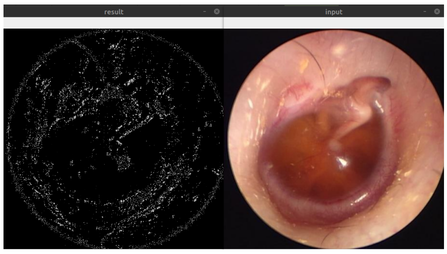
(a) The right image is the input image of type png. The left image is the result of applying a laplacian filter on it.



(b) The right image is the input image of type jpg. The left image is the result of applying a laplacian filter on it.



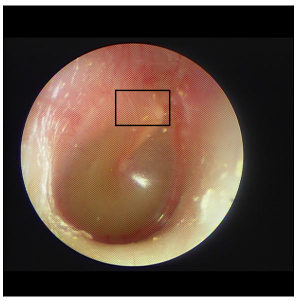
(c) The right image is the input image of type png. The left image is the result of applying a laplacian filter on it.



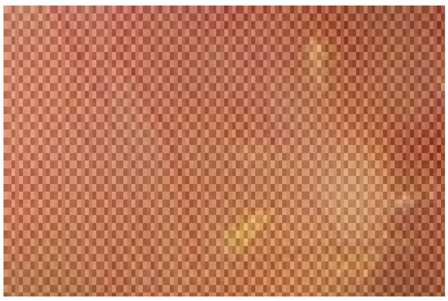
(d) The right image is the input image of type jpg. The left image is the result of applying a laplacian filter on it.

Figure 2: The above images displays the laplacian output of two different true sharp images of filetype png and jpg. The lower images represent a normal output of the laplacian filter, while the upper images produces outliers.

Looking closely at the original image (both jpg and png), one can see a filter-like pattern, which produces the very white output.



(a)

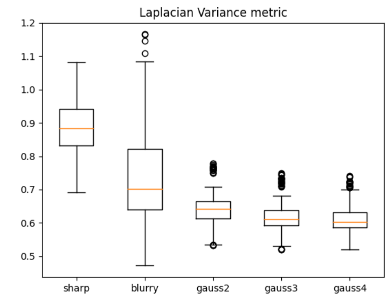


(b)

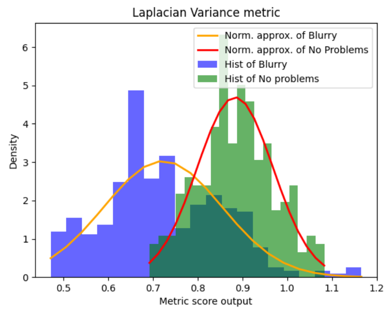
This image and the synthetic sharp data produced from it will be disregarded from the following analysis.

JPG

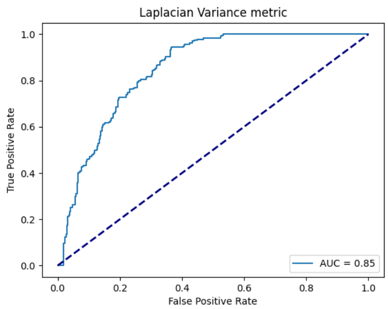
The scores on the jpg images are displayed in figure 4.



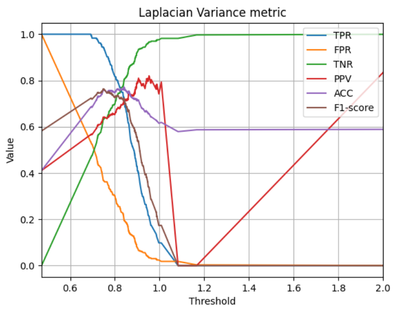
(a) Box plots on the sharp and blurry images including the rotated and mirrored images, and the gaussian blurred images with standard deviation of respectively 2, 3 and 4. The trend is that the median of the sharp images is above the median of the blurry images whose median is above all gaussian blurred medians. The medians of the gaussian images drops with the standard deviation, as expected.



(b) Density plot for sharp and blurry images without gaussian blur. The overlap between the scores of the two datasets is the smallest one seen at this point. The overlap represents the expected amount of falsely classified images.



(c) The outputs provide a fine roc curve with AUC = 0.85, the highest at this point.



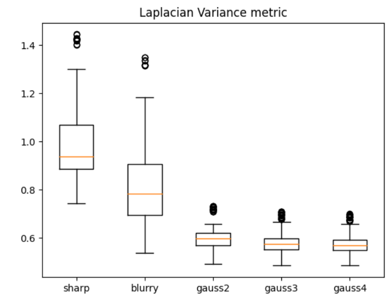
(d) All of the final values in the PPV curves have been set to 1, as TP and FP in the denominator are both 0 at the highest threshold used. Here, the highest threshold is 2, which is greater than the highest output of about 1.2

Figure 4: Visualization of the output of variance of laplacian metric scores on input images of type jpg.

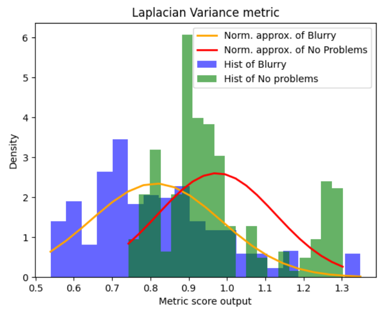
The metric scores very well on input images of type jpg.

PNG

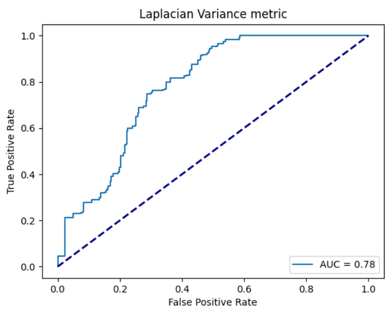
In figure 5 the outputs on the training data set of type png are displayed.



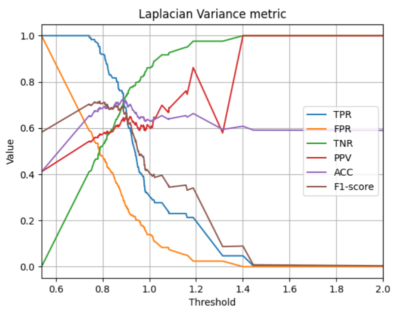
(a) As is also the case for the jpg's, the median drops as the images becomes more blurry, assuming that the gaussian blurred images are more blurred than the originally blurred images. This could very well be, as the gaussian filter applies a smoothening effect to the image, whereas parts of images in inmanipulated blurred images may be sharp. However, the shared interval of sharp and blurry is bigger for this box plot than for the jpg one.



(b) The overlapping area of this density plot constitutes a bigger part of the area under the pdf's than in the density plot on the jpg output. Thus, the metric performs worse on the training data of png's than of jpg's.



(c) This roc curve with an AUC = 0.78 agrees with the density plot, that the metric performs better on the jpg's.



(d)

Figure 5: Vizualization of different rate and accuracy values depending on different thresholds.

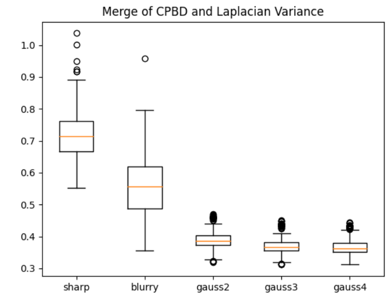
The metric performs better on the training set jpg images.

1.2 Combining metrics

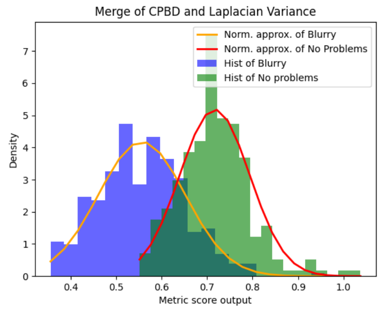
The three metrics, CPBD, Variance of Lapracian and Histogram Frequency-based metric, will be combined to form 4 new metrics. This is done by combining the outputs of the metrics as $m\_new_i = \alpha \cdot m1_i + (1-\alpha) \cdot m2_i$, where $m1$ and $m2$ are two different metrics and $i \in$ all outputs.

1.2.1 Merge of CPBD and Variance of Laplacian

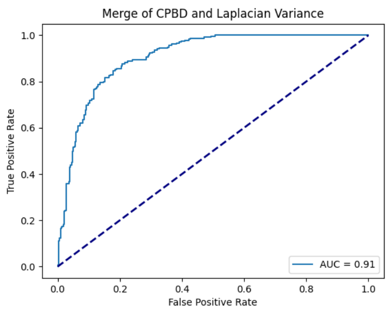
JPG



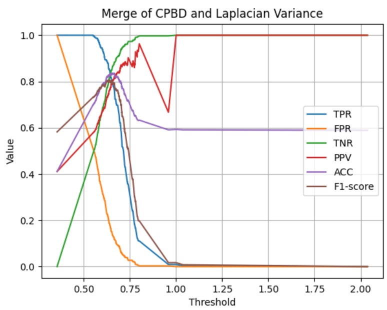
(a)



(b)



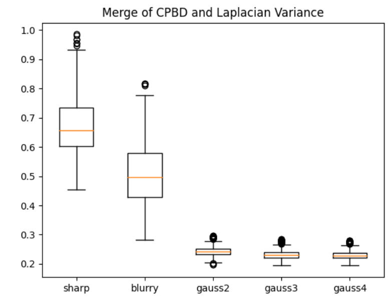
(c) AUC = 0.91



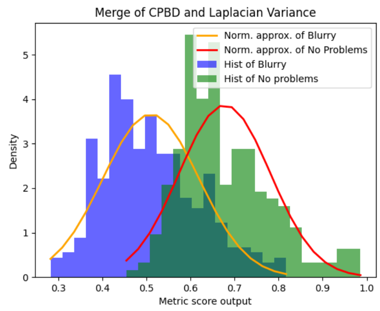
(d)

Figure 6

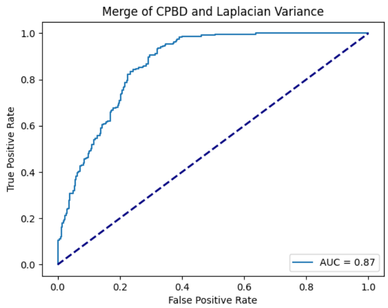
PNG



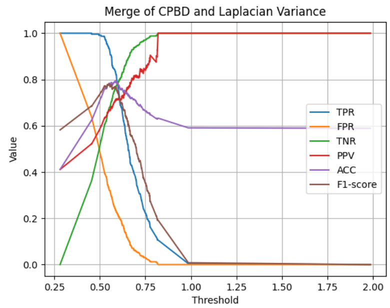
(a)



(b)



(c) AUC = 0.87



(d)

Figure 7

1.3 "Training" the metrics

The goal is to avoid false positives (**FP**: classified as sharp when blurry), as these can distort the results from the data later on, while still achieving some positive outputs.

Choosing the threshold:
Let the user choose an acceptable TNR (specificity), e.g. 98%.
After this, find the corresponding threshold and choose the metric with the highest accuracy, as it will provide a possibility that some images will be "accepted", that is, classified as sharp.

We could also calculate summed distance, d, of some rates... TPR (sensitivity), F1-score, precision... to the top and bottom border (1 and 0) according to which border is desired for that particular rate. Then choose the metric with smallest d.

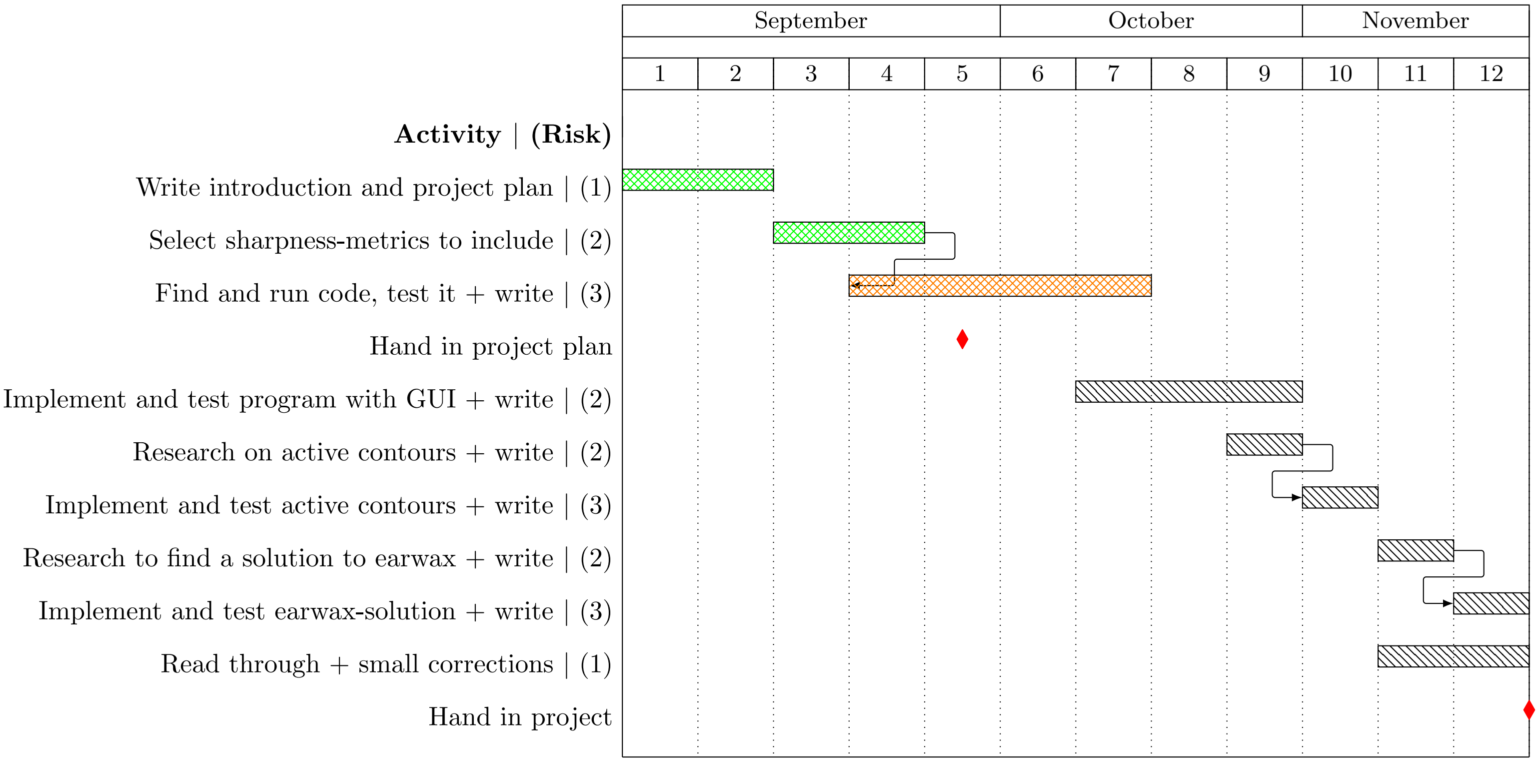
- maximize TNR (how many are blurry when blurry)
- maximize TPR (we would still like some possibility that images are classified "sharp")
- maximize PPV (part declared sharp when sharp)

- high f1-score for general score + do not produce FP (to not filter away too many negatives)

The following graphs displays output on the training data set excluding the Gaussian blurred images.

Project status according to the study plan

Overall Project Plan



Plan for the next weeks

1. Decide parameters to choose metric on (e.g. let user decide FNR and then minimize/maximize some parameter and choose the metric, that performs best on those criteria)
2. Implement choosing metric
3. Test the metrics on a test data set (the images not used for training) and report results
4. Start designing and implementing the final program

References