Bachelor's project

Image-based quality evaluation of otoscopy images

Weekly Report 7

Freja Rindel Peulicke, s185393

October 22, 2021

Literature

What has been done this week

1 Choosing the best metric

In this section, the four metrics will be given different parameters to make them perform their best.[1]

1.1 Tweaking parameters

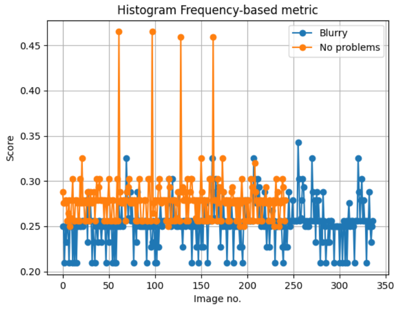
Some of the chosen metrics have constants that can be adjusted to optimize the correctness of the output scores. This concerns the histogram frequency-based and the frequency domain metrics.

The best possible constants have been chosen experimentally, looking at the AUC value of the outputs when running the metric on the training data set. The constants will be chosen between those that provide the highest AUC values on the training data set both including and excluding the Gaussian blurred images.

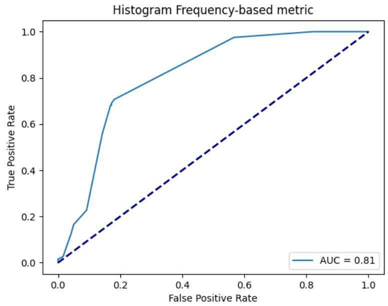
1.1.1 HF

In the histogram frequency-based metric, the two thresholds, $minDCTValue$ and $maxHistValue$, can be adjusted for optimizing the output.

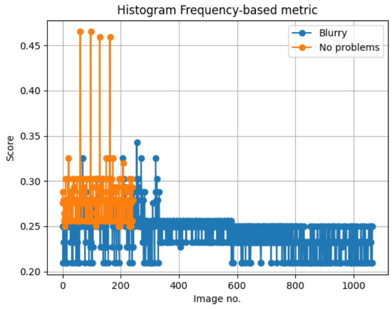
[1] rewrite



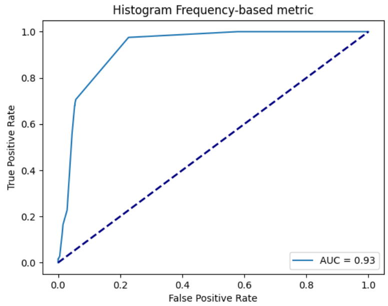
(a)



(b) One of the two best performances without Gauss have constants $minDCTValue = 1$ and $maxHistValue = 0.085$, having AUC$= 0.81$.

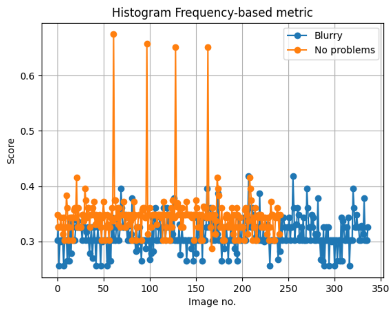


(c)

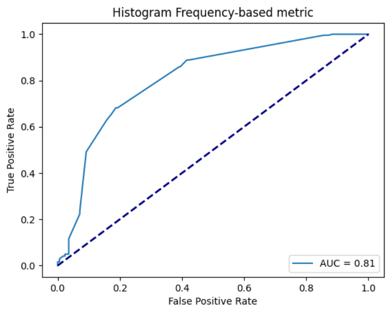


(d) The output on the training set including Gauss on constants $minDCTValue = 1$ and $maxHistValue = 0.085$, having AUC$= 0.93$.

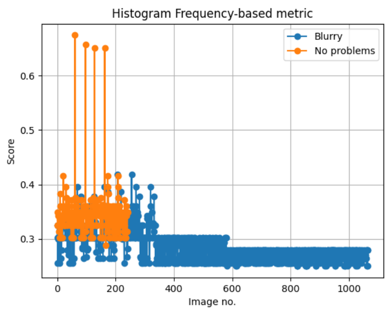
Figure 1: The output with $minDCTValue = 1$ and $maxHistValue = 0.085$.



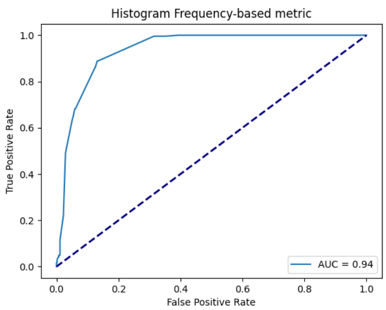
(a)



(b) One of the two best performances on the training set with $minDCTValue = 0$, $maxHistValue = 0.115$ and without Gauss, having AUC$= 0.81$.



(c)



(d) The best performance with Gauss is with $minDCTValue = 0$ and $maxHistValue = 0.115$, having AUC$=$ 0.94.

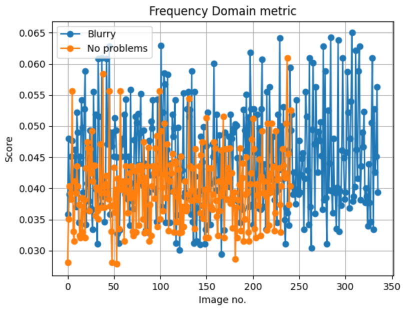
Figure 2: The output with $minDCTValue = 0$ and $maxHistValue = 0.115$.

The correctness of the output of the algorithm doesn't seem to be distorted by the Gaussian blurred images. In figure **??** the ROC curves with highest scores are displayed. As can be seen, the output of running the algorithm on the training set without Gaussian blurred images performs very well in two different situations. The correctness is also very high and almost identical for the results including the Gaussian blurred images. However, it is a little better with the lower $minDCTValue = 0$, whose purpose is filtering away noise. That is, when no noise is filtered away.

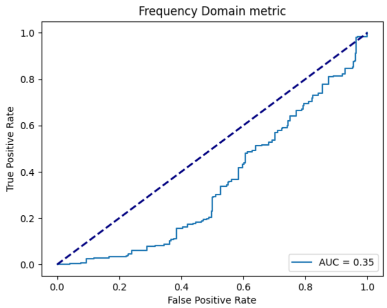
Knowing this, the $minDCTValue = 1$ and $maxHistValue = 0.085$ are chosen for the metric, as the purpose of the metric is not detecting Gaussian blur, but blur in non-manipulated images, where noise could occur, and as the worsening of the results on the Gaussian blurred images is insignificant.

1.1.2 FM

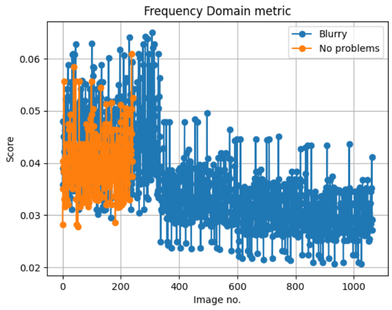
In the frequency domain image blur measure, the threshold $t = \dfrac{M}{1000}$ is experimentally estimated [2]. This will also be done in the following, where the constant for division (1000) is varied.



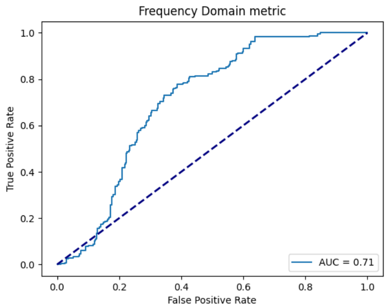
(a)



(b) The worst performance with constant 102 and no Gauss, having AUC= 0.35.



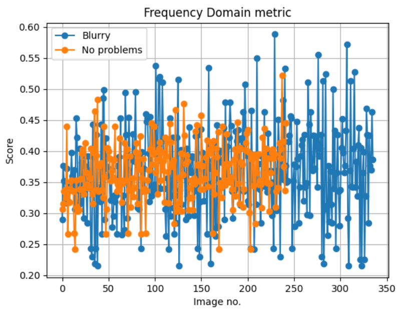
(c)



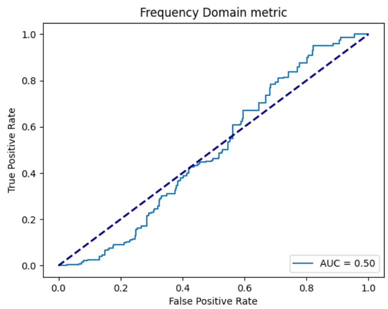
(d) The performance with constant 102 on the training set including Gauss, having AUC= 0.71.

Figure 3: The output with a constant of value 102

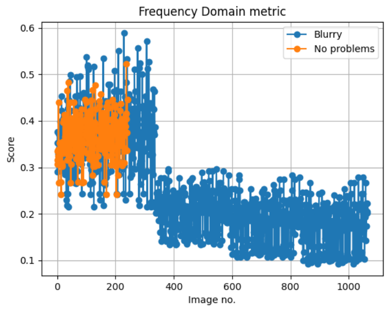
[2] *Experimentally it is observed that this particular threshold value gives a fairly accurate sense of image quality*[**FM**]



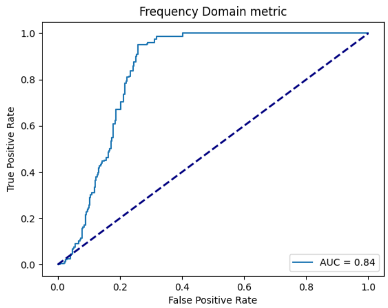
(a)



(b) The AUC= 0.5 on the training set without Gauss with the constant 2202.



(c)



(d) One of the best performances with Gauss is with the constant 2202, having AUC= 0.84.

Figure 4: The output with a constant of value 2202

The Gaussian blurred images will not be taking into account when choosing the constant for the metric, as they distort the result. This can be seen in the above images, where the metric performs excellent on the training set including the Gaussian images, while at the same time it performs quite badly on the training set without. The AUC in the tests on the training set without Gauss did not manage to exceed 53%, however, the lowest value was 35%, which gives an AUC of $100\% - 35\% = 65\%$. However, doing this, the metric would classify the Gaussian blurred images as sharper when applying more blur.

As this metric performs worst of all 4, it will not be taken into account in the following sections.

1.2 "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%.

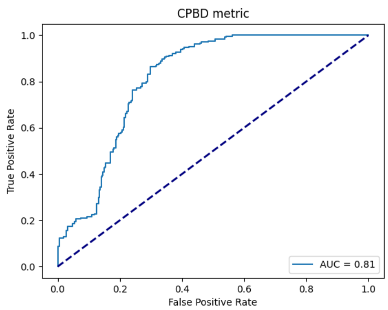
After this, find the corresponding threshold and choose the metric with the highest accuracy, as it will provide a possibility that some images will be ”accepted”, that is, classified as sharp.

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.

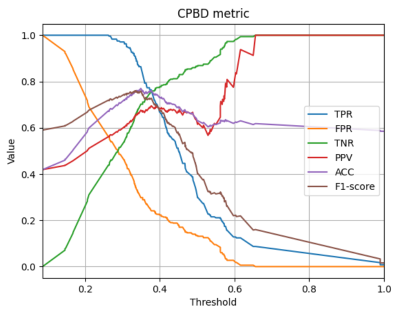
- maximize TNR (how many are blurry when blurry)
- maximize TPR (we would still like some possibility that images are classified ”sharp”)
- maximize PPV (part declared sharp when sharp)
- high f1-score for general score + do not produce FP (to not filter away too many negatives)

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

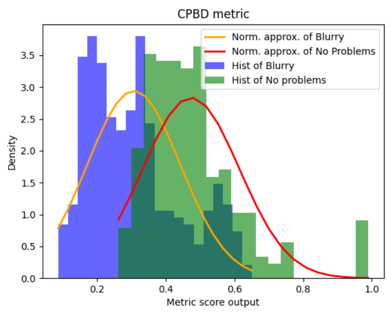
1.3 CPBD



(a)



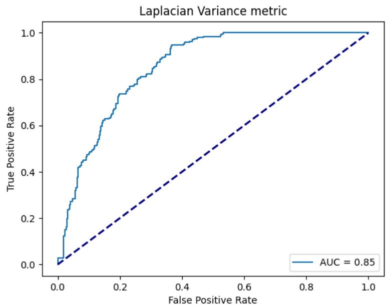
(b)



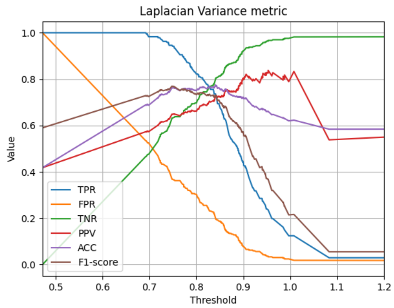
(c)

Figure 5

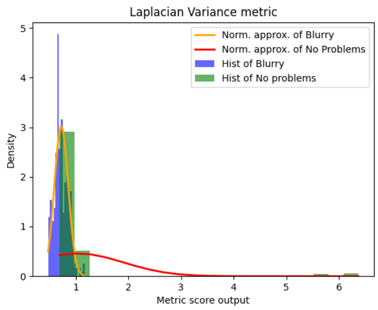
1.4 Variance of Laplacian



(a)



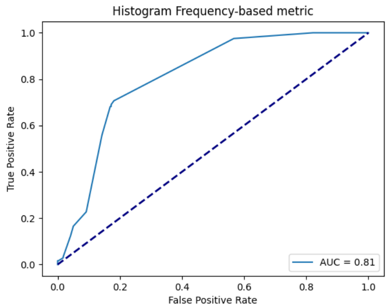
(b)



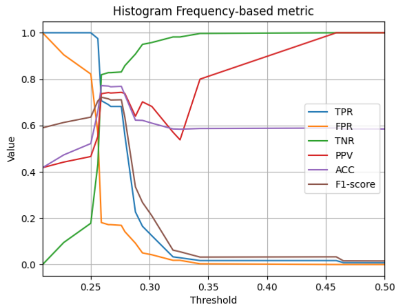
(c)

Figure 6

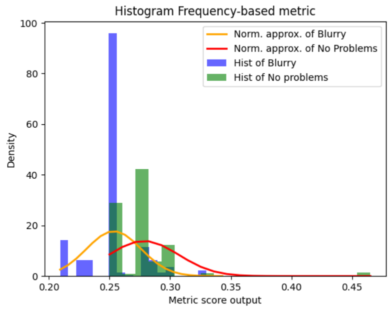
1.5 Histogram frequency-based metric



(a)



(b)

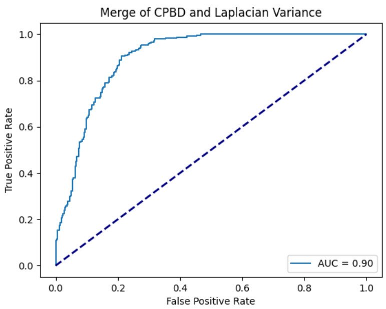


(c)

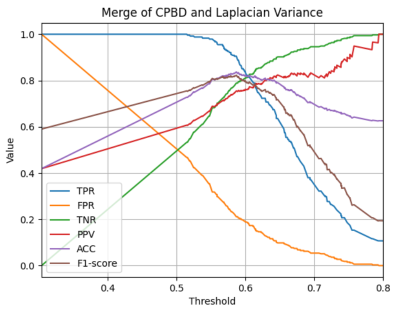
Figure 7

1.6 Merge of CPBD and Variance of Laplacian

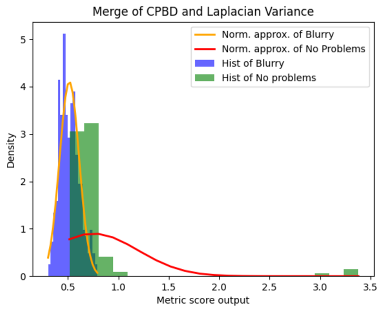
Highest AUC = 90%


Merge of CPBD and Laplacian Variance
True Positive Rate
False Positive Rate
1.0
0.8
0.6
0.4
0.2
0.0
AUC = 0.90

(a)


Merge of CPBD and Laplacian Variance
Value
Threshold
TPR
FPR
TNR
PPV
ACC
F1-score

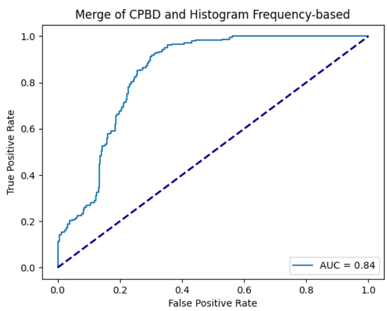
(b)


Merge of CPBD and Laplacian Variance
Density
Metric score output
Norm. approx. of Blurry
Norm. approx. of No Problems
Hist of Blurry
Hist of No problems

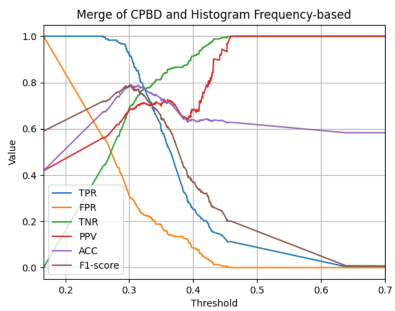
(c)

Figure 8

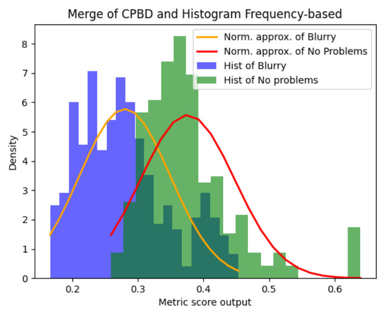
1.7 Merge of CPBD and Histogram frequency-based metric



(a)



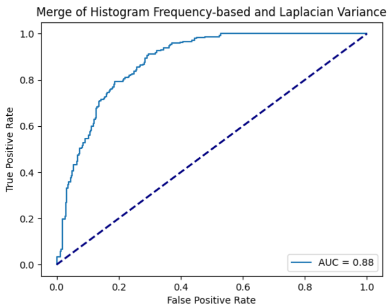
(b)



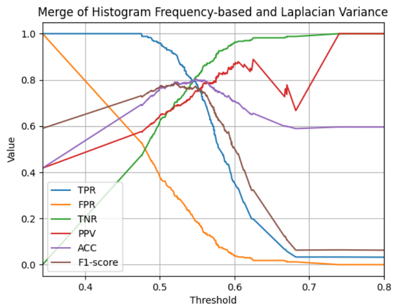
(c)

Figure 9

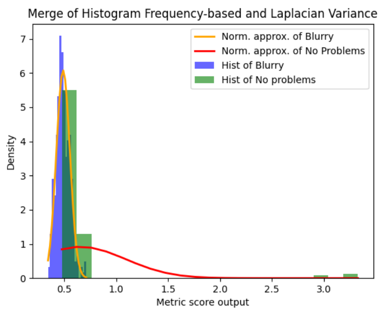
1.8 Merge of Histogram frequency-based metric and Variance of Laplacian



(a)



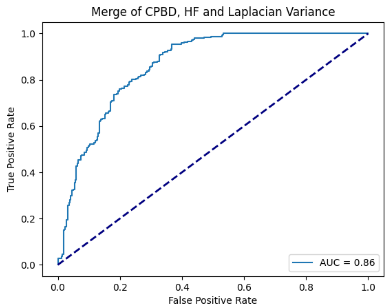
(b)



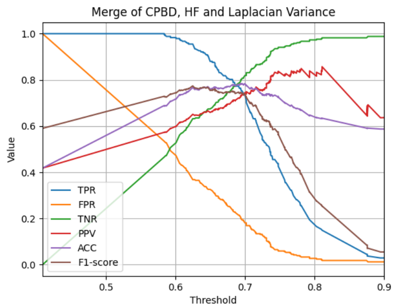
(c)

Figure 10

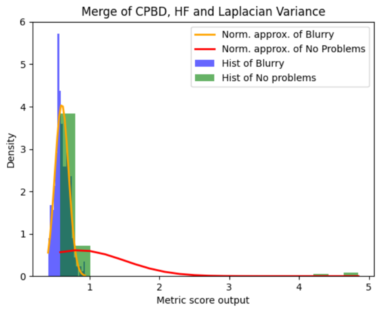
1.9 Merge of CPBD, Histogram frequency-based metric and Variance of Laplacian



(a)



(b)

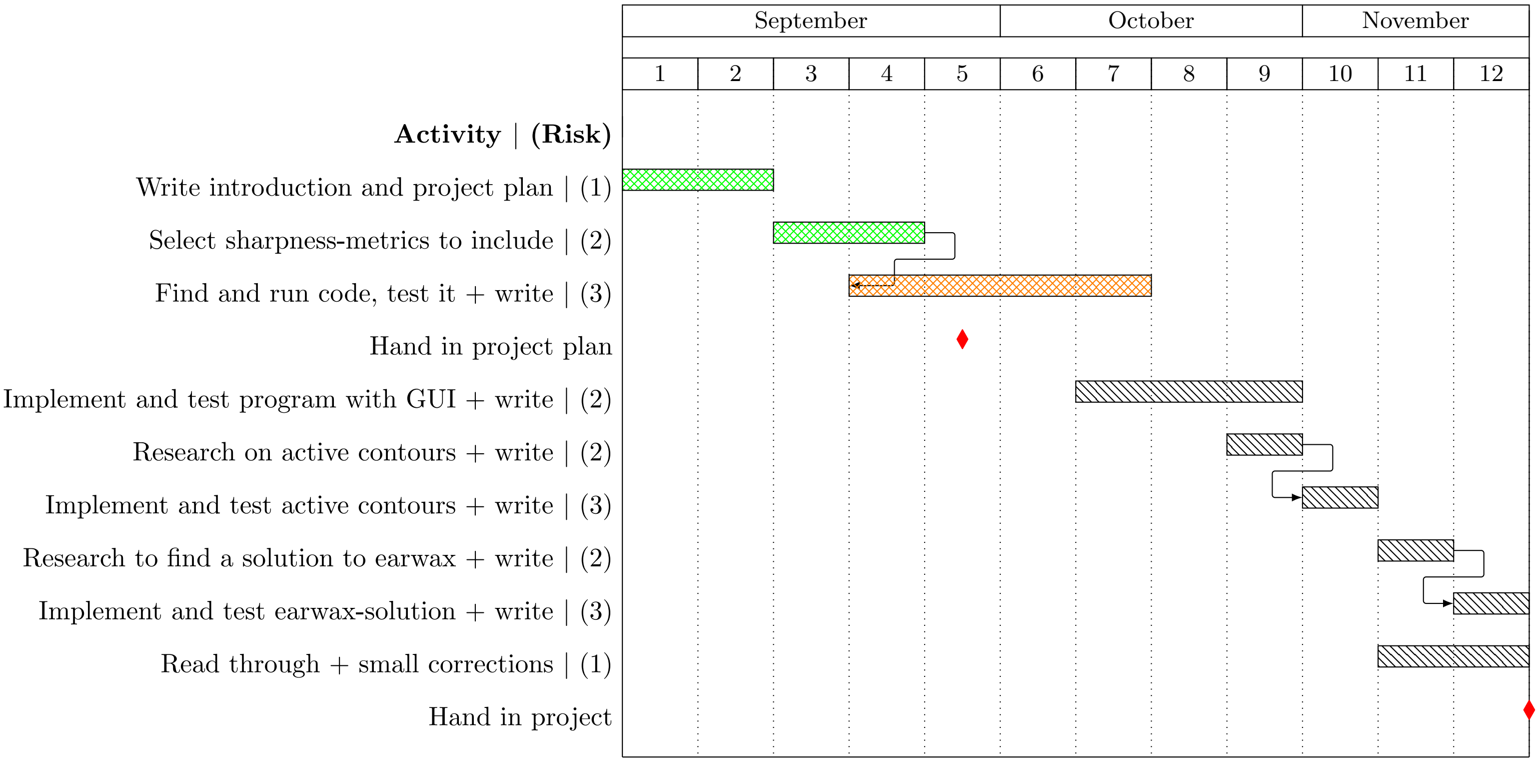


(c)

Figure 11

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