

# Exam in Computer Assisted Image Analysis I

Exam A, January, 2024

- Tools: This is an "open book" exam. You are allowed to bring and use the course literature (Any edition of Gonzales-Woods or Szeliski). Print-outs of the course literature are also allowed, as long as you have not made any additional notes on the print-outs.

**GOOD LUCK!**

/Filip

# Assignment

## Part 1

Imagine that you are tasked with implementing an optical character recognition (OCR) system, that takes a digital image of a scanned document and outputs a text file with the text contained in the scanned document.

You can assume that the scanned document contains only text, that the text is dark on a bright background, and that the text is written in a known font with fixed size. Furthermore, you can assume that the document has been scanned using a high quality document scanner, so that the illumination across the scanned document is even, and that the document is placed so that the orientation of the text lines is well aligned with the horizontal axis of the image.

The system should return a text file containing the text of the scanned document. The line breaks in the original document should be preserved, i.e., each text line in the original scanned document should be represented by a separate line in the resulting text document. The computation time of the text recognition system is not crucial, but your solution should be possible to implement in a reasonably efficient manner.

- Describe in detail how the image analysis pipeline of such a text recognition system might be designed, using methods covered in the course. What different steps would you propose to include, and what methods could be used at each step?
- Discuss how could the performance of such a system be evaluated? What performance metrics might be relevant?
- Discuss possible limitations of your proposed solution. What assumptions about the scanned images, besides those stated above, does your method rely on? When might those assumptions be violated, and how would your method perform in that case?

## Part 2

Imagine that after your proposed system has been implemented, the client wants to use the system on images of text captured with a another scanner of lower quality. In these images:

- the document may be rotated so that the text is no longer aligned with the horizontal axis of the image.

- the illumination across the image may be uneven.
- some of the documents may have handwritten notes written with a red ink pen, and the text recognition should not consider these notes.

How could your system be extended to handle these new conditions? Describe in detail how the image analysis pipeline from Part 1 could be modified to handle these additional challenges.

## Grading

Your answer will be graded as follows:

- *Grade 3:* You propose an image analysis pipeline for the task in Part 1, using relevant methods from the course. You provide reasonable arguments for why the proposed approach may work. The methods used in the pipeline are described correctly. You propose a relevant metric for evaluating the performance of the proposed approach. You discuss, at a basic level, the possible limitations of your proposed approach.
- *Grade 4:* (*In addition to the criteria for Grade 3*) You propose an extension of the image analysis pipeline to handle at least two of the three challenges introduced in Part 2, using relevant methods from the course. You provide a convincing motivation for your proposed solution, demonstrating a good understanding of the methods involved.
- *Grade 5:* (*In addition to the criteria for Grade 4*) You propose an extension of the image analysis pipeline to handle all three challenges introduced in Part 2, using relevant methods from the course. You provide a very convincing motivation for your proposed solution, and discuss the advantages and disadvantages of your solution compared to other possible approaches. You demonstrate a very good understanding of the methods involved. You discuss different ways in which the performance of the proposed system could be measured, and compare possible advantages/disadvantages of different performance metrics.

