Document Analysis
Exercises

a.o. Univ.-Prof. Dr. Robert Sablatnig

Schedule The exercise should be written in Python. Important dates are listed below:

- 19.04.2025 Deadline of assignment 1
- 28.04.2025 Presentation of assignment 1
- 23.05.2025 Deadline of assignment 2
- 02.06.2025 Presentation of assignment 2

Submission The assignments must be uploaded to the TUWEL course. Each assignment should be zipped (name.zip) and contain the subsequently enumerated items:

- **main.py** the main routine
- **assignmentX.pdf** the report

All files must contain the name of all group members. The report should explain the assignments and answer all questions. It should be a comprehensive but short (e.g. 6 pages) explanation of the algorithms used, their advantages, and drawbacks.

If you have any questions please write an email to: davu@cvl.tuwien.ac.at or post it in the TUWEL forum.

Assignment 1 – MSI Analysis and Binarization

Task A - Palimpsest

In the first exercise, you will analyze a synthetic palimpsest (`palimpsest-2025.jpg`, see TUWEL course). As task separate the 3 channels of the colour image into 3 individual images (R, G, and B). Thus you have a simple example of a

multispectral image. Your task is now to scale and combine the color channels to estimate the three classes of the object (Paper, Overwriting, and Underwriting).

Simple Method:

- Split color image in the R, G, B channels.
- Investigate the gray values of the 3 classes (e.g. sample, segment, ...) and select the 2 channels with the highest differences for the overwriting.
- You will need to find scale factor that makes parchment and overwriting subtract to same gray value (same contrast in both channels)
- Subtract to create a difference image with only underwriting text

In the report, document the steps you have done to separate the classes of your palimpsest image.

Task B - Binarization

As a second part, you must implement two binarization methods. First, the binarization method of Su et al. [2]. To show that your implementation is correct, use the DIBCO2009 dataset[1] and compare it with the results of the paper (as metric implement F-Score, pseudo-F-Score, P, R, and PSNR).

Furthermore, implement a U-Net-based [1] binarization[2]. You are free to choose the parameters yourself. As training data you can use datasets from other DIBCO competitions (except DIBCO2009, see also the link before). Compare the results of the U-Net to the binarization of Su et al.

Report the results of your binarization method on the DIBCO2009 dataset and show the comparison to U-Net. Describe the advantages and disadvantages of both methods and report sample images.

Literatur

[1] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015 - 18th International Conference Munich, Germany, October 5 - 9, 2015, Proceedings, Part III*, Lecture Notes in Computer Science, pages 234–241, 2015.

[1] `https://cloud.cvl.tuwien.ac.at/s/NPgtS6rnqYzxrFj`

[2] Example implementation for U-Net: `https://github.com/milesial/Pytorch-UNet`

[2] Bolan Su, Shijian Lu, and Chew Lim Tan. Binarization of historical document images using the local maximum and minimum. In *The Ninth IAPR International Workshop on Document Analysis Systems, DAS 2010, June 9-11, 2010, Boston, Massachusetts, USA*, pages 159–166, 2010.