Prepared by dr. D. Matuzevičius
2025 09 22

Laboratory Work No. **2**

Investigating Data Augmentation

The aim of this laboratory work is to explore and test methods of data augmentation (increasing the amount and diversity of data).

Main task [8 points]
When performing this LD, the following tasks must be completed:

1. Install the *albumentations* library:
https://albumentations.readthedocs.io/en/latest/index.html
https://github.com/albumentations-team/albumentations

2. Try out image transformation methods. Implement at least 5 different algorithms and test them with any images you choose. Choose from the list of transformations (https://albumentations.ai/docs/api_reference/full_reference/):

   a) from the group of intensity transformation algorithms (Pixel-level transforms);

   b) from the group of geometric transformation algorithms (Spatial-level transforms):
   – when only the image is transformed (Image),
   – when the initial image and mask are transformed (Image+Masks),
   – when the initial image and the framing rectangle are transformed (Image+BBoxes),
   – when the initial image is transformed together with significant points (Image+Keypoints).

3. Create a sequence of image augmentations that is suitable for training the convolutional neural network from the first homework assignment (*HW1: Semantic Segmentation of Medical Images*). Provide the results of this sequence of augmentations (the original image (input) together with the output image (target)).

Questions [2 points] [optional task]
1. What difficulties did you encounter while performing the task and how did you address/resolve them?
2. Which image augmentation methods are suitable and which are not when training an artificial neural network to segment blood vessels in retinal (fundus) images, and why?
3. How can you determine whether a particular image augmentation method is suitable for the task at hand?
4. Come up with your own question related to this LW, ask it to yourself, and answer it.

Contents of the Report

As a report, send the **script(s)** (*.py files or Jupyter Notebook file) prepared for the solution of this task and **additionally those files in PDF format**. In the scripts, provide explanations in the form of comments about what you did and why you did it in the most important steps you had to perform. Also, include the results of testing the algorithm/model you created to solve the task. If you submit a Jupyter Notebook file (and its PDF) and all the results of your task solution are clearly visible among the script execution results, it is not necessary to submit the test results separately. Also, include answers to the questions (answers to the questions are not a mandatory task).

Report Submission and Evaluation

This LW task should be completed and sent by email to dalius.matuzevicius@vilniustech.lt by 23:59 on 2025-10-07. Late submission of the report will result in a 25% reduction in the task.