Computer Vision (CV) (English) – JMU Wuerzburg, Master, Spring 2021

There are **4 home assignments** roughly corresponding to the teaching sessions. They are counting for a 10% bonus in the final grade for this course and are meant for familiarizing the students with the written exam which will contain similar tasks.
Each home assignment (HA) has 8 topics that are weighted equally. The solutions to each HA should be sent as a single *pdf* by email at radu.timofte@uni-wuerzburg.de no later than May 26 for HA1, June 10 for HA2, July 1 for HA3, and July 15 for HA4, respectively. Each student shall include with her/his email full name and the details of the study program. The Subject of the email should be of the form:
"CV 2021 HA[#home_assignment] [master_program] [lastname] [firstname]"
where [#home_assignment] should be replaced with the number of the home assignment, the [master_program] with the corresponding short abbreviated program (like XtAI, LuRI) and [lastname] and [firstname] are the student's last name and first name, respectively.
The grades on the home assignments will be provided together with the final course grade.
Plagiarism and misconduct are not tolerated!
The grading of Computer Vision (taught in English) is as follows:

- **The final CV (Eng) grade (maximum 100%)** is composed from 100% written exam.
- Up to **10% extra bonus** will be awarded for those attending the lectures.
- Up to **10% extra bonus** will be awarded for those attending the exercise sessions.
- Up to **10% extra bonus** will be awarded for those sending their solutions to the HAs.
- For **exam** the students are required to prepare the contents covered during the lectures, exercise sessions, home assignments and found on the provided slides and lab materials. Examples of potential exam topics are found in the HAs.

Home assignment III

(due July 1st, 2021)

17. What tracking method or combination of methods (region, point, template, tracking-by-detection, articulated model, on-line learned) would you use in each following application scenario? What limitations you may expect?

(i) App1. Medical: You wish to track the ultrasound probe, to relate images in 3D space.
(ii) App2. Autonomous driving: Tracking other nearby vehicles to adjust speed and course.
(iii) App3. Safety: In a lumber mill, you wish to use CV to stop the blade if a hand reaches nearby.
(iv). App4. Your favorite tracking application.
Justify your answers.

18. The task is to track a fish in a water tank. The camera is fixed, the background is fixed, the fish is moving. In the video the position of the fish is provided in the first frame for initialization of the tracker. The current tracker is a template-based tracker matching the initialization cropped image to any new frame in the video tracker.

(i) how would you handle the occlusions? (in the tank are also plants)
(ii) which information can be used to improve the tracker?
(iii) will the background modeling help? how?
(iv) will the modeling of the dynamics help? how?
Justify your answers.

19. What is the difference between linear and logistic regression?

20. What is the difference between classification and regression?

21. In a typical training scenario of a neural network with train, validation and test data, how do we detect under-fitting?

22. Why a MLP without activation functions is acting like a linear mapping?

23. In a MLP architecture for digit classification, the sequence of layers is as follows (no biases are added following the fully-connected layers):

```
input grayscale image with 16x16 pixels (total values: 16x16)
→ fully-connected layer with 128 output units → activation layer
→ fully-connected layer with 32 output units → activation layer
→ fully-connected layer with 16 output units → activation layer
→ fully-connected layer with 10 output units
```

What is the number of parameters of this specific MLP architecture?
(*for partial points you should provide intermediate number of parameters per each layer*)

24. In a MLP architecture for image classification, the sequence of layers is as follows (no biases are added following the fully-connected layers):

```
input RGB image with 64x64 pixels (total values: 64x64x3)
→ fully-connected layer with 128 output units → activation layer
→ fully-connected layer with 64 output units → activation layer
→ fully-connected layer with 20 output units
```

What is the number of parameters of this specific MLP architecture?
(*for partial points you should provide intermediate number of parameters per each layer*)