I. IDENTIFICATION DATA

Thesis name:	**Label propagation for one-shot video object segmentation**
Author's name:	**Davídek Hynek**
Type of thesis :	master
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Department of Computer Science
Thesis reviewer:	Siniša Šegvić
Reviewer's department:	UniZg-FER

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment	**challenging**

Evaluation of thesis difficulty of assignment.

The assignment is based on a recent research paper on one-shot segmentation. It proposes two groups of improvements over the baseline. Training improvements involve complementing cross-entropy with metric learning. Inference improvements involve a different label propagation strategy and several kinds of test-time ensembling.

Satisfaction of assignment	**fulfilled**

Assess that handed thesis meets assignment. Present points of assignment that fell short or were extended. Try to assess importance, impact or cause of each shortcoming.

The thesis meets the assignment. It reproduces the experiments from the literature and reports significant improvements. The greatest improvements come from probability propagation and increased input: 8.4 and 8.7 percentage points J&F (pp). Combination of these two techniques led to additional 0.6pp. The best metric learning approach contributed 4pp over the baseline. A combination of metric learning and increased input contributed 0.2pp. Unfortunately, combined approaches result in only slight improvement. This is not strange since the effort for achieving vision competence usually scales superlinearly with the target performance level .

Method of conception	**correct**

Assess that student has chosen correct approach or solution methods.

The chosen methodology is reasonable and suitable for a MS level. Saturation of combined approaches suggests that a radically different approach would be required to further improve the accuracy.

Technical level	**B - very good.**

Assess level of thesis specialty, use of knowledge gained by study and by expert literature, use of sources and data gained by experience.

Experiments are based on the public github repository of the original research paper. The performed interventions include:
- introducing triplet loss to the training script
- implementing and validating several triplet miners
- replacing one-hot label propagation with probability propagation

- inference on jittered images
- introducing a different backbone

The demonstrated technical level is suitable for a MS thesis.

Formal and language level, scope of thesis

B – very good.

Assess correctness of usage of formal notation. Assess typographical and language arrangement of thesis.

The thesis is well written with very few typographical errors. Written English is excellent.

Selection of sources, citation correctness

A – excellent.

Present your opinion to student's activity when obtaining and using study materials for thesis creation. Characterize selection of sources. Assess that student used all relevant sources. Verify that all used elements are correctly distinguished from own results and thoughts. Assess that citation ethics has not been breached and that all bibliographic citations are complete and in accordance with citation convention and standards.

The selection of cited sources is adequate. Suitable additional citation:

Alexander Hermans, Lucas Beyer, Bastian Leibe. In Defense of the Triplet Loss for Person Re-Identification. CoRR abs/1703.07737 (2017)

Additional commentary and evaluation

Present your opinion to achieved primary goals of thesis, e.g. level of theoretical results, level and functionality of technical or software conception, publication performance, experimental dexterity etc.

It would be helpful to describe what exactly happens when we skip the spatial term from equation (4.3).

It would be helpful to provide a breakdown on GPU RAM usage on page 30. Which tensors use most memory? Why would sparse matrices help? Could we expect improvement due to sparse matrices when dealing with probability propagation?

It would be helpful to indicate whether the dilation factor is increased after reducing the pooling stride from 2 to 1.

On page 41, the text suggests that the optimization uses SGD without momentum. However, the original implementation uses Nesterov acceleration with momentum 0.9 (according to the public github source). It would be helpful to clarify the reason for changing the optimizer or to correct the text.

The text does not explain the temperature hyper-parameter which has been validated in Table 5.1.

It would make sense to show a flip + rescale experiment in Table 5.4. Multi-scale inference would also make sense. It would be helpful to explain the poor result of backbone combination with probability propagation.

Copy-paste error in the caption of Table 5.7 should be corrected.

III. OVERALL EVALUATION, QUESTIONS FOR DEFENSE, CLASSIFICATION SUGGESTION

Summarize thesis aspects that swayed your final evaluation. Please present apt questions which student should answer during defense.

The thesis addresses a relevant problem and presents significant improvements with respect to a baseline published on a major interational conference.

Some implementation details could have been better described, while the proposed improvements have limited novelty.

The strengths of the thesis outweight the shortcomings.

I evaluate handed thesis with classification grade B - very good.

Questions:
1. Imagine two squares of the same size, G and M. The center of G is located at the vertex of M. What are the region similarity J(G,M) and the contour accuracy F(G,M)?
2. Describe optimization with Nesterov accelerated gradient.

Date: **14.6.2021**   Signature: