Skoltech
Skolkovo Institute of Science and Technology

Jury Member Report – Doctor of Philosophy thesis.

Name of Candidate: Egor Zakharov

PhD Program: Computational and Data Science and Engineering

Title of Thesis: Synthesis of human face and body images via generative adversarial networks

Supervisor: Associate Professor Victor Lempitsky

Name of the Reviewer: Tat-Jen Cham

I confirm the absence of any conflict of interest	**Date: 12-03-2023**

The purpose of this report is to obtain an independent review from the members of PhD defense Jury before the thesis defense. The members of PhD defense Jury are asked to submit signed copy of the report at least 30 days prior the thesis defense. The Reviewers are asked to bring a copy of the completed report to the thesis defense and to discuss the contents of each report with each other before the thesis defense. *If the reviewers have any queries about the thesis which they wish to raise in advance, please contact the Chair of the Jury.*
Reviewer's Report

This thesis presents an extensive body of research work, with five chapters covering original research on the topic of face / head synthesis in images and video, and one chapter on full-body synthesis. The quality of the research is high in terms of novelty (in relation to the times when the corresponding research papers were published), logical and well-justified architectural design decisions, extensive experimentation, as well as clearly written presentation of the works. All of these works have been validated by the international computer vision research community by having been published in top-tier venues, such as CVPR, ICCV and ECCV. Many of the works have influenced other researchers in terms of ideas and future directions, and in particular the work in Chapter 4 has been very well-cited. Compared to PhD theses published in my own university, which typically will have three to four chapters of original research, this thesis is an outlier in terms of the extensive research ground it covers, and will rate at a very high level.

Sequentially, the chapters provide a nice retrospective sense of the evolution of research ideas and capabilities since 2018, both in terms of the candidate's research, but which also reflects that in the research community as a whole. Because of the rapidity of this evolution, where ideas can quickly fall out of fashion, the chapters do appear quite siloed from each other. For example, the perceptual discriminator of Chapter 2 was not used in later chapters (or justified against its use), and likewise the low-high frequency decomposition and LGD methods from Chapter 5 in subsequent chapters.

Here are some questions and minor revisions to consider for improvement:

- In relation to my previous point, it will be good to at least have short explanations in the thesis as to why potentially-relevant methods in earlier chapters were not used in later ones, to provide a more cohesive flow to a thesis.
- Again, for improving the flow of a thesis, when citing the candidate's own work from a previous chapter, avoid citing the original paper but refer to the chapter directly.
- In the references, Zakharov 2019b appears to be the same as Zakharov 2019a?
- The chapters, in multiple instances, refer to "supplementary material", which is a holdover from the original conference publication, but does not apply to the thesis. Now these materials are actually important, because they provide additional details and other results that did not make it into the chapter proper (presumably not because of page limits, but because they would otherwise disrupt the presentation of the chapter). There are a few ways to handle this: (1) have appendices in the thesis, (2) refer to corresponding arxiv versions, or (3) refer to project websites (e.g. on GitHub). In relation to this, since video results are a strong component of the research, there should be some way to allow readers of the thesis to access such videos.
- In Chapter 5, what is the value of K, the number of embeddings? What is the equation for L_ {Total}^G, in Figure 5.3?
- For Chapters 4 and 5, why were landmarks converted into line-segment landmark images? Were there consideration for other types of representing landmark information, including explicit landmark positions, or landmark heatmaps?

Overall, this thesis is very well-written and reading it has been a pleasure, and I would like to compliment the candidate for the important research contributions covered in the thesis.

Provisional Recommendation

☒ *I recommend that the candidate should defend the thesis by means of a formal thesis defense*
☐ *I recommend that the candidate should defend the thesis by means of a formal thesis defense only after appropriate changes would be introduced in candidate's thesis according to the recommendations of the present report*
☐ *The thesis is not acceptable and I recommend that the candidate be exempt from the formal thesis defense*