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: Dmitry Dylov

I confirm the absence of any conflict of interest (Alternatively, Reviewer can formulate a possible conflict)	**Date: 12-03-2023**
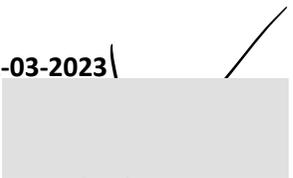
The purpose of this report is to obtain an independent review from the members of PhD defense Jury before the thesis defense. The members of PhD defense Jury are asked to submit signed copy of the report at least 30 days prior the thesis defense. The Reviewers are asked to bring a copy of the completed 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

The work is of exceptional quality and it promises to become one of the best Ph.D. theses ever defended at Skoltech. The author dedicated his effort to realistic image synthesis, with generative modeling, specifically with the generative adversarial networks (GANs), being the leitmotif of the research.

More specifically, the author studies how to influence the realism of a synthesized image of a person when only a single or several input images are given (few-shot synthesis). The author incorporated 3D constraints into the geometry-free models, pioneered to develop textured neural avatars, and proposed the one-shot method to create large high-resolution portraits using learnable meshes ('MegaPortraits').

Besides the unquestionable scientific value, the author also considers practical impact in a form of studying the challenging aspects of real-time initialization and inference. Please evaluate the text quality and the structure of the thesis: The technical rigor and the general story line in the thesis are drafted well, the reading is interesting and engaging, the conclusions and the limitations of each approach are given in each section, which creates a simple and immersed experience for a reader. The structure of the thesis text is satisfactory with only minor improvements suggested below.

Technical questions: In Section 2, besides CelebA, the competitors also consider 300-W and MAFL datasets. A comparison against those benchmarks could be useful. In perceptual metrics, Vid-to-Vid (few shot) performs similarly to proposed model in Fig. 21. What are the error bars (stDevs) and the reasons why LPIPS is not drastically better gain than the CSIM? Eq. (32): what are the reasons for assuming that gaze and face losses are separable? It is natural to assume a cross-term for these two attributes. The ultimate loss Eq. (36) looks a bit arbitrary, and it seems that only the cosine loss term was investigated w.r.t. the ablation effect. If the weights of the other terms, e.g. cycle consistency $w_{FM?}$, were also varied, the hyperparameter search should be elaborated.

Recommended improvements (all optional): 1 What is a unifying theme that connects all chapters in the thesis? The structure of the thesis could reflect this unifying theme in a form of a joint problem statement, from which all other sub-problems stem. One could even draw a visual map of connection between different chapters in the thesis to show how these connections help solve the major problem formulated in the beginning. Otherwise, the jump to Perceptual Discriminators is very sudden. The beginning of the thesis is a bit dry. It would be instrumental for the reader to see some examples of the images and avatars in the introduction. To have a strong finish, same applies to the Conclusions: what are the outcomes – it is an ideal place to show the stunning avatar. Consider renaming the subsection "Overview" (+ of…) Consider being more consistent with the structure of sub-sections. Sometimes Methods are placed as a separate section, and in some chapters, they are moved to Experiments. Same applies to dataset descriptions in different chapters.

Consider re-factoring references so that they start not from Ref. [147]. Consider using AMC [Last Name, Year] referencing format, as it would be more appropriate for a monography-style text to recognize contributions in the community. "For more details, please refer to the supplementary materials. " – this is, perhaps, an artefact from the article's text. "in submission" -> "under review" Please evaluate if the publications and conference presentations meet the PhD Program requirements: Mr. Zakharov exceeds the Ph.D. requirement for the publications by a large margin: by the time of defense, 6 papers will be published in the top-tier conferences (four A* publications at the time of writing this review with two other manuscripts nearly accepted). The author clearly states his contribution to these works, serving an absolutely crucial role for the acceptance of these papers (both technical and concept-devising contributions). The claims for the final defense of his Ph.D. thesis are justified and the defense can readily occur by the publication record.

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*