![Skoltech logo](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: Prof. Dr. Hao Li

I confirm the absence of any conflict of interest (Alternatively, Reviewer can formulate a possible conflict)	**Date: 13-03-2023**

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
Reviewers report should contain the following items: • Brief evaluation of the thesis quality and overall structure of the dissertation. • The relevance of the topic of dissertation work to its actual content • The relevance of the methods used in the dissertation • The scientific significance of the results obtained and their compliance with the international level and current state of the art • The relevance of the obtained results to applications (if applicable) • The quality of publications The summary of issues to be addressed before/during the thesis defense

A/ This Thesis presents a number of important advancements in GAN-based avatar synthesis techniques, that have already made an impact in the field. The thesis not only suggests the use of GANs to improve realism, easy of generating avatars, but presents new ways to address some of the challenges known for GAN-based synthesis, namely: the ability to control data more reliably (face), to synthesize more general content (bodies), to personalize trained models (using few shot techniques), to enable real-time performance on mobile devices, the ability to achieve high-resolution results, and hybrid CG and neural approaches for rendering. In each of these instances, the thesis has proposed a novelty that was published at one of the top conferences in computer vision (ECCV, CVPR, etc.).

B/ The Structure of the thesis is basically a concatenation of each paper, with an introduction of and a conclusion. The titles of each section directly reflect the titles of each paper, which may not be the most accessible for a non-expert in the field, or someone who knows all the related works. While this is acceptable and I have seen many thesis at top universities that are structured in this way, I would suggest the following structure to further highlight the importance of each findings and the proposed methods:

1/ Introduction
  * Motivation
    * Applications
    * CG vs Generative Models
  * Background on GANs
  * Challenges
    * Talk about Control
    * Talk about Performance
    * Talk about Resolution
    * Talk about Generalization (to Bodies, to Hair?, to Clothing?)
  * Thesis
    * Overview of Proposed Improvements

1/ Synthesizing and Controlling Faces
  * Motivation
  * Basic Techniques for Synthesizing/Manipulating Faces
  * Related Works for Synthesizing/Manipulating Faces
  * Proposed Approach with Perceptual Discriminators
  * Discussion
2/ Synthesizing and Animating Bodies
  * Motivation
    * going from faces to bodies
  * Basic Techniques for Synthesizing/Animating Bodies
  * Related Works for Synthesizing/Manipulating Faces
  * Proposed System of Textured Neural Avatar System
  * Discussion
3/ Personalizing Faces
  * Motivation
  * Basic learning approachs
    * few shot vs large training data
  * Related Works
  * Proposed Few Shot Approach with Applications in Neural Talking Head Models
  * Discussion
4/ Real-time Face Synthesis
  * Motivation

* Basic speed improvement approaches
* Related Works
* Proposed Bi-Layer Neural Synthesis Approach
* Discussion

5/ High-resolution Face Synthesis
* motivation
* Basic up-scaling techniques
* Related Works
* Proposed MegaPortrait System
* DIscussion

6/ Hybrid CG and Neural Head Rendering
* motivation
* Basic Idea for Hybrid CG and Neural Rendering
* Related Works
* Proposed One-Shot Mesh-Based Head Avatar Method
* Discussiong

7/ Conclusion
* Findings Summary and Broader Impact
* Discussion of Advantages and Disadvantages of GANS
* Limitations of Proposed Approaches
* Future Directions

Furthermore, the introduction and conclusion can benefit more from an in-depth analysis on the findings, advantages and disadvantages of GANs, limitations of the current approaches, and more discussions on how future directions could look like. How do these things compare with other generative methods, like stable diffusion? Potentially discuss briefly ethical considerations and how these technologies could be misused as they allow us to produce highly convincing media manipulations.

C/ The topic of the Dissertation is highly relevant to it s actual content, and I have no concerns about that, each chapter addresses a key challenge in GAN-based synthesis with a novel system, approach, methodology.

D/ The methods of the topic used in the dissertations are not only highly relevant, novel, and have already made an impact in both industry and academia, but they are simply impressive and very sophisticated solutions. As most publications are still quite recent, most of the techniques are still state of the art, and published at some of the top conferences in Computer Vision.

E/ The results from this dissertation demonstrate that GAN-based avatar synthesis are possibly the key for highly realistic telepresence technologies as well as visual effects system, and many of the proposed techniques can already impact consumer applications such as advanced filters in social media platforms (Snapchat, Tiktok etc.). The research published as part of this dissertations are at the highest international standards, and top conferences in computer vision.

Summary:

The way the Thesis is presented is absolutely acceptable as it is, but since the content and approaches are so good, I would recommend to spend a week and make another pass at improving the structure based on my suggestions above:

1/ adopt the title if the author likes it.

2/ adjust the titles of each section and adopt the structures of the subsections.

3/ add a bit of background material for each of the topics

4/ write a little bit more in the introduction (applications, GAN background, what are the challenges, etc.): 3 more pages

5/ add a little bit more into the conclusion, in terms of analysis, discussion, limitations, and future work 4 more pages.

6/ small typos here and there, use something like grammarly to fix them, also there are some errors in the bibtex (see citation of PaGAN)

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*