Skoltech
Skolkovo Institute of Science and Technology

Thesis Changes Log

Name of Candidate: Aleksei Shcherbak

PhD Program: Computational and Data Science and Engineering

Title of Thesis: Diagnosis of the human neurological movement disorders based on the analysis of heterogeneous data using ml methods

Supervisor: Associate Professor Andrey Somov

The thesis document includes the following changes in answer to the external review process.

Improvements:

1. Conclusions were rewritten
2. Some of typos and grammar mistakes were fixed
3. Two figures were added to chapter 3.3 (Wait-n-Sync time diagram and experimental testbed)
4. Figure was added to chapter 4.1.3 (Handwriting analysis workflow)
5. Some of images with low resolution were replaced with images with higher resolution (data processing pipelines in chapter 4 and chapter 5
6. The row with an average values was added to tables 6.1, 6.2 and 7.1-7.4
7. Information about the state-of-the-art gaps was added to the end of chapter 2
8. Scheme of data collection procedure and camera locations was added to chapter 3
9. Equations in chapter 3.3.2 were relocated
10. Figures 3.2 and 3.3 were described in text in chapter 3
11. Abstract was edited

Q&A section for reviewer's comments:

Maxim Sharaev:

1. **Q:** In Intro section there are two statements "Symptoms overlapping with other diseases especially ET" and "Diagnosis can be confirmed after an autopsy only". Moreover, next "State-of-the-art diagnostic practices in fact have probabilistic character and cannot guarantee success." How this relates to labelling by clinicians, i.e., if it is hard to solve diagnostic task manually - how you can be sure that your labelling is correct?

   A: Indeed, the data labelling is not 100% correct, and mistakes are possible and acceptable here because of specificity of this field of study

2. **Q:** In Section 1.2 again "there were many attempts to use additional instruments to diagnose PD such as positron emission tomography (PET), computed tomography (CT) with a radioactive tracer (SPECT), magnetic resonance imaging (MRI), electroencephalography (EEG) and others." – in these attempts the aim is to differentially diagnose the reason, i.e. true Parkinson disease. So, how exactly this is connected to symptoms?

   A: This focus here was on diversity of ML applications in PD field, not on application of ML methods to specific symptoms

3. **Q:** Section 1.2 “Additionally, proposed in this work ML-based approaches potentially can register not visible tremor and promotes early PD diagnostic,” How exactly this works? Should be justified (invisible tremor diagnostics)

   A: I replaced “ not visible” with “unnoticeable” to make it more clear. The thought here is that the tremor amplitude can be too small to be noticeable (visible) even for patient, which is possible at the earliest state.

4. **Q:** Section 3.1 “the collected data and medical conclusion from neurologists, who examined the patients before experiments” Info about clinical objectiveness and concordance between neurologists should be provided.

   A: The data was collected in clinics during standard intensive admission days, thus no special procedure to keep concordance was conducted, except for cases, when diagnosis was unclear and neurologists held discussion.

5. **Q:** “In such situations, clearly defined clinical diagnoses came later and expressed a group opinion of several neurologists (3-5 persons), relying on additional comprehensive examinations.” Here at least a bit of discussion will be valuable

   A: Thank you for this comment. I agree with it.

6. **Q:** Section 4.1.1 “ For bradykinesia, the related features were extracted within a range from 0 to 3 Hz.” Before it was stated that high-pass filter was used from 0.5Hz

   A: Yes, but it is not necessary here to keep boundaries of the filters. 0 Hz is not a strict bottom boundary of bradykinesia spectrum span, it is more a trend and justified simplification.

7. **Q:** “For this, the mean and standard deviation (STD) of all the amplitudes and frequencies of the full range of spectrum (mean and STD of amplitudes)”. Is there any clinical interpretation for other than frequency features? please elaborate more on clinical meaning?

   A: In data-driven approach, not all features have exact clinical interpretations, but presence of clinical meaningful is very beneficial.

8. **Q:** Section 4.1.3. “Image analysis is the most complicated part here.” Why? For me it is not clear. Why not to compare versus standard image classification methods (DL without feature extraction?) - I removed this subjective part and a little bit rewrote this part.

   A: I removed this subjective comment from the thesis. Indeed, it is not so challenging.

9. **Q:** Section 4.2.2 “For video data collection and analysis, all exercises except for 3, 6, 10, 11 and 15 were used.” Why? Some statements lack explanation.

   A: Within my PhD study we carried out research on 'efficiency' of proposed exercises, i.e. exercise which best provoke the PD symptoms. That is why some of 'inefficient' exercises have been excluded from data analysis

10. **Q:** Section 4.2.3 “..are not applicable in practice due to the poor result on the Parkinson Drawing Dataset.” Why? Could you show the figures with examples of misclassification? For me it seems strange, the difference between HC and PD writing looks huge (fig 4.5).

    A: Cannot provide you with examples of misclassification, because it was several years ago and there is no special place, where I saved specifically these pictures. However, I can notice here that, in my opinion, the problem with data reproducibility is common for this field of study (ML-based PD diagnostic). I highlighted this moment in conclusions.

11. **Q:** Section 7.1.3 For future tests I’d suggest intermediate fusion (not early or late) with smth like autoencoders

    A: I will pass your idea to my PhD colleague at Skoltech, who continues this topic. Thank you.

12. **Q:** Conclusions. “Clinical checkup time matters in healthcare and it can be further reduced by relying on them.” I think here comparison with traditional examination will be valuable: can this methods find early stages when clinician fails?

    A: Potentially, yes, it can. However, I did not test it.

13. **Q:** Minor: Several typos, i.e. Section 2.1 “gate” – gait, “data dollection”, “all acientific” and etc. Figures quality could be improved (see fig. 4.4 – low resolution)

    A: Thank you for this comment. Fixed.

Ahmed Bouridane

14. The methodology of each contribution is not well articulated to highlight/describe the novelty in a way that can capture the significance and scientific rigor. Rather the candidate dives directly into the tools and methods used in the experimental results.
15. **Q:** Another concern is that the findings are not compared to some (2 or 3) existing and similar works (if any). Such an analysis would enhance the quality of the thesis and the contributions made.

    A: Results comparison is not so meaningful in this field of study, I explained this moment in conclusions
16. **Q:** Some chapters can be combined. Example Chapter 7 can be combined with Chapter 5 and Chapter 6 to make it more informative and larger one.

    A: Chapters were organized according to logical division of the materials, and I would prefer to keep them as they are, although your idea is clear for me
17. **Q:** Chapter 2: The style should be re-visited in my opinion. The narrative is long and is diluted. I suggest that some images/diagrams/Tables are added to show the evolution of the works for each type of works. Normally, for each investigation research gaps should be identified and highlighted to link them to the current investigation.

    A: I added the research gaps to this section, indeed, it is very important. However, adding some images is problematic here. At the end of this chapter, comparison table exists which represents diversity of relevant studies.
18. **Q:** Chapter 3: No title is given. No motivations/objectives? Methods section is too early.

    A: I fixed this imperfection, here and in other chapters.
19. **Q:** The formulation on Page 49 is too simplistic and needs attention. Define all parameters and clearly relate them to what is needed. Can you prove this against your assumptions?

    A: For me, definitions and limitations of parameters, described in this section, are clear. However, I added diagram, which shows relations between parameters.
20. **Q:** The results look like what you have obtained and lack a detailed analysis. A comparative analysis using various parameters would enhance the quality of the findings.
    A: In the results, I proved that the proposed method is working and compared its performance in different situations (parameters). In theoretical section, I identified which parameters and what values of them limit the performance of my method. Of course, the comparison can be extended more and more, but it is not the main topic of my thesis.
21. **Q:** A discussion of the results will be useful. Are there issues when the results are not actually good? When will these happen? When won’t your method work?
    A: Partially, answers for these questions are in theoretical section.
22. **Q:** A mathematical modelling is necessary. What are you capturing (feature extraction) and analyzing (feature selection)? What are you looking for? What are the metrics used? Are you combining (fusing) the results together for a better outcome? If not why not?
    A: In this study, a data-driven approach is used to diagnose Parkinson’s disease via methods of machine learning. Thus, no mathematical modelling is required. Answers for others questions are in the thesis.
23. **Q:** Again, an analysis of the results is necessary? How confident are you with the results? Are they repetitive?
    A: I am confident, and the results are repetitive, but only for my dataset. I highlighted the problem of reproducibility in conclusions of the thesis.

24. **Q:** Conclusion Chapter: it is too short. It does not clearly showcase the findings obtained. This chapter should be re-visited by clearly discussing the contributions made. A short section of future work would be a good idea, too.
    A: Absolutely agree. Conclusions were revised, everything was added.

Roberto Passerone

25. **Q:** Candidate is encouraged to make parallels and comparison to the existing work when presenting the proposed techniques. This could help better pinpoint the novelties and the differences, as well as advantages in terms of accuracy and the other metrics that have been used to evaluate the approach. If desired, these considerations could be collected in a discussion section that could be part of the conclusions.
    A: Thank you for this comment. I agree with it.
26. **Q:** The conclusions, too, could be extended. In particular, the candidate could summarize more precisely the limitations of the techniques adopted in this work, and, more importantly, give a more extensive account of the potential future work that could help overcome them. This would be of help to other researchers, to provide new ideas and perspective.
    A: Agree, conclusions were rewritten.
27. **Q:** Another suggestion is to add a little bit of context at the beginning of each section, before jumping into the technical details. For instance (but this is true at other points, too), Section 4.1 is empty and is followed immediately by Section 4.1.1. A paragraph introducing what the section is about could help raise the correct expectations.
    A: A context was added to all such places.
28. **Q:** Section 3.2: - Perhaps it would be easier for the reader if the equations (3.1) - (3.7) were placed in the text right after they are mentioned, so that one doesn't have to move continuously in the text to look at the equations.
    A: Thank you for this comment. Fixed
29. **Q:** Section 4.1.3: what is the "exam template"? Do you mean the ground truth? If so, why don't you call it ground truth?
    A: Yes, in the text I explained that the exam template is the ground truth.

Nikolay Brilliantov

30. **Q:** Could you please provide more details about the operation of the proposed time synchronization method? Does it use timing or sequence diagram? Is it already published or planned to be published elsewhere?
    A: I added timing diagram to explain better the operation of the proposed method. It is planned to be published.
31. **Q:** Typically, it is a good practice to discuss the obtained results in a comparison with the respective results of the relevant research work in the field. Although the state-of-the-art analysis has been performed, the respective discussion has not been provided. Could you please clarify this point – do similar research papers exists, and how your results differ from these of the other scientists?
    A: Yes, similar researches exist, but there is a problem of data reproducibility, which makes results comparison not meaningful enough. I discussed this moment in conclusions.
32. **Q:** It would be great to add a picture demonstrating schematically the data collection procedure and cameras position. Besides, there is no information about the functioning of smartphone-based application for the data collection.
    A: Thank you for this comment, appropriate image was added.
33. **Q:** In Chapter 7, in the results for joint sensors-video approach, the acronym “CI” was not explained.
    A: The acronym was explained in another section. I added it to the list of acronyms.
34. **Q:** The Conclusions section should be more clear and structured. In current conclusions, the focus jumps from sensors to video and back too frequently, therefore the text is difficult to perceive. Besides, although the analysis of the spiral-based method was performed as well, no attention was paid to this method in the conclusions. Similar criticism refers to the proposed time synchronization method.

A: Conclusions were rewritten and these moments were addressed.