Meeting - Week 2

Progress, Challenges & Next steps

Simon Winther & Hjalte Bjoernstrup,
Department of Computer Science,
University of Copenhagen (DIKU),
19–02–2025.

Questions? (1/5)

1. Is it permitted to stream while coding on platforms like Twitch or YouTube, not for providing help, but simply because it's fun and motivating to stream while working on a project?

2. Are we allowed to install missing PIP libraries, such as Hydra and other modules, if they are not found in the local environment, or should we ask for permission before doing so?

3. Are we allowed to use Docker? We attempted to use it, but it required the sudo command, which we are not allowed to use. Would it be possible to be added as Docker users, or is there another way to run it without sudo? Also, do we need to use Docker for the UNET 3+, and if so, does it impact performance (e.g., making it slower), or is it fine?
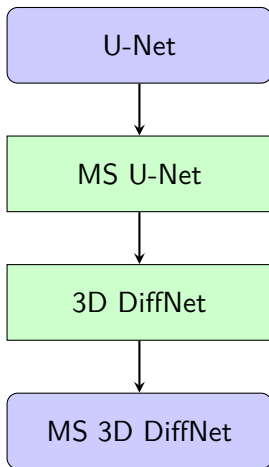
Questions? (2/5)

4. Are we allowed to use multiple GPUs for training? Specifically, if we notice that two GPUs are currently inactive or idle, can we use them for training the model? Or would you prefer that we always use only a single GPU?

5. We noticed that many papers had excellent modeling and visualizations of CNN and U-Net architectures. We were wondering which tools you and other authors used to create these high-quality figures so that we could generate similar ones ourselves.

Model Progression

Model Hierarchy

To the right, you can see a figure illustrating our proposed model progression, to ensure that it aligns with what we talked about in the last meeting.

- U-Net
- MS U-Net
- 3D DiffNet
- MS 3D DiffNet

Figure: Model progression from U-Net to MS 3D DiffNet

Questions? (3/5)

6. So, we wanted to ask whether the proposed model progression aligns with what we discussed in our last meeting as mentioned already.

7. After successfully training the U-Net 3+ model, should we evaluate that one first—storing its performance, using validation/testing data, and checking metrics like score or log-likelihood—before proceeding to train other models, or should we take a different approach?

8. Would it be possible to get an overview of the project timeline for the next few weeks? This would help us structure our focus areas. If this isn't something that can be provided, that's completely fine. We were simply just curious.

9. We find that papers formatted in two columns look very professional, and we wanted to ask if it would be acceptable to use this format for our paper, or if it would be a bad idea and something we should avoid.

Questions? (4/5)

11. When do you think we should start writing the paper, and what key aspects should we include? Would it be relevant to present the architectures of the multi-scale U-Net 3+ and the multi-scale 3D Diffusion Network, similar to how other papers illustrate model architectures? Should we also cover the mathematical foundations of diffusion models, or what would you recommend we focus on?

12. Since we haven't been able to run the training yet, we're unsure if there's already a built-in tool for visualizing and storing training progress—something like Weights & Biases, where you can track loss and other metrics. If such a system isn't included, would it be okay for us to integrate a library like Weights & Biases and add logging to track the training process?

Questions? (5/5)

14. Just to confirm, we only need to modify the code for three models. First, we clone the U-Net 3+ repository, train it as is, and store its performance in a table—this serves as the baseline U-Net 3+ model. At this point, no code changes are made; we simply run the model and record its accuracy. Next, we modify U-Net 3+ by adding multi-scaling, train it, and store the performance results to compare it with the baseline. Then, we clone the DiffNet repository, modify it from 2D to 3D, train it, and store its performance—this serves as the 3D DiffNet baseline. Finally, we take the 3D DiffNet and add multi-scaling, train it, and record its performance as well. In total, we will have four models in our comparison, but only three require actual code modifications: adding multi-scaling to U-Net 3+, converting DiffNet to 3D, and adding multi-scaling to the 3D DiffNet. Have we interpreted this correctly, or is there something we misunderstood?