

Feedback intermediate

- illustrative images in section 2 - make them smaller. Now you waste a lot of space and they are not so informative, to need so much space. For example, images 1.1. and 1.2 could be made smaller and placed next to each other. Same thing for 1.3, 1.5, 1.6 etc, these are huge given how much information content they carry.
 - careful about text formatting, e.g. page 5, or numbered list in page 9. What environment you use for this? Check some that would not leave so much empty space around. The whole visual feeling plays a role. It should look nice as a pleasant book to read.
 - page 11 - this is misunderstanding! These look like technical questions that you have asked yourself throughout the thesis and (at least partly) answered. You shall explain in the text the solutions for these but not present them as questions. Research questions are more general broader questions that your thesis aims to answer. Basically they define the objectives and scope of the thesis. There should be 1-3 (well max 5) of those. You can formulate them as questions but if difficult you may formulate them as aims or objectives.
 - section 5 images - again far too big
 - general for images and tables. The captions shall be much more informative. Basically, I shall understand what the image or table shows just from reading the caption without needing to read the text. I know this is tricky cause you cannot put everything so you need to find a compromise. For example for image 5.1 it could read something like "Figure 5.1 Examples of bright-field microscopy images with valid and invalid (unexpectedly elongated) morphology of tumor models.
 - Figure 5.5 - 16 vs 8 bits - yes, for calculation it matters but as soon as you display them, you probably always display only 8bit representation. Or not? What I mean to say is that for a human we cannot see any difference between the 16bit and 8bit by visually expecting them. Or shall we?
- Page 21 - a bit random comment - you spend a lot of time talking about 16 vs 8 bit and random jitter and 3000 epochs. But is not clear at all at this point why are you even jittering and what jittering even is and why you need 3000 epochs for it. Think about balancing your thesis. You shall spend a lot of time talking about the important critical things that really matter for the thesis - the objectives, the problem, the results. Technical details that you had to learn about and solve are interesting but do not sell them as the main results of the thesis. When too detailed and too technical, move it to Appendix. Think about textbooks - e.g. a book on machine learning would have a chapter on the mathematical pre-requisites typically not as part of the main text but only in the appendix.
- Section 6 - it seems to be very focused on the simCLR features. I have thought that you tried similar approaches (e.g. the clustering) over some other features, e.g. directly the original images or the ResNet features. Is it not the case? Do you have a baseline model to compare to?
- pages 34-35 aren't these repetitions from pages 22-23?
 - figure 7.7 why is anchor the last image and not the first in the left?

- Table 8.1.1 - are you saying that you classification accuracy is 100%? That seems far too good to be true. Why do the train runs have different number of epochs? And what do you mean by "Test Epoch"?!
 - general comment to tables with result - after giving the tables you need to write a text which helps us to interpret it. What are the most important number we shall look at? What shall we take out from the table as a message. For example something like (please do not use this directly. I have no clue if this is the message you want us to take away from the table) "In Table 8.1. we show that when using the features Before Projection Head all augmentation methods reached 100 train and test accuracy. The Resize approach needed the fewest epochs."
 - table 8.2 can you somehow bring this into table 8.1 e.g. as another column? Would be easier to read. The first column "Augmentation Type" could be made smaller if you wrap the text so that it is on multiple lines.
 - tables 8.4. - the heading in the tables are too long and the table then needs to use very little font. Two solutions - wrap text so that it is on multiple lines or use abbreviations that you then explain in the caption E.g. (U), (B) ... (U) stands for Unbalanced. Or FD, 40S, CFD, C40S.
 - Tables 8.4-8.6 - again titles (captions) should be self-explanatory. Evaluations? Of what? K means? Also, what are the numbers? What metric? Accuracy? Something else. It is also good to help the reader by saying what is actually better, higher or lower value. Especially if not obvious.
 - page 47 table in top has no number? What does it actually show?
 - table 8.7 - Metrics? What metrics?
 - page 48 - CAE - every time you use an acronym, make sure that you have first spelled it out completely.
 - Conclusions - yes, what are they? What is the outcome. Important - it is fine to say that despite your best efforts the methods do not seem to deliver useful results. Perhaps they do not beat the baseline, this is very common in research and is perfectly fine, correct and useful to admit. Or they actually seem to perform randomly and are not reliable. This may be due to not enough data and is also ok to say, admit and point out. Or they do something but you are not actually sure, how to use for the final objective or ranking. This is also ok and you shall explain why you are not convinced. Perhaps the ranking is simply a badly formulated problem and it should not be approached this way at all. This is also ok to say and be open about it. Simply, even if it "does not work" it is still VERY useful if we understand what you have done and why you think it does not work.
 - References - clean these. Correct references shall have not only the names and year but also the publication venue (journal, conference, arxiv... etc.)