Peer review of Erik Vasquez Petersson & Robert Forsgren

By Mikael Melander

Background

First off, I would like to say that I like the project, it is interesting and something I, myself, is not very familiar with. And with that said I think that you have done a really good job describing the background, you give good explanations to how things work, why they work the way they do, and what is a problem and how you want to change it.

You present a valid research case based on the background. AI is the new in computer science and I think that, if used in a correct way really could change the software testing industry. When you talk about the problem, you mention that it is to solve problematic cases that can appear, and that it is to save time, effort and money. And you mention the clear drawbacks from other forms of testing compared to what this project aims to create, like having to redo complete tests because changes to the ID's in the doom for example. So you mention their cons against your solution, but not much more reflection is made there.

- What pros do they have that you cannot complete with?
- And how often is the IDs in a doom changed counter, how often an imaged is changes to much to not be valid anymore?
- What about and AI solution for recognizing changes to an ID?

As said above, you mention cost as a motivator to use this kind of solution, since time is money. But you do not really go in to any cost situation, you say your solution uses Google Cloud Vision, and that, that is free to try. But not much more, is it really expensive after you try it? And what about the competitors that you compare to, are they free or what do they cost? It is not so much about the specific prizes here, but more to get a sense of it, since you mention it as a pro.

When talking about the competitors you mention they use some kind of "Similarity threshold" to account for differences in rendering, to detect changes, and that this is a similarity to pixel values. You say this can be really ineffective because a different pixel can be black counter white and that saying you think its ok to have a 10% change, can cause a lot of differences. But then later on you mention you using a threshold to decide if your results are reliable or not, but this threshold is not explained in any real way. I would like an explanation for this in the report, so we get to know what it does more in depth and why that is better than what you previously mention as something unreliable.

I would also have liked to see some more examples of the test suit, you show an image of what would be seen as a valid change in Figure 2.2. But I would have liked to see at least one of the before and after images of what is used when testing, it could have been as an appendix. This was even more clear during your presentation when you showed an demo, it explains a lot more then text can, especially in these sort of situations.

Overall I really like the report, you did a good job from explaining the background to your conclusion. I actually learned quite a bit. You explain your method and your choices in a good way, your layout is from what I can see correct, as well as your references. Good work!

Spelling and grammar

The use of “ is made in more than one place, for example 1.2 Related work, line 3. That should only be used for citations as mention on the course web place. It however is used correctly together with some references as well.

Under problem formulation, you say “mentioned in related research”, I am guessing you mean related work.

Figure 3.2: Explanation sentence is a bit weird, “REGTEST’s main window with a test selected, and one if its test steps, which is of the type IMAGESNAP.”

It is mostly small stuff, but I figure since I noticed it, it would be good to mention to you.