

Group: Taylor Berg-Kirkpatrick
 Feedback provided by: Christine
 Writing Hub visit: No

| | Grade | Comments |
|--|------------------|---|
| Context and Motivation | Pass | Your motivation is pretty clear, but it's not clear why OCR can help with finding exact publishers for anonymous documents or why they can help understand linguistic practices. |
| Problem Statement | Not Pass | There is no clear problem statement in part 1. You talk about how dhSegment "did not achieve comparably great results" but I want to understand more specifically what it lacks. That is the problem you seem to be proposing to address. Be more specific and add more detail in the first section. Some of this text can be taken from the second section (see comments in the proposal). |
| Literature Review | Pass | Your work includes appropriate references to previous approaches, though it would help to better position the graph-based methods and the dynamic programming methods relative to the NN based methods. |
| Proposed Solution | Not (quite) Pass | It's clear to me that you are going to extend the model used in dhSegment to enforce consistency over the labels using a second model, so that's good. But after that, I get a little lost. I don't really see how the two models fit together, or which one is new vs. existing. When you are more precise about the problem you are solving, this will be easier to explain and for the reader to understand. |
| Presentation of Technical Details | Not Pass | <p>A big challenge I had understanding your approach is the details of how it will work. I have several questions in the text. Most importantly:</p> <ul style="list-style-type: none"> • Be more clear about what Y is in your model. I can't understand what $P(X Y)$ or $P(Y)$ is without this information • Give more detail on what a probability map is and what it means for a pixel to be determined to be part of a line. Which line? Any line? <p>There are also a few technical terms that could use more clarity, eg. "baselines", CNN architectures (how do they relate here?)</p> |
| Evaluation Plan and Timeline | Pass | I'd like to understand more about your test set and how you will quantify accuracy. See the comments in the text. Your timeline looks good. |
| Writing Mechanics | Pass | <p>The main things to work on in improving your writing are:</p> <ul style="list-style-type: none"> • Making sure paragraphs stick to a single topic. You sometimes have paragraphs that do too much, and other times you have sentences that should be placed with the subsequent paragraph. • Tightening up the wording. You tend to use more words than you need to throughout. Be direct and to the point. |

| | | |
|----------------|----------|---|
| Overall | Not Pass | You have a good start here. When you clarify your specific problem, organize your writing, and be really clear about what everything means in the models you are proposing, this will be a strong proposal. |
|----------------|----------|---|