Reflective Essay

Chris Oh

Computer Vision (CV) is an interdisciplinary field that uses formal learning to teach computers how to recognize patterns in digital images and videos. CV is revolutionizing the way we interact with technology. Currently, CV is gaining traction in fields like traffic monitoring, video games, and self-driving cars.

The initial problem with Artstor was that users were unable to search for specific art images using the images' metadata. The research topic was to solve this problem by generating and adding meaningful metadata to these art images using CV. This would enhance the user experience of Artstor and make more data reachable to its users, by allowing users to query images with the known metadata. During the first ideation of this project, our goal was to dive straight into developing CV algorithms to generate metadata from images. However, we later realized that no research currently exists around the topic of analyzing art images using CV techniques. So, we decided to first change the scope of our project and answer the question if it was even possible to extract and generate metadata from art pieces. We also set the duration of the project to be two academic Quarters long. I met with the team once a week and worked on the research individually for five to ten hours a week.

We researched different CV methods and found a method called Deep Mapping, which was used in an existing research project regarding the classification of scientific figures (1). When we successfully found methods to analyze and extract information from figures, we were able to move forward with our initial project scope and answer the question, "How do we generate metadata for images?".

The initial step of finding useful resources to reference was a difficult task due to CV being a relatively new field of study. However, being involved with the DataLab supported my research project because there was ongoing research related to CV by other teams within the DataLab.

When the project was being formulated, I needed resources to reference and to learn more about the field of indexing and extracting metadata at a large scale. The first thing I decided to do was visit the Suzzallo & Allen Library. I asked the librarian about my research topic and what would be a good place to start looking. I was told about the library's database, which holds references to other publishers that the UW network has access to. Using the database and querying "Computer Science" I was able to find "Google Scholars", which I already knew, but I also found "ACM Digital Library" and "IEEE Xplore", both which were new to me. From those sources, I was able to find an abundant amount of rich content and information that supported my research.

One of the greatest leads was learning about existing conferences related to CV such as, CVPR, ICCV, and ECCV. I found all the papers that were published in these conferences, and also learned that the UW provides access to these published papers. Once I discovered where to find the most relevant and up to date information, I started to read through all the findings for more understanding of the field and related research.

Another valuable technique I learned during my project was querying for information. Rather than using basic searching with keywords, I learned to use advanced searching. Some filters I user were terms to group, excluding unwanted terms, limit a range of dates, and sorting by reference counts. It allowed me to control my queries similarly to writing SQL statements.

I also found it crucial to document my processes. Coming up with meaningful search queries was important but recording all the queries I used was even more important for reproducibility reasons. I had trouble finding some papers which I found earlier on because I didn't document the terms used to search for the particular papers.

Since this research was related to a relatively new field, it was challenging to find existing papers which conducted a similar research. The one research found that was extremely helpful was "Deep Mapping of the Visual Literature" (1), which conducted a similar study by analyzing scientific figures in thousands of research papers.

This project also involved collecting variations of art images for training computer models. The biggest challenge with this was finding and collecting reliable sources. About half of the research period was spent collecting and cleaning data for processing. Some of the sources used for image data were "The Metropolitan" and "Artsy", which are museums with large collections of publicly accessible images. Choosing the two as a reliable data source required validation of the various metadata that were already labeled. Even if sources contained abundant amounts of metadata available, if none of them included fields that our project required, such as categories, material, etc, they were considered not useful.

The entire process of determining a research topic, finding valuable resources, evaluating sources, and learning how to research itself was a challenging but rewarding experience. Conducting this research in the DataLab also enhanced my understanding of the rather abstract concept of CV and deep learning through hands-on application. I am able to continue to fulfill my curiosity for the field of CV through research projects like this. Also, this project experience will allow me to further apply my knowledge on other adjacent fields to CV. I am confident that the knowledge gained from taking part of this research will serve as a stepping stone for my overall goal of pursuing a doctorate degree in the field of CV.

1. Bill Howe, Po-shen Lee, Maxim Grechkin, Sean T. Yang, and Jevin D. West. 2017. Deep Mapping of the Visual Literature. In Proceedings of the 26th International Conference on World Wide Web Companion (WWW '17 Companion). International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 1273-1277.