**Computer Vision Project Stage 2**

Problem Statement: Our group’s focus will be tackling the problem given to us by Sonic Automotive, extracting mileage data from Odometer readings.

**2.1 Research**

Your research should be spread across a few key areas. Depending on your project, your research may be more "business focused" or more "academically" focused. You should address at least 10 of the questions below, and must choose and review 3 academic papers.

**2.1.1 Business/Customers**

* What problem will your Computer Vision solution solve, and for whom?
  + Our Computer Vision solution will solve the crucial problem of identifying odometer readings for Sonic Automotive.
* What value will it provide them? What are their pain points?
  + Sonic Automotive frequently sells used vehicles at a price partially based off of the odometer readings. Right now, each used vehicle is being processed manually which is slow due to the high volume of used vehicles being processed for selling.
* How big is the potential market?
  + The market of this could potentially be worth thousands to tens of thousands of dollars, as hundreds of hours of manual labor would be saved.
* Do other similar solutions exist?
  + There have been very similar problems solved by other researchers. For example, in the article *Mileage Extraction From Odometer Pictures for Automating Auto Insurance Processes,* researchers found a “a novel solution to the insurance-related problem of extracting mileage readings from odometer images” (Acharya, S. & Fung, G. 2019).
* Would your business have any competitors? Who are they? How are they doing?
  + Right now there are no known competitors to this area of business. However, there is a possibility of businesses addressing the same problem to streamline their business processes.
* How are potential customers dealing with these issues now?
  + Right now, customers are not dealing with these issues as they are completely dealt with by employees at Sonic Automotive.
* Are your customers individuals or businesses? Be sure to cite the sources you reference in your research.
  + Our customers would be businesses. While we are working with Sonic Automotive to help solve their problem, this solution could be used and applied by other businesses with a similar problem.

**2.1.2 Academic Literature Review**

A thorough literature review can save weeks of wasted time implementing dead-end solutions or re-doing work that others have already done. You should pick 3 academic papers (we can help you with this, if needed), and do a "deep-dive" on these publications.

* What academic work is relevant to your project topic? Pick 3 papers, ask us for help if you need it.
  + Goodfellow, I. J., Bulatov, Y., Ibarz, J., Arnoud, S., & Shet, V. (2013). Multi-digit number recognition from street view imagery using deep convolutional neural networks. *arXiv preprint arXiv:1312.6082*.  
    [<https://arxiv.org/pdf/1312.6082.pdf>]
  + Acharya, S., & Fung, G. (2019). Mileage extraction from odometer pictures for automating auto insurance processes. *Frontiers in Applied Mathematics and Statistics*, *5*, 61.  
    [<https://www.frontiersin.org/articles/10.3389/fams.2019.00061/full>]
  + Wang, X., Meng, L., & Xue, Y. (2020, January 01). Automatic bounding-box-labeling method of occluded objects in virtual image data. Retrieved March 13, 2021, from [<https://dl.acm.org/doi/pdf/10.1145/3381271.3381292>]
* What makes these papers important/relevant?
  + These papers are important because they address methods that we can do to help identify the values within the images and accurately identify the odometer readings

**2.1.3 Open Source**

* What open source code is available that are relevant to your topic?
  + There are a couple of open source programs on github that offer similar strategies, such as an android application that can find phone numbers from an image using google OCR - link: <https://github.com/CCalvin457/image-scanner>

**2.1.4 Industry Solutions**

* What companies are solving similar problems to yours?
  + There are current applications that drivers can use to track their mileage, however these are mainly for car consumers not for main auto companies
* Has anyone reverse engineered these products?
  + Not based off any information that could be found

**2.2 Data Collection**

We have taken the dataset provided by Sonic Automotive and performed some preprocessing and labeling to the images. We have manually extracted the values of the mileage/kilometers and units corresponding to each picture and stored these for use in accuracy checking later on. We have also created a program to put bounding boxes around areas of text that appear in the picture and plan on optimizing this in the future to only read the mileage/kilometers count in the images.