# Group 9: Pedestrian Traffic Prediction


**Summary:**

- The goal of this project was to identify pedestrians in intersections and collect data about their movement. This is meant to make intersections a safer place for pedestrians, and also benefit drivers. CVAT, YOLO, and RNN algorithms were used to annotate, identify pedestrians, and predict whether or not they crossed the street respectively. The models were trained on hours of footage from intersections containing pedestrians. The YOLO model was able to identify pedestrians accurately, but the RNN struggled with classifying whether or not they crossed the street, possibly due to data limitations. In the future this project could be expanded by using additional footage from a wider variety of intersections, and also by exploring some different types of algorithms, like transformers.


**High Level Problem Description: P-2**

- A good description of the problem and what could be helped when utilizing this technology. We think it would have been nice to give an exact example of a pedestrian safety project that this might be able to be applied to because just reading this it doesn't immediately come to mind what kind of project this could be applied to. Additionally, we see how this technology could improve the efficiency, but would it really improve the accuracy over a human annotating it?

**Technical ML Problem Description: ML-1**

- Some of the problem is described, but a general machine learning level problem is not explicitly stated. The objectives of the overall project are stated, the models used are stated, but only the input and output portions of this section really describe the machine learning problem. This section has mostly information about how the problem was solved, not what it is.

**Data Used: DT-2**

- The data used is clearly stated and makes sense for this type of project. Regarding the portion about removing 60 clips due to noise or too many pedestrians, while this might add noise or be hard to annotate, wouldn't this also make the model not able to generalize well to real situations? Additionally, what was the thought process behind not including scenes where there were no pedestrians crossing? Was it just a time constraint thing or made it harder to run the model? We thought this may help with the classification problems discussed in the results section.

**Algorithms Implemented/Why: A-1.5**

- A good description of the algorithms used, however, we think it would be useful to describe what YOLO is made of–is it a trained machine learning model? If it is, what kind of model is it? Why is the approach it takes much faster? Additionally, was there a reason behind only having two LSTM layers for the RNN? Computation speed? Led to better accuracy? We think having answers to those may have made this section a bit more robust. It is clear what the advantages of each algorithm are but some further explanation of how they improve upon other algorithms would be helpful as well.

**Final Results: R-2**

- There is a helpful example and also loss graphics that clearly show the results of the model training. An example of how the models identify the pedestrians and where it goes from there would have been cool to see but there are some clear results given from the semester's work.

**Discussion of Results: D-2**

- The limitations of the models are clearly discussed throughout the report. The scope of the project is clear. There are explanations for some of the shortcomings in model performance. We would have liked to see some more about what the model does well, especially compared to other similar projects, but based on the results it seems as though some parts of the project did not turn out exactly as planned due to data and time limitations.

**Next Steps: NS-1.5**

- A couple good examples of what next steps might be, but we feel like they're a little bit lacking. Yes, more data and trying a different model are good next steps, but what could be some next steps specific to the group's problem? Are there any next steps that the group could look into to solve some of the issues related to their current model. We noticed they didn't mention diversifying the dataset, which would be one of our first next steps if we were in their shoes. Additionally, for next steps with something like a transformer, there was no discussion around how that would help them solve their problem. It's important to try new things, but it would have been interesting to hear why something like a transformer may help the project in the future.

**Feedback: F-2**

- Feedback was addressed well, however, the portion about how it would be helpful is still a bit vague. Yes it would be more efficient than people annotating, but what kind of projects could this be applied to? What specific pedestrian safety tasks would this improve the efficiency/effectiveness of?

**Novelty: N-0.5**

- There doesn’t seem to be a ton of novelty to this project, mostly coming from the particular videos that were used and the RNN that was constructed. Detecting pedestrian crossing and pedestrian trajectory has been done before using CNNs and RNNs, though, perhaps the use of YOLO in combination with the RNN is novel as we could not find anyone else using that combination. The results didn’t seem particularly novel either, though, something might have been if there was more diversity in the dataset and if better results were able to be obtained. We’d be interested to know if the novelty of this idea was researched beforehand at all and what the group came up with.

**Learnings: L-0**

- As far as we could tell, there weren't a ton of mentions about what the group learned. We think that it would have been beneficial to the report to maybe reflect more on the dataset and the limitations of it and what the group took from that. Additionally, we’d like to have really seen what the group learned overall about the problem: i.e. how difficult it is to detect pedestrian crossing, what are the natural limitations to this style of pedestrian detection, did it end up being good enough to replace humans, etc…

**Future Work Suggestions:**

- If this project were to be continued, we’d like to see how well it would fare in more noisy circumstances. It was mentioned in the data preprocessing portion that videos with a lot of people or weird weather were removed. We’d like to see results with those videos included, even with just the current models that the group has, just to see if the model would learn any better with those videos included than it did without them.
- We would also like to see some more exploration on the specific use cases of this technology. We’d like to see more tests on which projects might benefit from it the most. Additionally, we’d like to see if it works better to have the model both detect pedestrians *and* their path, or if it would work better for the model to just detect pedestrians and cut down video to the points where pedestrians are crossing in order to allow humans to find the paths. Discussion around this would be interesting to see if the project was continued.
- We would like to see an exploration of how the RNN model could classify differently. It is mentioned that the model would only predict one class, which is attributed to the small data set, however we feel as though there may be alternative methods for this classification that could give new results. It would be interesting to see if the model was “considering” classifying differently but the probability just kept the one class as the maximum for every case. This would give more insight into how well the model actually worked despite the problems and might give clues as to what is missing to make improvements.