

CS 566 Project - Midterm Report

Zachery Fleming

Webpage: <https://zsfleming17.github.io/566-Project/>

Current Progress:

- I collected and labeled 231 photos from 4 MLB stadiums (AmFam field, Wrigley field, Oracle park, Fenway park).
- Created a CSV file to track image paths, labels, and stadiums.
- Implemented automated train/validation/test split scripts, including random 70/15/15 and Leave-One-Stadium-Out (LOSO) splits.
- Built training pipeline in PyTorch using MobileNetV3-Small with ImageNet weights.
- Developed preprocessing and augmentations such as random crop, flip, blur, color jitter, perspective.
- Conducted 2 key tests:
 - Random Split (n=35 test) – Accuracy 83%, Clear Recall = 1.00, Obstructed Recall = 0.57.
 - LOSO-Wrigley (n=54 test) – Accuracy 93%, Clear Recall = 1.00, Obstructed Recall = 0.85.
- Generated visualizations:
 - Sample grid of dataset images.
 - Confusion matrices for random and LOSO results.
 - Set up project webpage with those visualizations.

Current Results:

- Baseline model *mostly* successful when distinguishing clear vs. obstructed seats.
- Model achieves high accuracy on the “clear” class but can struggle with not so obvious obstructions (e.g., thin rails, light poles).

- Cross stadium testing (LOSO) shows strong generalization. Unseen stadiums were shown to still be classified accurately.
- Confusion matrices confirm that most errors were false negatives, meaning some obstructed seats were predicted as clear. After I visually inspected these misclassifications, they mostly involved the outfield being obstructed rather than the infield.

Difficulties and Adjustments:

- **Ambiguous labels:** Some images contain people or partial obstructions that don't block the actual field of play or aren't part of the stadium architecture (meaning they should be labeled clear). During training and testing I haven't seen major issues with these "edge cases" being misclassified, but the model currently behaves conservatively when predicting obstructions. I plan to adjust to this by refining label definitions for these "edge cases".
- **Limited dataset:** With only 230 images, the model occasionally overfits and struggles with not so obvious obstructions. I plan on getting much closer to a 50/50 split, which may take more stadiums that tend to have more obstructions.
- **Imagine imbalance:** Clear seats outnumber obstructed ones (120 vs. 90), slightly biasing predictions toward "clear." It was a bit more difficult than anticipated finding obstructed seats that varied compared to clear.
- **Changes from proposal:**
 - Added structured CSV metadata instead of manual labeling.
 - Added LOSO cross-validation earlier than planned to analyze generalization.
 - I deferred Grad-CAM and obstruction labeling (obstruction_type, blocks_field) to the post-midterm.

Next Steps:

- Expand dataset to 400–600 images across additional stadiums.
- Refine labeling policy to separate structural vs. human/temporary obstructions.

- Add Grad-CAM visualizations to interpret model focus areas (confirm if poles/netting drive decisions).
- Run LOSO tests for each stadium and compare performance variance.
- Explore stronger augmentations to boost obstructed recall.
- Update the webpage with new figures and analysis by the final report.