Hockey Puck Tracking: Report

Dataset Analysis and Preprocessing

Aspect Ratio Correction

During initial dataset inspection, we identified a significant aspect ratio distortion:

- Original video aspect ratio: 1920:1080 (16:9)
- Training images aspect ratio: 1240:1240 (1:1)
- This stretching effect distorts the natural shape of objects, potentially affecting model performance
- Implementation: Created aspect ratio transformation functionality in `preprocess.py` to restore correct proportions.

Bounding Box Analysis

Analysis of annotations using `bbox_check.py` revealed inconsistencies in bounding box dimensions:

Statistical Analysis Results:

Mean Dimensions:

- Width: 0.0187 (normalized)
- Height: 0.0208 (normalized)

Standard Deviations:

- Width: 0.0153 (81.8% of mean width)
- Height: 0.0146 (70.2% of mean height)

These high standard deviations indicated significant inconsistency in annotation sizes for the same object (puck), necessitating standardization.

Preprocessing Implementation

Combined preprocessing steps in `preprocess.py`:

- Aspect ratio correction: Transforms images from 1240x1240 to 1920x1080
- Bounding box standardization: Sets all boxes to mean dimensions (0.0187 x 0.0208)

Model Training

Training Configuration

- Model: YOLOv11x

- Input resolution: 640x640

- Epochs: 10

- Dataset: Preprocessed AllHockey.v1i.yolov11

Training Results

Validation:	**P**	**R**	**mAP50**	**mAP50-95)**
	1	0.809	0.904	0.383
Testing :	**P**	**R**	**mAP50**	**mAP50-95)**
	0.84	0.656	0.731	0.312

Refer to the runs folder for detailed information.

Inference and Testing

Tracking Implementation

Tested two tracking approaches:

1. ByteTrack (default)

2. BOT-SORT

- Better handling of fast movement

- IOU threshold: 0.7 for similar object discrimination

- Average inference time: 42ms per frame

Performance Analysis

Despite promising dataset metrics (mAP50 = 0.731), real-world testing revealed several challenges:

1. Camera Angle Discrepancy

- Training data: Focused field of view on game play
- Test video: Wider environmental view

- Impact: Reduced puck visibility and clarity

2. False Positives

- Increased environmental objects in frame
- Similar-looking objects being misclassified as pucks
- Impact: Reduced precision in real-world scenarios

Proposed Improvements

SAHI (Slicing Aided Hyper Inference) Implementation

SAHI could address current limitations by:

- Slicing large images into smaller sub-images
- Running detection on each slice
- Merging results

Benefits:

- Better small object detection
- Reduced impact of varying camera angles
- More robust to environmental variations

Resolution Enhancement

Increase model input resolution to 960:

- Current: 640x640
- Proposed: 960x960

Benefits:

- Better feature extraction for small objects
- Improved detection of distant pucks
- Trade-off: Increased computational cost