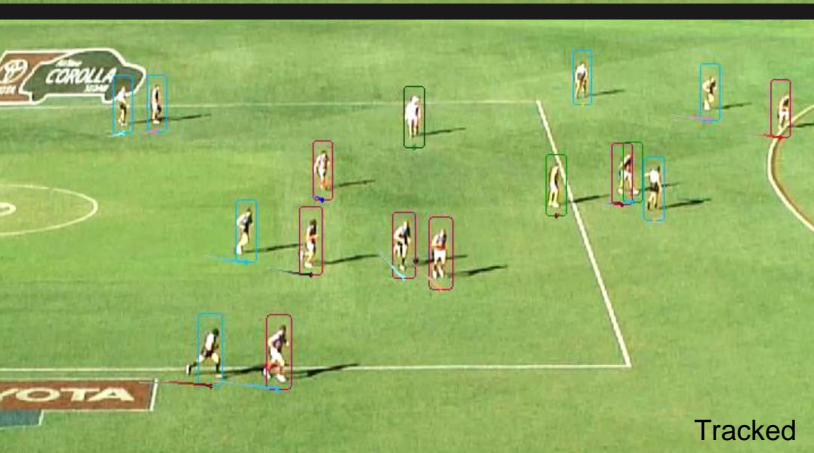




Original



Tracked



Digital Image Computing Techniques and Applications, 2015

AFL Player Detection and Tracking

Hayden Faulkner, Dr. Anthony Dick

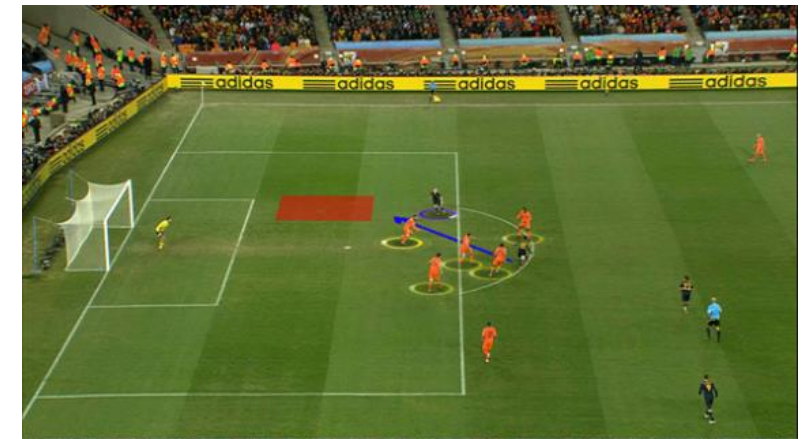


THE UNIVERSITY
of ADELAIDE

seekLIGHT

Introduction

- Pedestrian detection and tracking widely researched
 - Surveillance, vehicle navigation, human/computer interaction
- But what about sports?
 - Provide a foundation for automated game statistics, including player movements, interactions and events
- Sports player detection and tracking researched only in a few cases
 - Mostly soccer and basketball
- In Australia, the most popular sport is Australian Rules Football (AFL)
 - AFL funded the work
 - Vision detection and tracking methods yet to be applied and studied



Top Image: <http://www.eurocarnews.com/media/pictorials/2294/13861.jpg>
Middle Image: http://www.tvtechnology.com/BE_Files/uploads/2013/08/Orad%20patent.jpg
Bottom Right Image: <http://www.tourstogo.com.au/images/photos/zoomed/d6535cea.jpg>

Challenges

- Large field size leads to low resolution of players on far side of field (~ 40pixels in height)
 - Adelaide Oval (167m x 124m) 2x size of Hindmarsh soccer stadium (120m x 80m)
- Outdoor play leads to great exposure variations across different field positions
- Large number of people 18 per team, 9 umpires, ~20 other officials
- Extreme pose and appearance variation
- Regular occlusion of multiple players
- Fast speed and direction changes of players



Mild Pose Variation



Extreme Pose Variation



Low Resolution

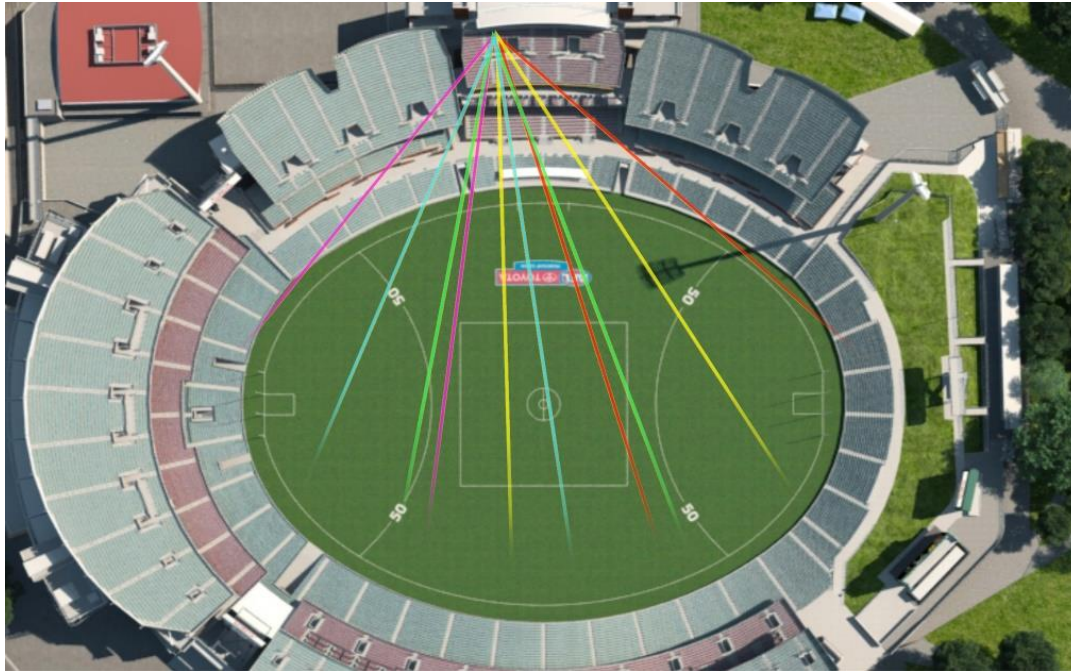


Exposure Problems

Occlusion Difficulties

Footage Capturing

- Broadcast footage not suitable:
 - Lack of control, constant cuts, camera movements, etc.
- Captured our own from top of grandstand:
 - Single vantage point - easy setup and monitoring, however not ideal for optimal resolution possibilities
 - Panorama style - cameras pointed at different areas of field
 - Height provides some relief from occlusion (see over players)

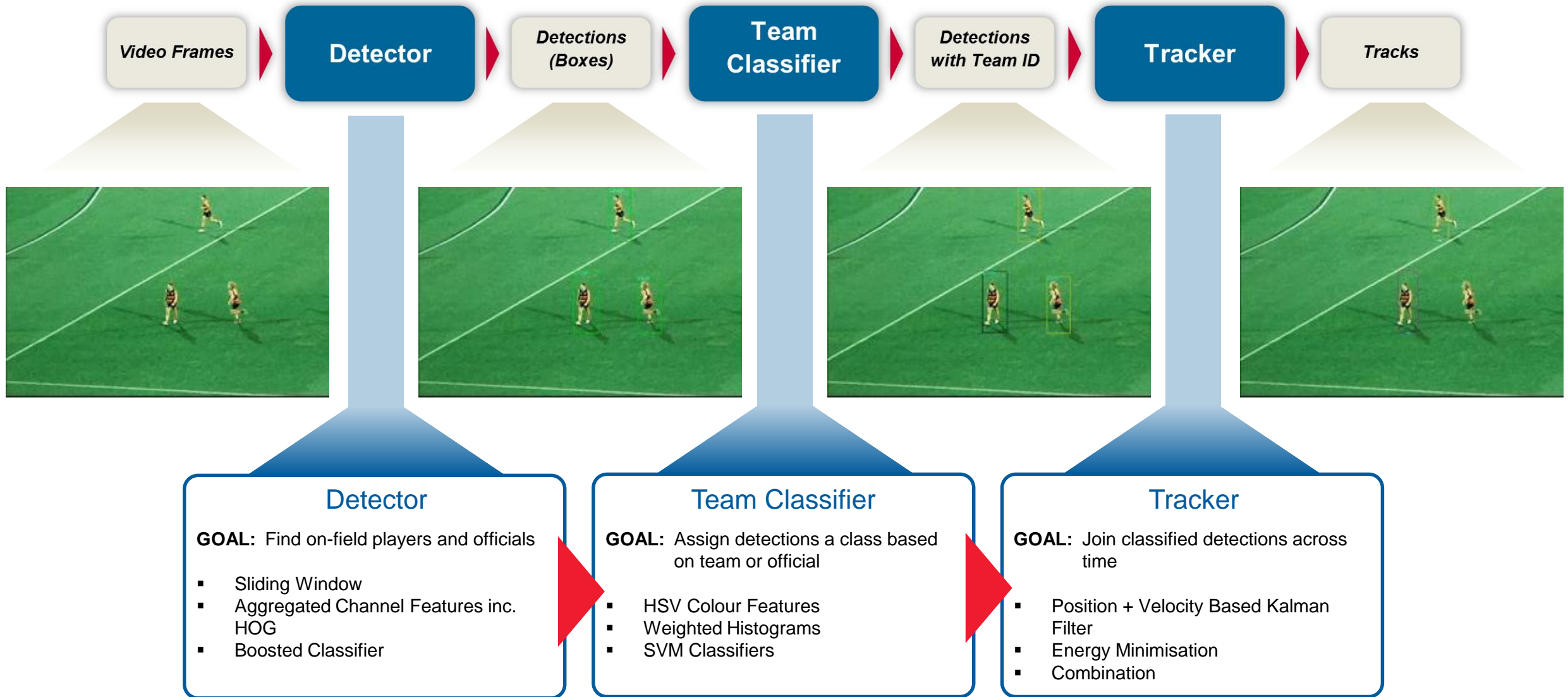


Top View of Oval showing Camera Overlap



Camera Setup

The Pipeline: Overview



The Pipeline: Player Detection Module

Goal: Find the position and size of on field players and officials, marking them with a bounding box

- It's important that the detection module be as accurate as possible
 - It is relied on by the following team classification and tracking modules



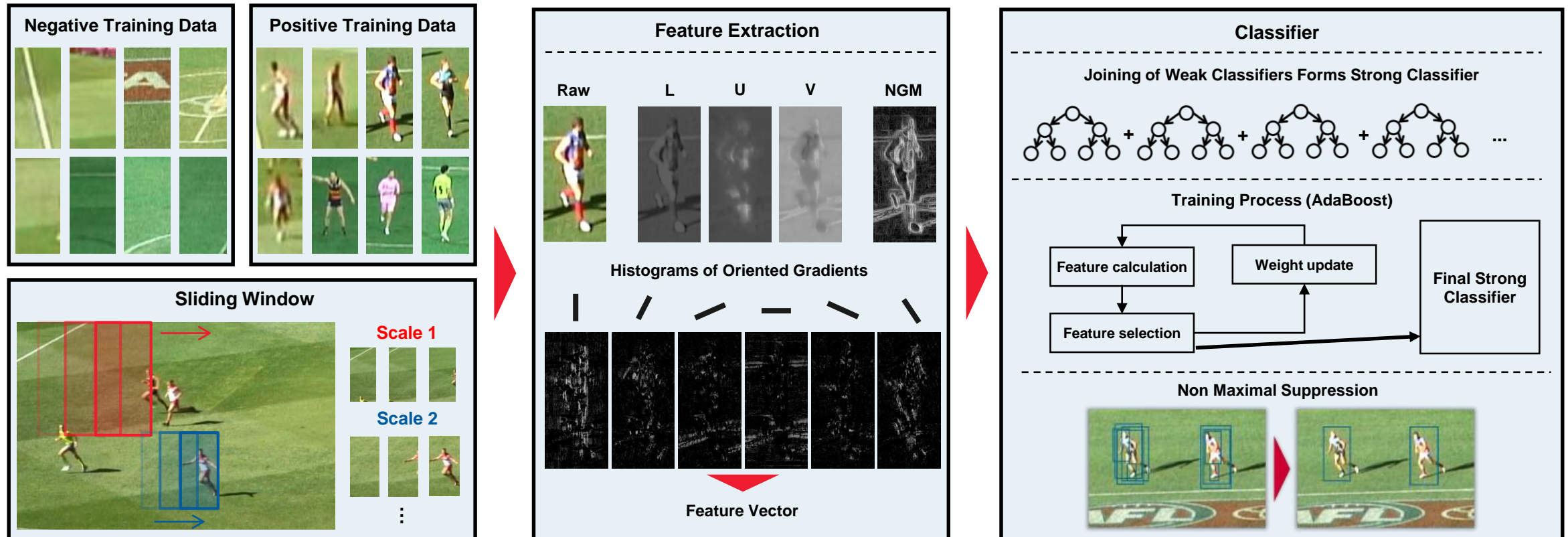
Original



Detections

The Pipeline: Player Detection Module

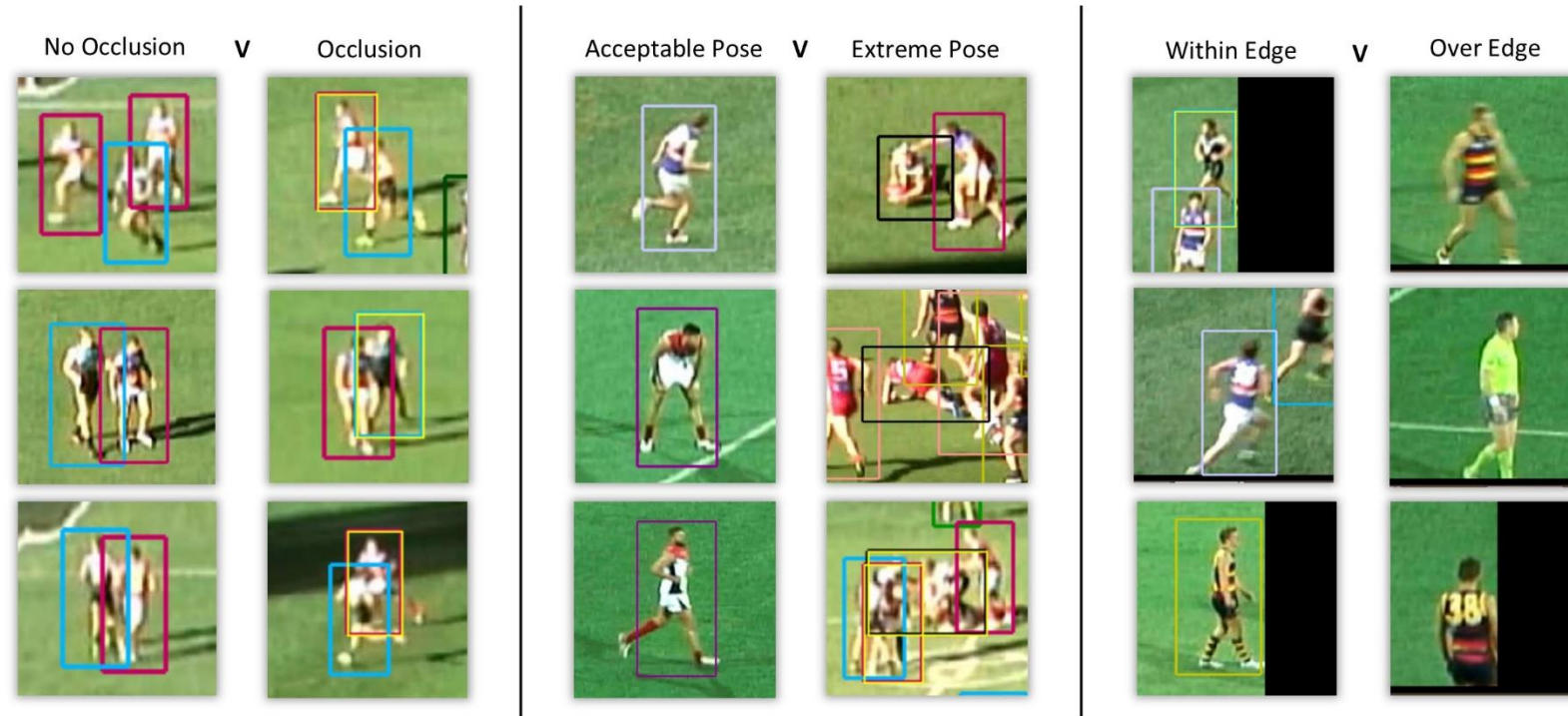
- Rescaling sliding window over entire frames
- Aggregated Channel Features (*Dollár et al. 2009/2010*):
 - Histograms of Oriented Gradients (HOG), Normalised Gradient Magnitudes, and LUV colour channels
- Boosted classifier of 2048 depth-two decision trees as weak classifiers (*Dollár et al. 2014*)
- Non-Maximal Suppression for combining multiple positive detections



Representative Overview of the Player Detection Module

Footage Annotation

- 13001 (1046 occluded) training samples; 5620 (423 occluded) testing samples
- 1:2 ratio bounding boxes
- Marked into team, marked with flag for occlusion
- Extreme poses not bounded to 1:2 bounding boxes
- Negatives sampled randomly from non-positive areas (30% positive overlap permitted)



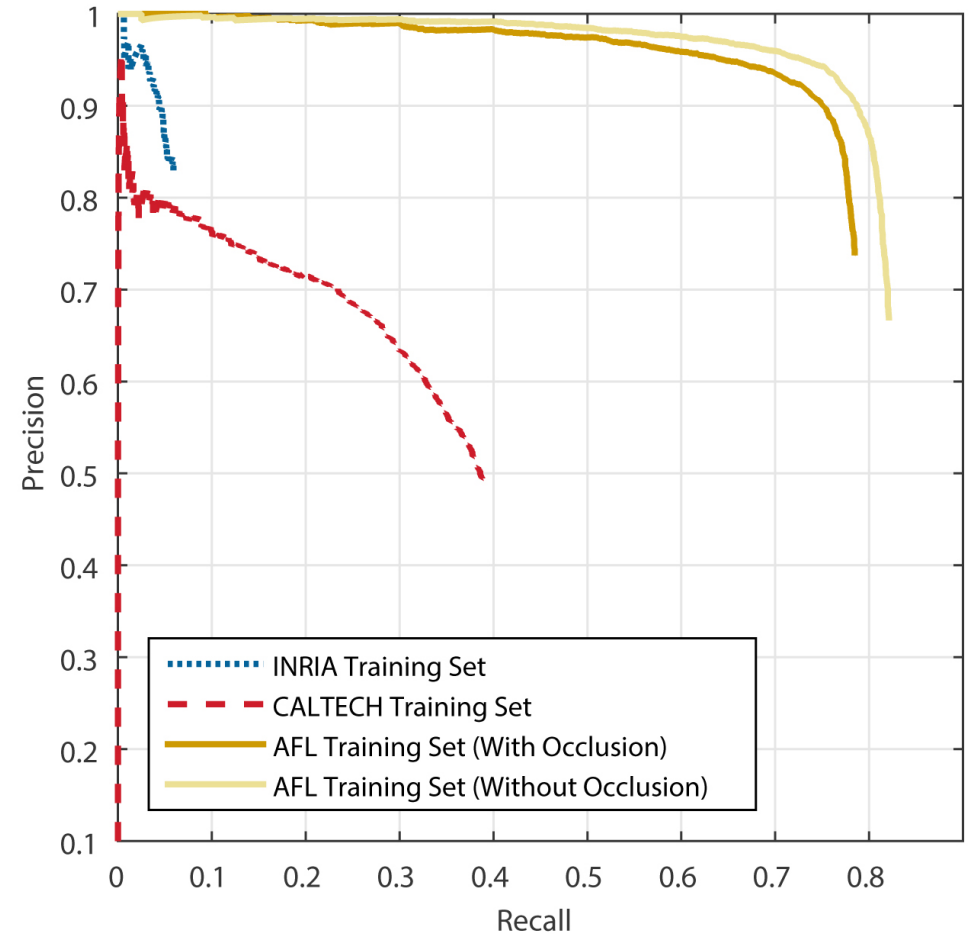
Player Detection Module Experiments

Comparison of Models Trained on Particular Data

- INRIA* vs CALTECH** vs AFL trained models
- Highlights need for particular AFL detector, trained with relevant data:
 - Pedestrians found in distinctly different environments
 - Pedestrians much more limited pose
 - Pedestrians captured from side-on / ground level



AFL Trained Model on some Test Images from INRIA



Precision-Recall Curve of Models Trained on Different Datasets

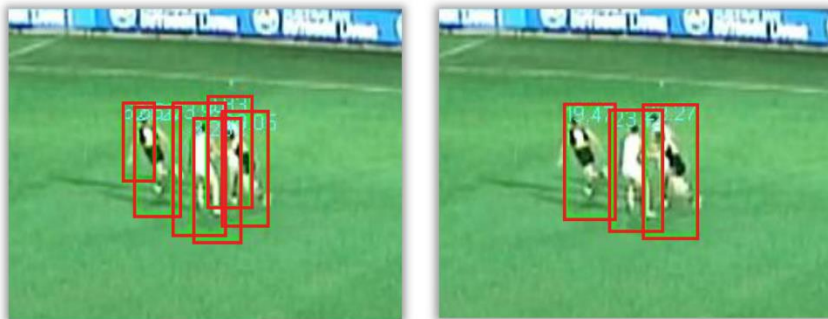
* INRIA: <http://pascal.inrialpes.fr/data/human/>

** CALTECH: http://www.vision.caltech.edu/Image_Datasets/CaltechPedestrians/

Player Detection Module Experiments

Comparison of Models Trained on Particular Data

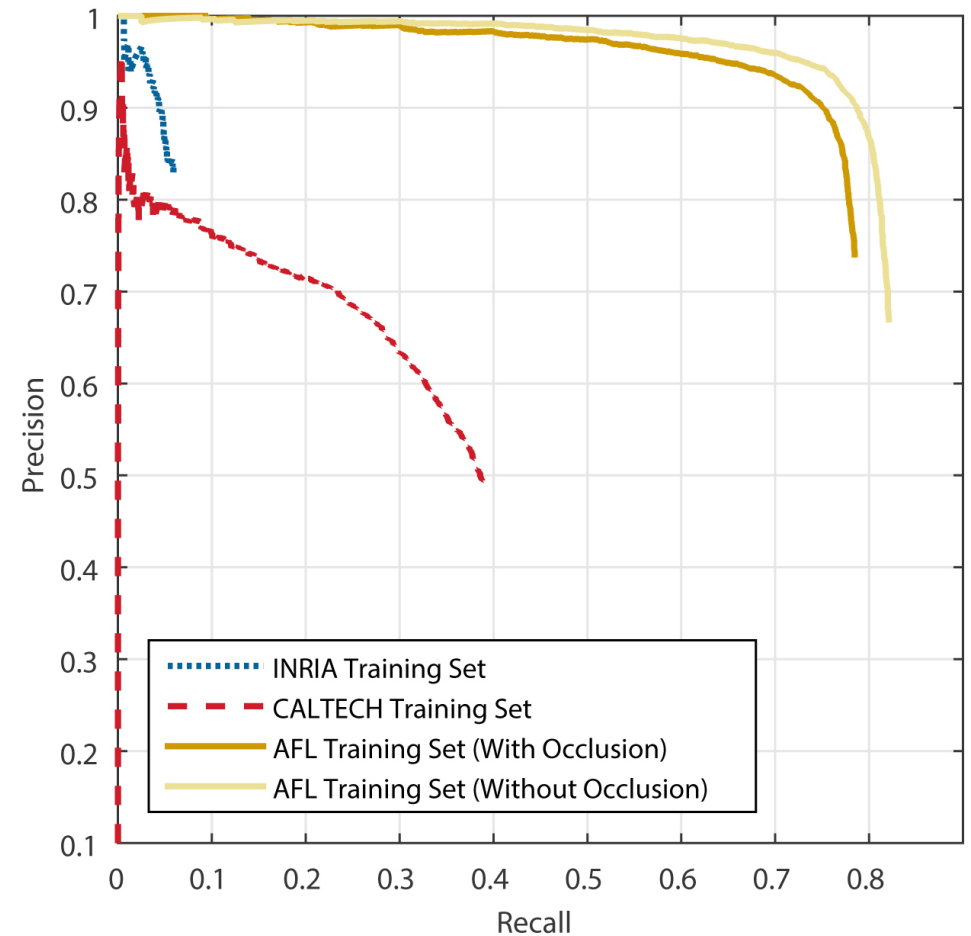
- Inclusion of occlusion samples results in detection of many more false positives
 - Especially in crowded, heavily occluded areas
 - Model has lessened ability to distinguish individual players



With Occlusion Samples

Without Occlusion Samples

Test Results of AFL Models Trained with and without Occlusion



Precision-Recall Curve of Models Trained on Different Datasets

The Pipeline: Team Classification Module

Goal: Assign each detection a class - team A, team B, umpire, runner, or other

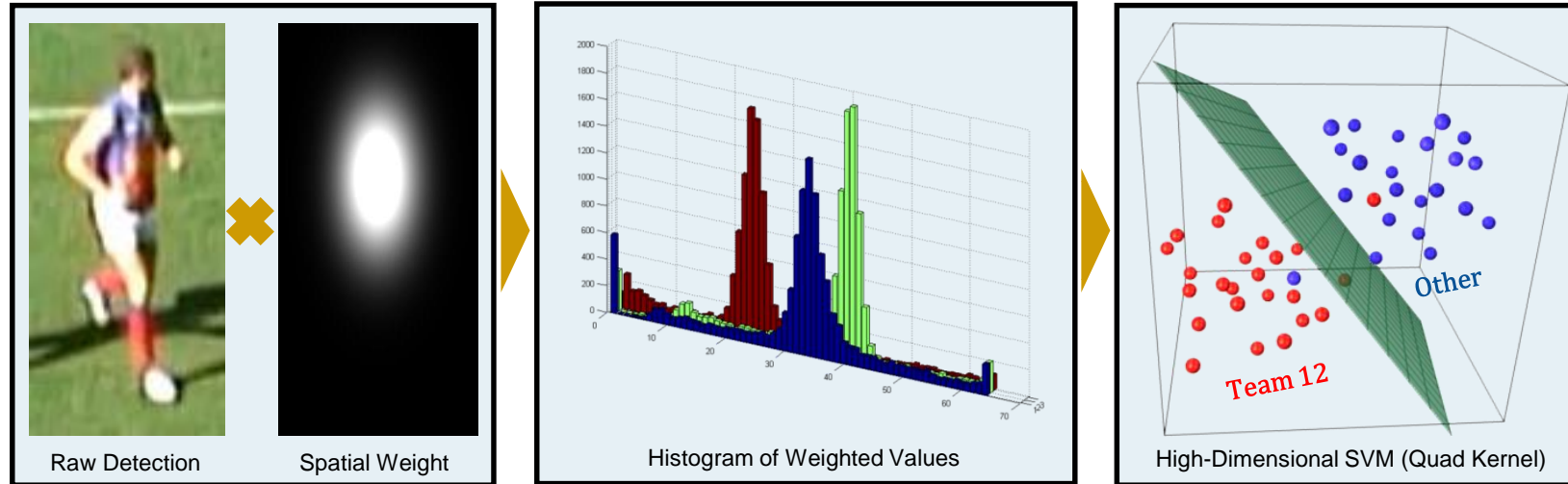
- Each class wears its own uniform consisting of different colours and patterns
 - However variation in pose, lighting conditions and resolution cause the same uniform to appear very different



Teams Captured Uniforms. They Look Different from Different Perspectives and in Different Conditions.

The Pipeline: Team Classification Module

- Histograms of HSV pixel intensities (64 bins for each channel)
 - HSV used over RGB as found better discriminative power 84.31% vs 90.37% MAP
- Support Vector Machine Classifiers with Quadratic Kernel
- Players Guernsey only covers 5 – 15% of detection box
 - Apply static spatial weighted mask (2D Gaussian) – increased MAP from 73.49% (no mask) to 90.37%

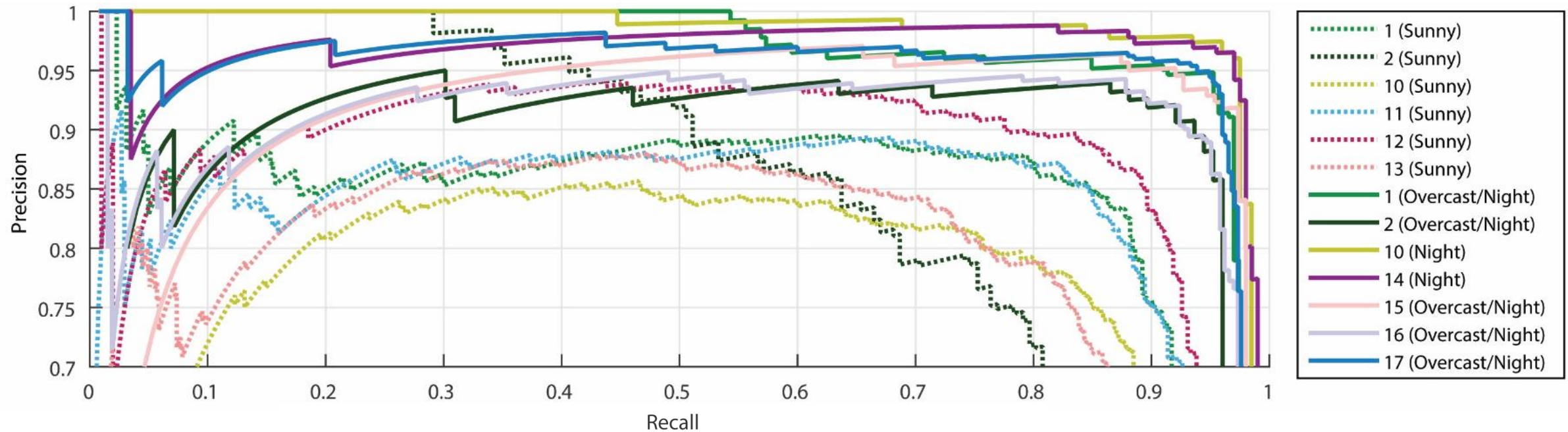


Representative Overview of the Team Classification Module

The Pipeline: Team Classification Module

Comparison of Teams and Environmental Conditions

- Some teams are classified better than others
- Main factor is lighting conditions with the night and overcast matches having best results
- Trained per team models with data from multiple matches, and per match models with data from only a particular match
 - Per Team MAP 85.98%, Per Team Per Match MAP 90.37%

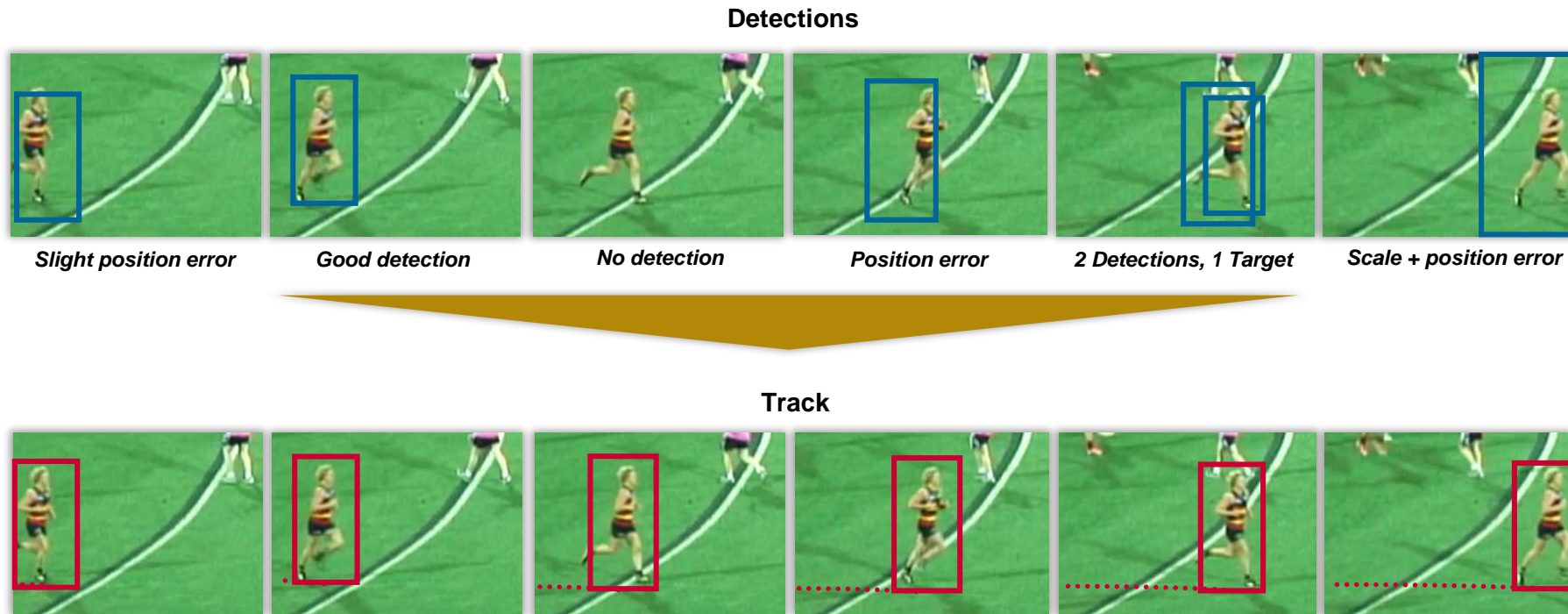


Precision-Recall Curve of Models Trained on Different Teams (Numbers/Colours) in Different Lighting Conditions (Line Style)

The Pipeline: Player Tracking Module

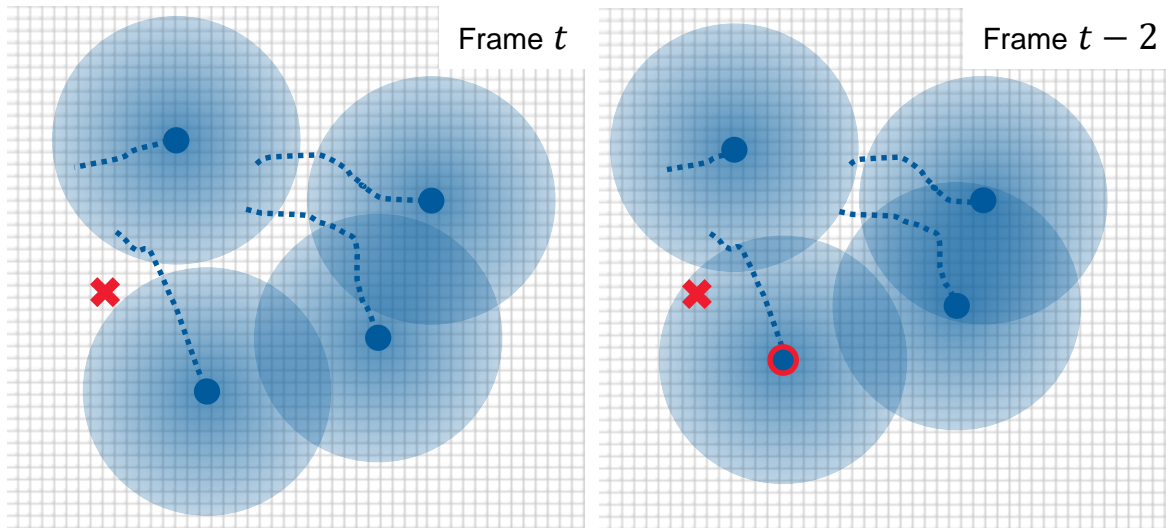
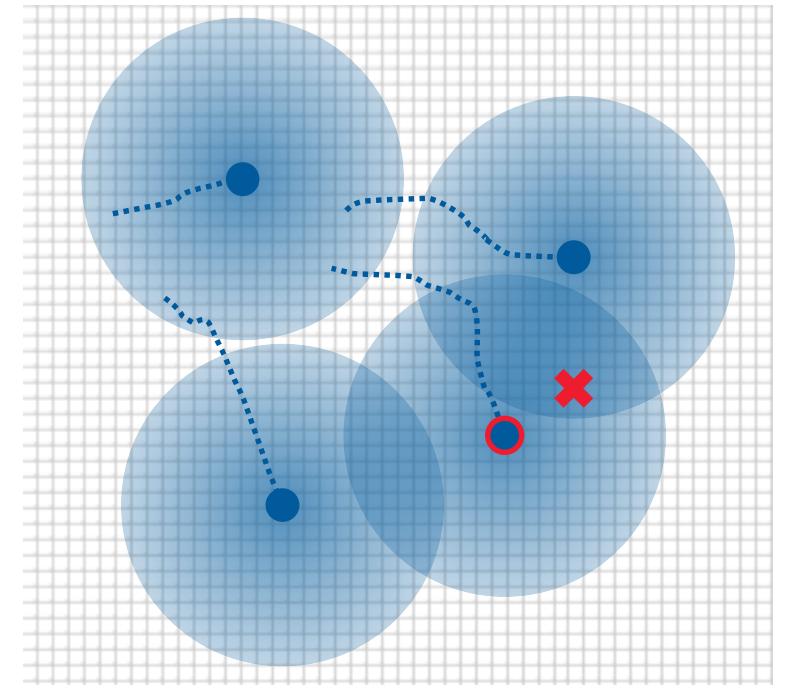
Goal: Join detections of the same players across time building tracks

- Tracking by Detection approach
- Can be difficult as detections aren't perfect:
 - Position or scale errors, false positives or false negatives

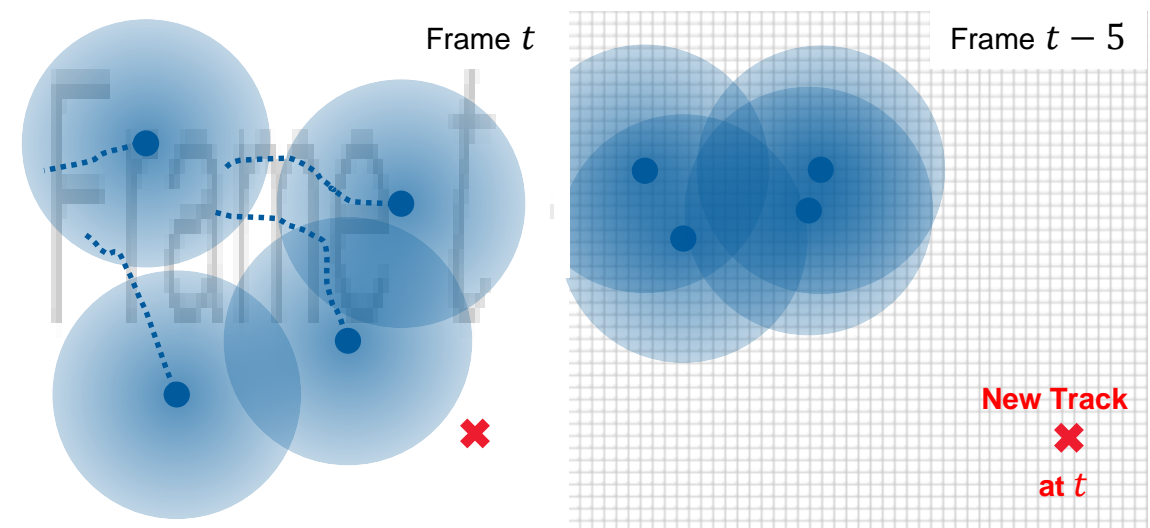


The Pipeline: Player Tracking Module

- Local Kalman Filter, based on velocity and position of players
 - Assigns detections greedily based on Euclidean distance within 20 pixel radius of predicted position of each track
 - If unassigned, checks past frames (up to 5), if still unassigned initialises a new track
 - Tracks lasting less than 20 frames are removed, assumed false positives



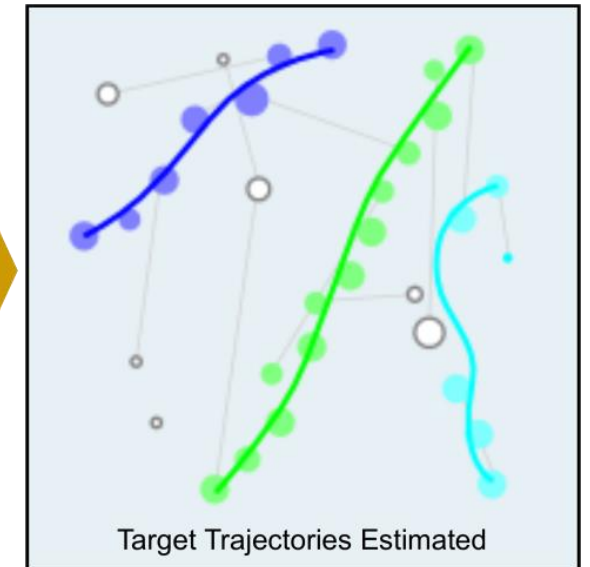
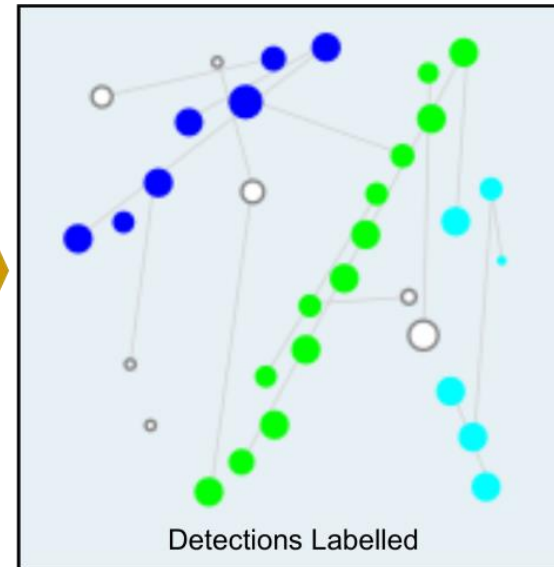
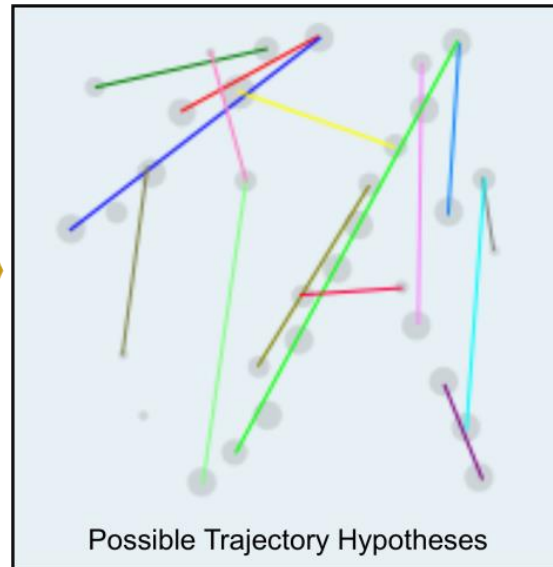
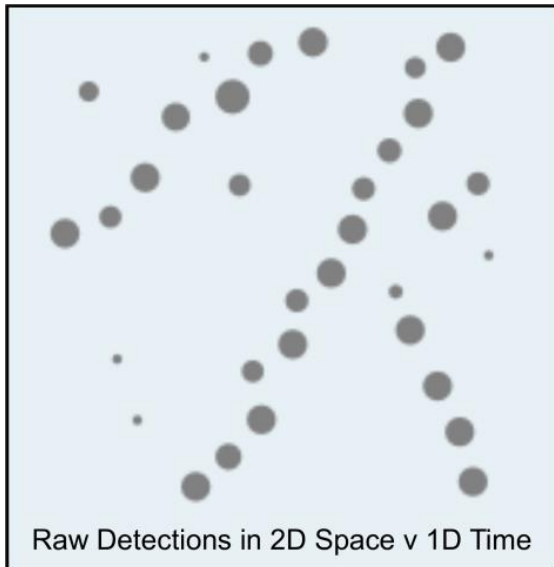
Detection assigned to track after looking back a few frames



Detection not assigned within 5 frames, new track initialised

The Pipeline: Player Tracking Module

- Global Discrete Continuous Energy Minimisation approach (Milan et al. 2014)
 - How would this global method compare to the local method?
- Combined, Energy Minimisation refines an initial Kalman Filter solution



Representative Overview of the Player Tracking Module

Player Tracking Module Experiments



Kalman Filter



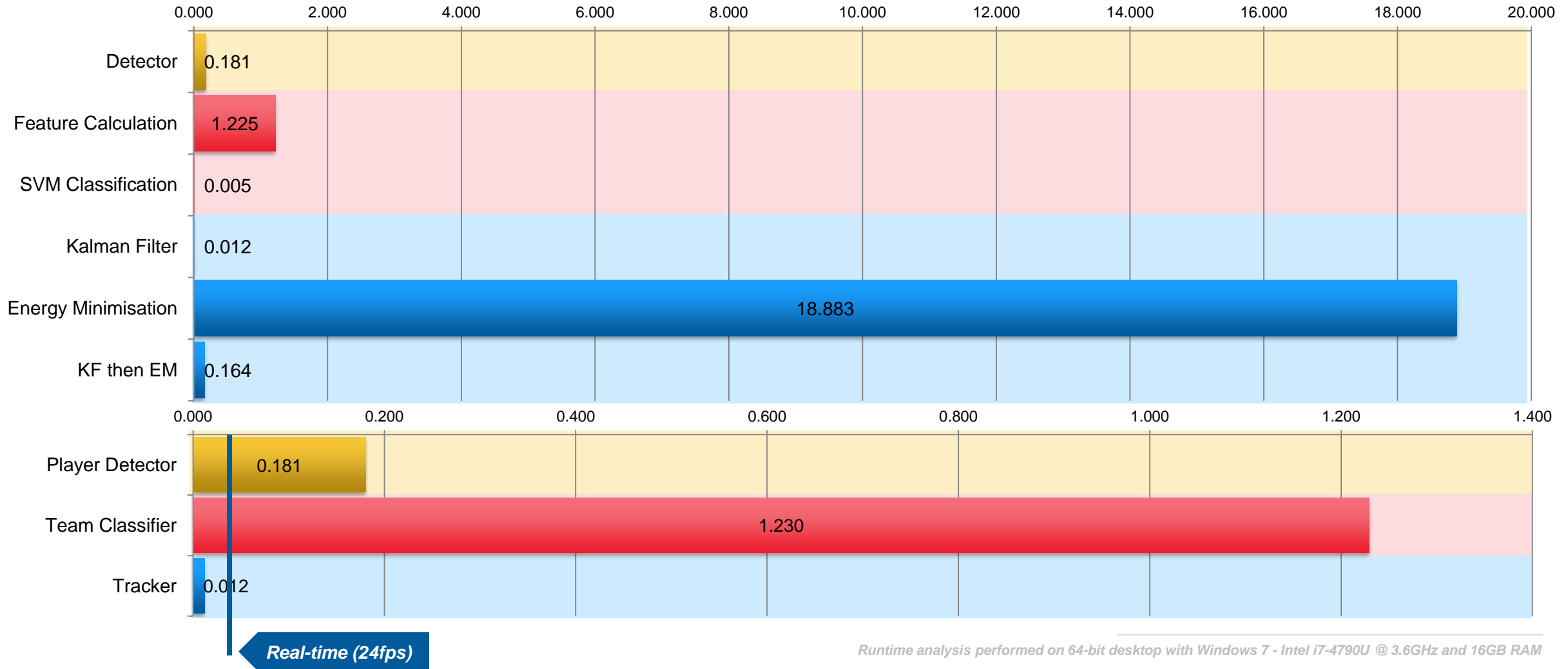
*Energy
Minimisation*



Combination

Runtime Analysis

- Time per single frame (seconds) assuming ~20 detections per frame



Conclusions and Further Work

- Pedestrian detection approaches inadequate
 - Vital to train AFL player detector, with AFL training data
- HSV Colour channels are sufficiently discriminative with weighted histograms
 - The SVM classifiers can separate teams with 90% MAP
 - However highly susceptible to lighting conditions, may need classifiers for particular lighting conditions
- The Kalman Filter tracker creates relatively accurate short tracks and is much faster than the Energy Minimisation approach
- Longer tracks can be captured with the use of the global Energy Minimisation approach, however requires Kalman Filter tracks as initializer to be feasible in both accuracy and runtime
- Improve runtimes, investigate other detection methods, different condition invariant features for team classification and more state-of-the-art tracking methods