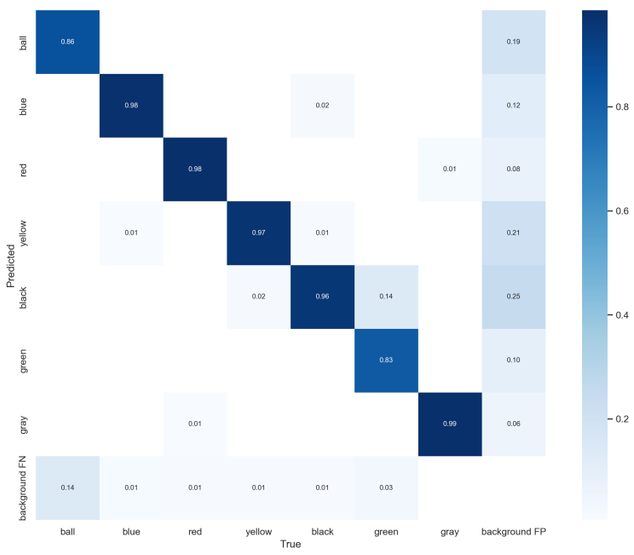
German Research Center for Artificial Intelligence

# Video Analysis Challenge

Thomas Röfer, Arne Hasselbring, Maximilian Wick, Simon Werner, Ayleen Lührsen, Lennart Heinbokel

## Object Detection

- Trained YOLOv5 [1] with an input resolution of 1920x1080 with bounding boxes and player colors
- Trained on 682 images from the data set provided by the SPL with a 70/20/10 split
- Detection of ball and robots works well, however there are some false positives in referees
- Detection of player colors depends on the lighting conditions being similar to the ones in the training set
- Inference takes ~55 ms per image on M1 Apple Neural Engine




*Confusion matrix of ball and player colors in the validation set*

## Extrinsic Camera Calibration

- Adapted code provided by the team Berlin United [2]
- Still relies on initial guess and working thresholds for the detection of green and white
- Calibration is determined before first playback of game video and cached for later uses

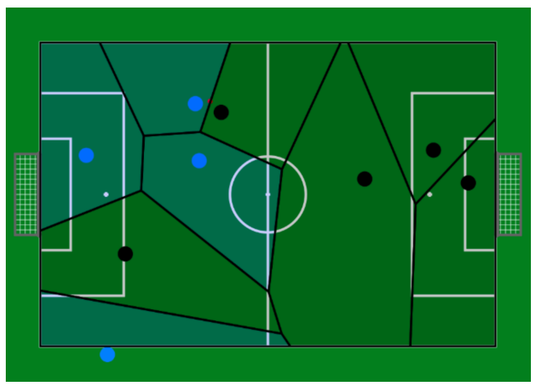

*Application showing the video of a half not in the training set with positions of ball and players (big circle indicates fallen robot). The statistics determined from the GameController log and from the video are shown next to the video.*

## World Model

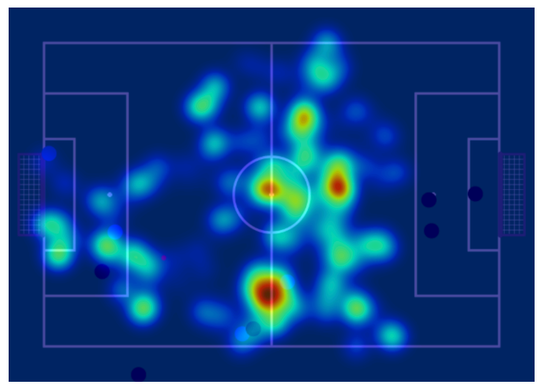
- Tracking the bounding boxes of players using a centroid multi object tracker [3] ignoring colors in YOLOv5's NMS
- Keeping the recent history of colors per track and assigning the most frequent one as player color
- Wrong ball detections are filtered out based on assumed maximum speed
- Centers of bounding boxes are projected into the world based on an assumed height (ball: 5 cm, players: 26 cm)
- If the projection of a player position back into the image is outside the bounding box, the player is assumed to be fallen

## Application

- Views: Game video, 2D field, ball position heat map
- Integrated GameController log to create statistics about game events while video is played back
- Creates statistics per team about fallen robots, distance walked, ball possession, distance ball moved, average ball distance to own goal, and average size of controlled areas from the video



*2D field view with controlled areas*



*Ball position heat map view*

## References

1. Glenn Jocher et al.:ultralytics/yolov5: v3.1 – Bug Fixes and Performance Improvements. Zenodo. https://doi.org/10.5281/zenodo.4154370
2. Benjamin Schlotter (2020): Analyse von RoboCup Spielen Erkennen und Lokalisieren von Nao Robotern. Studienarbeit. Humboldt-Universität zu Berlin. https://www.naoteamhumboldt.de/wp-content/papercite-data/pdf/2020_studienarbeit_schlotter.pdf
3. Aditya M. Deshpande (2021). Multi-object trackers in Python. https://adipandas.github.io/multi-object-tracker/

Universität Bremen

B-Human

**Contact:**
DFKI Bremen
Cyber-Physical Systems

Dr. Thomas Röfer
Phone: +49 (0) 421 218-64200
E-Mail: Thomas.Roefer@dfki.de
Website: www.b-human.de