

ABSTRACT

Basketball is a game where the movement of the players are very dynamic, that the difference between an event occurring and not occurring can be small and occasionally unnoticeable, and the sequence of frames of which the dribble is done matters tremendously in order to determine where a violation takes place. As such, a spatio-temporal action verification model is needed to assist referees in making decisions regarding whether or not a violation takes place, as it is almost instantaneous in nature. The challenge is to achieve balance between speed and accuracy, since the violation detection has to be fast enough for real-time use. Mediapipe Pose Estimation is chosen for its lightweight nature and capability for human joint feature extraction, and YOLO is trained to detect and extract basketball coordinates and person in order to pinpoint a region of interest for Mediapipe Pose Estimation in the case of failure to detect a player. The raw dataset consists of videos of basketball player(s) dribbling properly and committing a travelling violation. OpenCV is used for frame preprocessing, before passing each frame to the Mediapipe Pose and YOLO models for feature extraction. LSTM will be used to model the temporal dynamics of the movements the player(s) make, in the form of a binary classification as to whether or not the violation occurs. Hyperparameter tuning will be done on the amount of landmarks considered for feature extraction, and various hyperparameters of the LSTM model, such as the batch size, layers, number of units, optimiser, dropout rate, loss function, etc. Stacked LSTM architecture is found to be the most suitable model, at an accuracy of 0.8690. The resulting pipeline manages to run at at least 10 FPS, with enough overhead to theoretically achieve a higher framerate, after various methods of optimisations.

Key Words: Mediapipe, YOLO, LSTM, Action Verification, Basketball Travelling Violation

ABSTRAK

Basketball adalah permainan yang mana gerakan para pemain sangat dinamis, sehingga perbedaan antara terjadinya suatu peristiwa dan tidak terjadi bisa sangat kecil dan kadang-kadang sulit terlihat. Selain itu, urutan frame di mana dribble dilakukan sangat penting untuk menentukan di mana pelanggaran terjadi. Oleh karena itu, diperlukan sebuah model verifikasi aksi spasial-temporal untuk membantu wasit dalam membuat keputusan apakah pelanggaran terjadi atau tidak, karena kejadian ini hampir bersifat seketika. Tantangannya adalah mencapai keseimbangan antara kecepatan dan akurasi, karena deteksi pelanggaran harus cukup cepat untuk digunakan secara real-time. Mediapipe Pose Estimation dipilih karena sifatnya yang ringan dan kemampuannya untuk mengekstraksi fitur persendian manusia. Selain itu, YOLO dilatih untuk mendeteksi dan mengekstraksi koordinat bola basket dan manusia untuk menentukan daerah minat bagi Mediapipe Pose Estimation dalam kasus ketidakmampuan mendeteksi pemain. Dataset mentah terdiri dari video pemain basket yang melakukan dribble dengan benar dan melakukan pelanggaran perjalanan. OpenCV digunakan untuk preprocessing frame, sebelum setiap frame diberikan kepada model Mediapipe Pose dan YOLO untuk ekstraksi fitur. LSTM akan digunakan untuk memodelkan dinamika temporal gerakan yang dilakukan oleh pemain, dalam bentuk klasifikasi biner apakah pelanggaran terjadi atau tidak. Penyetelan hyperparameter akan dilakukan pada jumlah landmark yang dipertimbangkan untuk ekstraksi fitur, serta berbagai hyperparameter dari model LSTM seperti ukuran batch, jumlah layer, jumlah unit, optimiser, tingkat dropout, fungsi loss, dan lain-lain. Arsitektur Stacked LSTM ditemukan sebagai model yang paling cocok, dengan akurasi sebesar 0.8690. Pipeline yang dihasilkan mampu berjalan dengan setidaknya 10 FPS, dengan cukup overhead untuk secara teoritis mencapai frame rate yang lebih tinggi setelah berbagai metode optimisasi.

Kata Kunci: Mediapipe, YOLO, LSTM, Verifikasi Aksi, Pelanggaran Travelling dalam Bola Basket