

ABSTRAK

Bulu tangkis merupakan olahraga dengan kecepatan *shuttlecock* tertinggi di antara olahraga dengan bola lainnya, sehingga dalam pelacakan *shuttlecock* secara otomatis menjadi tantangan di bidang *computer vision*. TrackNetV3 adalah model deep learning berbasis CNN yang dirancang untuk melacak *shuttlecock* menggunakan informasi *input* dari beberapa frame berurutan (*sequence*). Tetapi, konfigurasi 8 frame *input* yang digunakan pada penelitian asli belum divalidasi pada dataset dengan karakteristik video yang lebih bervariasi, mencakup pertandingan kategori *Amateur* dan *Professional*. Penelitian ini bertujuan mengimplementasikan TrackNetV3 dengan variasi panjang *sequence input* yaitu 4 frame, 6 frame, 8 frame, dan 10 frame, serta menganalisis pengaruh terhadap akurasi pelacakan *shuttlecock*.

Penelitian menggunakan 201 video *rally* pertandingan bulu tangkis dengan pembagian data latih, validasi, dan pengujian sebesar 70:20:10. Pengujian dilakukan dengan empat skenario yang ditambahkan secara bertahap yaitu model dasar TrackNetV2, penambahan estimasi latar belakang, penambahan *Mixup augmentation*, dan penambahan *Rectification Module* berbasis InpaintNet. Evaluasi dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-Score*.

Hasil pengujian menunjukkan bahwa *sequence* 6-frame memberikan performa terbaik pada skenario model dasar TrackNetV2 dengan akurasi 91,21% dan *F1-Score* 94,71%. Penambahan estimasi latar belakang meningkatkan akurasi menjadi 92,16% dan *recall* menjadi 94,08% pada *sequence* 6-frame. Penambahan *Mixup augmentation* menggeser konfigurasi optimal ke *sequence* 8-frame dengan akurasi 92,12% dan *F1-Score* 95,30%, sekaligus menurunkan FP1 ke nilai terendah sebesar 0,53%. Penambahan *Rectification Module* menghasilkan *recall* tertinggi sebesar 94,15% pada *sequence* 4-frame, sedangkan *sequence* 8-frame menghasilkan TP tertinggi sebesar 80,26% dengan *recall* 94,11% dan *F1-Score* 95,11%, namun sedikit menurunkan *precision* menjadi 96,14% karena model menjadi lebih sensitif dalam mendeteksi *shuttlecock*. Secara keseluruhan, *sequence* 8-frame dengan kombinasi estimasi latar belakang, *Mixup augmentation*, dan *Rectification Module* menghasilkan performa paling seimbang antara kemampuan deteksi dan pengurangan kesalahan.

Kata kunci: bulu tangkis, *deep learning*, pelacakan *shuttlecock*, *sequence input*, TrackNetV3

ABSTRACT

Badminton is a sport with the highest shuttlecock speed among ball-based sports, making automatic shuttlecock tracking a significant challenge in the field of computer vision. TrackNetV3 is a CNN-based deep learning model designed to track shuttlecocks using temporal information from consecutive input frames (sequence). However, the 8-frame input configuration used in the original study has not been validated on datasets with more varied video characteristics, covering both Amateur and Professional match categories. This study aims to implement TrackNetV3 with input sequence length variations of 4-frame, 6-frame, 8-frame, and 10-frame, and analyze their effects on shuttlecock tracking accuracy.

The study used 201 badminton rally videos with a 70:20:10 split for training, validation, and testing. Experiments were conducted in four progressively added scenarios: baseline TrackNetV2, addition of background estimation, addition of Mixup augmentation, and addition of an InpaintNet-based Rectification Module. Evaluation was performed using accuracy, precision, recall, and F1-Score metrics.

Results show that the 6-frame sequence achieved the best performance in the baseline TrackNetV2 scenario with an accuracy of 91.21% and F1-Score of 94.71%. The addition of background estimation improved accuracy to 92.16% and recall to 94.08% on the 6-frame sequence. The addition of Mixup augmentation shifted the optimal configuration to the 8-frame sequence with an accuracy of 92.12% and F1-Score of 95.30%, while reducing FPI to its lowest value of 0.53%. The addition of the Rectification Module achieved the highest recall of 94.15% on the 4-frame sequence, while the 8-frame sequence achieved the highest TP of 80.26% with a recall of 94.11% and F1-Score of 95.11%, though precision slightly decreased to 96.14% as the model became more sensitive in detecting the shuttlecock. Overall, the 8-frame sequence combined with background estimation, Mixup augmentation, and Rectification Module produced the most balanced performance between detection capability and error reduction.

Keywords: badminton, deep learning, shuttlecock tracking, input sequence, TrackNetV3