

ABSTRACT

Eye tracking technology has become a focal point in research and industrial applications, especially in the field of assistive technology. Gaze-based interaction enables innovative assistive technologies for individuals with motor impairments. Recent advances in deep learning can identify eye movements—fixation, saccade, and smooth pursuit—from eye tracking data, providing reliable inputs for gaze-based interaction, especially for smooth pursuit. Various deep learning methods are widely used for eye movement classification. However, several issues remain, such as inaccuracies in classifying smooth pursuit with an F1 score of 76.2%. This inaccuracy leads to errors in detecting smooth pursuit movements, affecting the reliability of gaze-based interaction. Smooth pursuit eye movement is crucial for calibration-free eye tracking applications as it allows for smooth tracking of moving objects, creating natural, intuitive, and efficient interactions. This study aims to address these challenges by improving the classification of eye movements, particularly in detecting smooth pursuit.

To tackle these issues, this research proposes a deep learning model based on a 1D-CNN-Transformer model with Hyperband parameter optimization for eye movement classification. The use of 1D-CNN-Transformer enables more effective extraction of temporal and spatial features compared to other deep learning models, while Hyperband optimization enhances efficiency and accuracy in hyperparameter selection. This study utilizes two types of datasets, GazeCom and HMR, due to their large data volumes and their nature as open datasets.

Using K-Fold Cross Validation, the method in this study improved classification of eye movements with F1 scores of 95.53%, 92.84%, 82.89%, and 83.84% for fixation, saccade, smooth pursuit, and noise, respectively, on the GazeCom dataset. For the HMR dataset, the F1 scores achieved were 97.58%, 92.93%, 97.27%, and 93.35% for fixation, saccade, smooth pursuit, and noise, respectively.

These achievements represent a significant step forward in eye movement classification, particularly in the context of integration with gaze-based input. Improved performance in eye movement classification has a direct and significant impact on assistive technology, enabling users to control devices more accurately and efficiently using their gaze, thus enhancing accessibility and independence in various daily activities.

Keywords : 1D-CNN-Transformer, deep learning, hyperband, hyperparameters tuning, eye tracking

INTISARI

Teknologi *eye tracking* telah menjadi sorotan dalam penelitian dan aplikasi industri, terutama dalam bidang *assistive technology*. Interaksi berbasis tatapan mata membuka peluang pengembangan *assistive technology* yang inovatif bagi individu dengan gangguan gerak. Pendekatan *deep learning* terbaru dapat mengidentifikasi gerakan mata seperti *fixation*, *saccade*, dan *smooth pursuit* dari data gerakan mata, yang menjadi *input* andal untuk interaksi berbasis tatapan mata, terutama pada gerakan *smooth pursuit*. Berbagai metode *deep learning* banyak digunakan untuk klasifikasi gerakan mata. Namun, beberapa masalah masih ada, seperti ketidakakuratan dalam mengklasifikasikan *smooth pursuit* dengan *F1 score* sebesar 76.2%. Akibat ketidakakuratan ini, terjadi kesalahan dalam mendeteksi gerakan *smooth pursuit*, yang dapat mempengaruhi keandalan interaksi berbasis tatapan mata.

Untuk mengatasi permasalahan tersebut, penelitian ini mengusulkan model *deep learning* berbasis model 1D-CNN-Transformer dengan optimasi nilai parameter *Hyperband* dalam melakukan klasifikasi gerakan mata. Penggunaan 1D-CNN-Transformer memungkinkan ekstraksi fitur *temporal* dan *spatial* yang lebih efektif dibandingkan model *deep learning* lainnya, sementara optimasi *Hyperband* meningkatkan efisiensi dan akurasi dalam pemilihan *hyperparameter*. Penelitian ini menggunakan dua jenis dataset, yaitu dataset GazeCom dan dataset HMR, karena jumlah data yang besar dan sifatnya sebagai *open dataset*.

Dengan menggunakan *K-Fold Cross Validation*, metode dalam penelitian ini meningkatkan klasifikasi gerakan mata dengan *F1 score* sebesar 95.53%, 92.84%, 82.89%, dan 83.84% untuk *fixation*, *saccade*, *smooth pursuit*, dan *noise*, secara berurutan, pada dataset GazeCom. Untuk dataset HMR, *F1 score* yang dicapai adalah 97.58%, 92.93%, 97.27%, dan 93.35% untuk *fixation*, *saccade*, *smooth pursuit*, dan *noise*, secara berurutan.

Pencapaian yang dihasilkan merupakan langkah penting dalam mengklasifikasi gerakan mata, terutama dalam konteks integrasi dengan input berbasis tatapan. Peningkatan performa klasifikasi gerakan mata memiliki dampak langsung yang signifikan pada *assistive technology* dengan memungkinkan pengguna yang membutuhkan untuk mengendalikan perangkat dengan lebih akurat dan efisien menggunakan tatapan mata mereka, meningkatkan aksesibilitas dan kemandirian dalam berbagai aktivitas sehari-hari.

Kata kunci – 1D-CNN-Transformer, *deep learning*, *hyperband*, *hyperparameters tuning*, *eye tracking*