

ABSTRAK

Perkembangan teknologi *deep learning* telah mendorong pemanfaatan *gaze estimation* sebagai komponen penting dalam sistem *human-computer interaction* (HCI). Namun, penerapan model *gaze estimation* berbasis *deep learning* pada lingkungan nyata masih menghadapi tantangan besar, khususnya ketika dijalankan pada platform *edge computing* yang memiliki keterbatasan sumber daya komputasi. Oleh karena itu, penelitian ini bertujuan untuk menerapkan model EfficientNet-B0 sebagai *backbone gaze estimation* pada platform *edge computing* dengan tetap mempertahankan akurasi prediksi dan efisiensi inferensi.

Penelitian ini menggunakan dataset MPIIGaze *normalized subset* yang terdiri dari lebih dari 213.000 pasangan citra mata atau sekitar 427.316 citra mata setelah penggabungan data mata kanan dan kiri, sebagai data pelatihan dan pengujian. Metodologi yang digunakan mengacu pada kerangka kerja CRISP-DM, dengan tahapan meliputi *business understanding*, *data understanding*, *data preparation*, *modeling*, dan *Evaluation*. Proses *preprocessing* mencakup *resize* citra ke ukuran 36×60 piksel, normalisasi piksel, serta augmentasi data berupa rotasi ringan dan *brightness adjustment*. Model dilatih menggunakan tiga skenario jumlah *epoch* (30, 50, dan 70 *epoch*) dengan *optimizer* Adam dan fungsi loss *Mean Squared Error* (MSE) untuk menentukan konfigurasi pelatihan terbaik.

Hasil penelitian menunjukkan bahwa model EfficientNet-B0 dengan pelatihan 70 *epoch* menghasilkan performa terbaik dengan *Mean Angular Error* sebesar 1,46°, *Validation loss* sebesar 0,000493, *Average Inference Time* sebesar 0,000093 detik, dan kecepatan pemrosesan mencapai 10.803,74 FPS. Hasil tersebut menunjukkan bahwa model EfficientNet-B0 mampu memberikan keseimbangan yang baik antara akurasi prediksi arah pandangan dan efisiensi komputasi, sehingga layak diterapkan pada platform *edge computing* untuk mendukung aplikasi HCI berbasis *real-time*.

Kata kunci: *gaze estimation*, EfficientNet, *edge computing*, *deep learning*, analisis kelayakan implementasi.

ABSTRACT

The advancement of deep learning technology has driven the utilization of gaze estimation as a critical component in human–computer interaction (HCI) systems. However, deploying deep learning-based gaze estimation models in real-world environments remains challenging, particularly when executed on edge computing platforms with limited computational resources. Therefore, this study aims to apply the EfficientNet-B0 model as a gaze estimation backbone on an edge computing platform while maintaining prediction accuracy and inference efficiency. This study utilized the MPIIGaze normalized subset dataset comprising more than 213,000 eye image pairs, or approximately 427,316 eye images after combining left and right eye data, for training and Evaluation. The research methodology followed the CRISP-DM framework, encompassing business understanding, data understanding, data preparation, modeling, and Evaluation phases. The preprocessing pipeline included image resizing to 36×60 pixels, pixel normalization, and data augmentation techniques such as slight rotation and brightness adjustment. The model was trained under three epoch configurations (30, 50, and 70 epochs) using the Adam optimizer with Mean Squared Error (MSE) as the loss function to identify the optimal training configuration. The results demonstrate that the EfficientNet-B0 model trained for 70 epochs achieved the best performance, with a Mean Angular Error of 1.46°, a Validation loss of 0.000493, an Average Inference Time of 0.000093 seconds, and a processing speed of 10,803.74 FPS. These findings indicate that EfficientNet-B0 successfully delivers a strong balance between gaze prediction accuracy and computational efficiency, making it well-suited for implementation on edge computing platforms to support real-time HCI applications.

Keywords: gaze estimation, EfficientNet, edge computing, deep learning, implementation feasibility.