

INTISARI

ARSITEKTUR HYBRID DENGAN VISION TRANSFORMER DAN CNN UNTUK KLASIFIKASI RETINOPATI DIABETIK

Oleh

Ariefan Dipokusumo Wibowo

23/525137/PPA/06592

Retinopati diabetik (RD) merupakan salah satu penyebab utama kebutaan di dunia yang memerlukan deteksi dini dan klasifikasi yang akurat. Penelitian ini mengembangkan model *deep learning* hibrida yang mengintegrasikan *Convolutional Neural Network* (CNN) dan *Vision Transformer* (ViT) untuk klasifikasi RD. Model hibrida ini menggabungkan keunggulan CNN dalam ekstraksi fitur spasial dengan mekanisme *global attention* dari ViT untuk meningkatkan akurasi dan keandalan klasifikasi.

Pengujian dilakukan menggunakan dataset APTOS 2019 yang telah melalui tahap *preprocessing*, termasuk augmentasi data dan penyeimbangan kelas untuk mengatasi distribusi data yang tidak seimbang. Hasil penelitian menunjukkan bahwa model hibrida memberikan performa terbaik dengan akurasi sebesar 84,00%, *macro average* F1-score 67,87%, dan *weighted average* F1-score 83,08%. Model ini juga unggul dalam mendeteksi kelas Non-DR dengan F1-score di atas 96,00%, serta memiliki *recall* tertinggi pada kelas *Moderate DR* (90,00%) dan F1-score terbaik pada kelas *Severe DR* (40,00%). Namun, terdapat tantangan pada kelas dengan representasi minoritas, seperti *Mild DR* dan *Proliferative DR*, yang masih perlu ditingkatkan.

Dengan hasil ini, model hibrida yang dikembangkan diharapkan dapat berkontribusi dalam deteksi dini retinopati diabetik serta memberikan solusi yang lebih efektif untuk aplikasi di bidang medis.

Kata-kata kunci : retinopati diabetik, model hibrida, convolutional neural network, vision transformer, klasifikasi citra, deep learning, citra medis, class imbalance.

ABSTRACT

HYBRID ARCHITECTURE WITH VISION TRANSFORMER AND CNN FOR CLASSIFICATION OF DIABETIC RETINOPATHY

By

Ariefan Dipokusumo Wibowo

23/525137/PPA/06592

Diabetic retinopathy (DR) is one of the leading causes of blindness worldwide, requiring early detection and accurate classification. This study develops a hybrid *deep learning* model that integrates *Convolutional Neural Network* (CNN) and *Vision Transformer* (ViT) for DR classification. The hybrid model combines the strengths of CNN in spatial feature extraction with the *global attention* mechanism of ViT to improve classification accuracy and reliability.

The evaluation was conducted using the APTOS 2019 dataset, which underwent *preprocessing* steps, including data augmentation and class balancing, to address the issue of imbalanced data distribution. The results indicate that the hybrid model achieved the best performance, with an accuracy of 84,00%, a *macro average* F1-score of 67,87%, and a *weighted average* F1-score of 83,08%. The model also excelled in detecting the Non-DR class with an F1-score above 96,00%, and it achieved the highest *recall* for the *Moderate DR* class (90,00%) and the best F1-score for the *Severe DR* class (40,00%). However, challenges remain in detecting minority-represented classes, such as *Mild DR* and *Proliferative DR*, which require further improvement.

With these results, the proposed hybrid model is expected to contribute to the early detection of diabetic retinopathy and provide a more effective solution for medical applications.

Keywords : diabetic retinopathy, hybrid model, convolutional neural network, vision transformer, image classification, deep learning, medical imaging, class imbalance.