ABSTRAK

Hidayat, Acep Taufik (2025), Perbandingan Model ResNet50 dan EfficientNetB0 dengan Augmentasi GANs untuk Klasifikasi Tingkat Keparahan Retinopati Diabetes

Kata Kunci: Retinopati Diabetis, *ResNet50*, *EfficientNetB0*, Augmentasi Data, *Generative Adversarial Networks (GANs),* Klasifikasi Citra

Retinopati Diabetik (DR) adalah komplikasi diabetes utama penyebab kebutaan, sehingga deteksi dini dan klasifikasinya krusial. Tantangan diagnosis otomatis terletak pada keterbatasan dan ketidakseimbangan dataset citra medis. Penelitian ini membandingkan *ResNet50* dan *EfficientNetB0* untuk klasifikasi lima tingkat keparahan DR (*Normal*, *Mild*, *Moderate*, *Severe*, *Proliferative DR*) menggunakan dataset IDRiD, yang diperkaya dengan augmentasi *Generative Adversarial Networks* (*GANs*). Augmentasi GANs berhasil mengatasi ketidakseimbangan data, meningkatkan kuantitas dan keseimbangan dataset training secara signifikan. *EfficientNetB0* secara konsisten mengungguli *ResNet50*. Pada dataset asli, *EfficientNetB0* mencapai akurasi testing 47% (presisi 47%, *recall* 47%, *f1-score* 45%), sedangkan *ResNet50* 41% (presisi 45%, *recall* 41%, *f1-score* 38%). Penggunaan data sintetis saja menghasilkan akurasi *testing* sangat rendah (23%) untuk kedua model. Namun, kombinasi data asli dan sintetis meningkatkan generalisasi model pada data *testing*: *ResNet50* mencapai 44% (presisi 49%, *recall* 44%, *f1-score* 42%), dan *EfficientNetB0* mencapai 47% (presisi 50%, *recall* 47%, *f1-score* 45%), dibandingkan dataset asli saja. Meskipun akurasi testing tertinggi 47% (*EfficientNetB0*) masih *sub-optimal*, hal ini disebabkan kompleksitas klasifikasi multikelas DR dengan ambiguitas visual (terutama *Mild NPDR* dan *Moderate NPDR* yang memiliki *f1-score* rendah, 0.21 dan 0.40) serta keterbatasan ukuran *testing* set asli (103 citra) yang memicu *overfitting*. Penelitian ini menegaskan potensi augmentasi GANs sebagai suplemen berharga untuk dataset medis yang terbatas, menunjukkan peningkatan kinerja signifikan, dan mengidentifikasi *EfficientNetB0* sebagai model yang lebih unggul. Hasil ini menjadi fondasi penting untuk diagnosis DR otomatis yang lebih baik, dengan fokus pada optimasi data sintetis dan penanganan *overfitting* di masa depan.

ABSTRACT

Hidayat, Acep Taufik (2025), *Comparison of ResNet50 and EfficientNetB0 Models with GANs Augmentation for Classification of Diabetic Retinopathy Severity Level*

Keywords*: Diabetic Retinopathy, ResNet50, EfficientNetB0, Data Augmentation, Generative Adversarial Networks (GANs), Image Classification*

Diabetic Retinopathy (DR) is a major diabetes-related complication and a leading cause of blindness, making early detection and classification crucial. However, automated diagnosis faces challenges due to limited and imbalanced medical image datasets. This study compares ResNet50 and EfficientNetB0 for classifying five DR severity levels (Normal, Mild, Moderate, Severe, Proliferative DR) using the IDRiD dataset, enriched with Generative Adversarial Networks (GANs)-based augmentation. The GAN-based augmentation effectively addressed data imbalance by significantly increasing the quantity and balance of the training dataset. EfficientNetB0 consistently outperformed ResNet50. On the original dataset, EfficientNetB0 achieved a testing accuracy of 47% (precision: 47%, recall: 47%, F1-score: 45%), while ResNet50 achieved 41% (precision: 45%, recall: 41%, F1-score: 38%). Using synthetic data alone resulted in very low testing accuracy (23%) for both models. However, combining original and synthetic data improved model generalization on the testing set: ResNet50 reached 44% (precision: 49%, recall: 44%, F1-score: 42%), and EfficientNetB0 achieved 47% (precision: 50%, recall: 47%, F1-score: 45%), compared to using the original dataset alone. Although the highest testing accuracy (47% with EfficientNetB0) remains sub-optimal, this is attributed to the complexity of multiclass DR classification with visually ambiguous cases (especially Mild NPDR and Moderate NPDR, which had low F1-scores of 0.21 and 0.40) and the limited size of the original test set (103 images), which increases the risk of overfitting. This study underscores the potential of GAN-based augmentation as a valuable supplement for limited medical datasets, demonstrates significant performance improvements, and identifies EfficientNetB0 as a more effective model. These findings provide a critical foundation for improving automated DR diagnosis, with future work focusing on optimizing synthetic data and addressing overfitting.