UNIVERSITAS
GADJAH MADA
Integrasi Convolutional Block Attention Module dan Auxiliary Segmentation Branch pada ResNet-50 untuk Klasifikasi Biner Microaneurysm dan Lokalisasi Class Activation Map pada Citra Fundus
Muhammad Dzaki Zhafiri, Anugerah Galang Persada, S.T., M.Eng.; Prof. Ir. Hanung Adi Nugroho, S.T., M.Eng., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

INTISARI

Diabetic retinopathy (DR) merupakan komplikasi mikrovaskular diabetes yang dapat menyebabkan kebutaan, dan *microaneurysm* (MA) adalah tanda klinis paling awal dari penyakit ini. Sistem klasifikasi DR berbasis *deep learning* telah menunjukkan performa yang menjanjikan, namun kemampuan model untuk secara konsisten mengarahkan perhatiannya ke lokasi lesi yang relevan masih belum terverifikasi secara kuantitatif, khususnya untuk lesi mikro seperti MA. Penelitian ini menganalisis pengaruh penambahan *Convolutional Block Attention Module* (CBAM) dan *auxiliary segmentation branch* (AuxSeg) terhadap dua aspek secara simultan, yakni performa klasifikasi biner MA dan kualitas lokalisasi *heatmap Class Activation Map* (CAM) pada arsitektur ResNet-50.

Empat skenario arsitektur dievaluasi (*Baseline*, +CBAM, +AuxSeg, +AuxSeg+CBAM) menggunakan dataset FGADR Seg-set (1.842 citra fundus dengan anotasi *pixel-level* MA) pada tiga resolusi masukan (224×224, 512×512, 1024×1024). Kualitas lokalisasi diukur menggunakan tujuh metode CAM dari tiga paradigma berbeda (*gradient-based*, dekomposisi spektral, dan hibrida) terhadap *ground truth* MA *mask*, dengan metrik MIoU, *Pointing Game*, dan *Pixel AUC-ROC*.

Hasil menunjukkan bahwa AuxSeg meningkatkan seluruh metrik secara konsisten: F1-Score naik dari 0,8266 menjadi 0,8684, Cohen's Kappa naik 43,3%, dan MIoU rata-rata meningkat 42,7% di seluruh tujuh metode CAM pada resolusi 1024. Sebaliknya, CBAM *standalone* menurunkan kualitas lokalisasi secara drastis (MIoU $-65{,}8\%$). Hasil ini konsisten dengan hipotesis bahwa mekanisme *channel attention* menekan fitur MA yang berukuran sangat kecil sehingga secara statistik menyerupai *noise* pada aktivasi *channel*, meskipun mekanisme tersebut tidak diverifikasi secara langsung dalam penelitian ini. Ketika dikombinasikan dengan AuxSeg, dampak destruktif CBAM terhadap lokalisasi dapat dimitigasi (MIoU $+2{,}9\%$), meskipun interaksi terhadap klasifikasi tetap destruktif (Kappa $-37{,}7\%$). Resolusi terbukti menjadi faktor kritis: *Pointing Game* bernilai 0% pada resolusi 224 di seluruh model, menunjukkan fenomena kebutaan spasial total, sementara peningkatan MIoU dari resolusi 224 ke 1024 mencapai 336–1048%. Metode CAM berbasis non-gradien (spektral dan hibrída) yang belum pernah dievaluasi secara kuantitatif pada domain DR mendapat manfaat terbesar dari AuxSeg ($+$45–77% MIoU), mengonfirmasi bahwa perbaikan bersifat fundamental pada representasi fitur *backbone*.

Penelitian ini menyimpulkan bahwa supervisi spasial eksplisit melalui AuxSeg secara konsisten memperbaiki klasifikasi dan lokalisasi MA, sedangkan mekanisme *attention* tanpa panduan spasial justru kontraproduktif untuk lesi mikro. Temuan ini memberikan panduan desain arsitektur bagi pengembangan sistem diagnosis DR yang tidak hanya akurat tetapi juga *interpretable*. Seluruh temuan diperoleh dari satu dataset (FGADR Seg-set) dengan pelatihan *single-seed*; konfirmasi signifikansi statistik melalui validasi multi-*seed* dan validasi eksternal lintas-dataset menjadi arah pengembangan lanjutan.

Keywords: *diabetic retinopathy, class activation map, microaneurysm, convolutional block attention module, deep learning*

UNIVERSITAS
GADJAH MADA
Integrasi Convolutional Block Attention Module dan Auxiliary Segmentation Branch pada ResNet-50 untuk Klasifikasi Biner Microaneurysm dan Lokalisasi Class Activation Map pada Citra Fundus
Muhammad Dzaki Zhafiri, Anugerah Galang Persada, S.T., M.Eng.; Prof. Ir. Hanung Adi Nugroho, S.T., M.Eng., Ph.D
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

Diabetic retinopathy (DR) is a microvascular complication of diabetes that can lead to blindness, and microaneurysms (MA) are the earliest clinical sign of the disease. Deep learning-based DR classification systems have shown promising performance, yet whether models consistently attend to relevant lesion locations remains quantitatively unverified, particularly for micro-lesions such as MA. This study analyzes the effects of adding a Convolutional Block Attention Module (CBAM) and an auxiliary segmentation branch (AuxSeg) on two aspects simultaneously, namely binary MA classification performance and Class Activation Map (CAM) heatmap localization quality on a ResNet-50 architecture.

Four architectural scenarios (Baseline, +CBAM, +AuxSeg, +AuxSeg+CBAM) were evaluated on the FGADR Seg-set dataset (1,842 fundus images with pixel-level MA annotations) at three input resolutions (224×224, 512×512, 1024×1024). Localization quality was assessed using seven CAM methods from three distinct paradigms (gradient-based, spectral decomposition, and hybrid) against ground truth MA masks, measured by MIoU, Pointing Game, and Pixel AUC-ROC.

Results show that AuxSeg consistently improved all metrics: F1-Score increased from 0.8266 to 0.8684, Cohen's Kappa rose by 43.3%, and mean MIoU improved by 42.7% across all seven CAM methods at 1024 resolution. Conversely, standalone CBAM drastically degraded localization quality (MIoU -65.8%). This result is consistent with the hypothesis that the channel attention mechanism suppresses the very small MA features that statistically resemble noise in channel activations, although this mechanism was not directly verified in this study. When combined with AuxSeg, CBAM's destructive effect on localization was mitigated (MIoU $+2.9\%$), although the interaction remained destructive for classification (Kappa -37.7%). Resolution proved critical: Pointing Game was 0% at 224 resolution across all models, indicating complete spatial blindness, while MIoU improvement from 224 to 1024 resolution ranged from 336% to 1048%. Non-gradient CAM methods (spectral-based and hybrid-based), which had never been quantitatively evaluated in the DR domain, benefited most from AuxSeg ($+45$–77% MIoU), confirming that improvements were fundamental to backbone feature representations.

This study concludes that explicit spatial supervision through AuxSeg consistently improves both MA classification and localization, whereas attention mechanisms without spatial guidance are counterproductive for micro-lesions. These findings provide architectural design guidelines for developing DR diagnostic systems that are not only accurate but also interpretable. All findings were obtained from a single dataset (FGADR Seg-set) using single-seed training; confirming statistical significance through multi-seed validation and external cross-dataset validation remains future work.

Keywords: *diabetic retinopathy, class activation map, microaneurysm, convolutional block attention module, deep learning*