UNIVERSITAS
GADJAH MADA
Pengembangan Metode Complementary-Stream Sum Fusion Untuk Segmentasi Semantik Pada Citra Relief Candi Borobudur
Lukman Awaludin, Prof. Drs. Agus Harjoko, M.Sc., Ph.D.; Wahyono, S.Kom., Ph.D.; Dr. Dra. Niken Wirasanti, M.Si.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

INTISARI

Pengembangan Metode Complementary-Stream Sum Fusion Untuk Segmentasi Semantik Pada Citra Relief Candi Borobudur

Lukman Awaludin
21/484857/SPA/00803

Relief candi merupakan salah satu warisan budaya yang memiliki nilai historis tinggi. Proses segmentasi semantik pada relief menghadapi tantangan akibat kompleksitas latar belakang, pencahayaan tidak merata, degradasi material, serta ambiguitas batas visual antara objek dan latar belakang. Penelitian ini mengusulkan metode Complementary-Stream Sum Fusion (CSSF), sebuah pendekatan segmentasi semantik berbasis *deep learning* yang menggabungkan informasi tekstur permukaan (RGB-D) dengan representasi fitur spasial dan struktural (Edge-Depth) melalui mekanisme fusi.

Dataset penelitian terdiri atas 248 panel relief Candi Borobudur yang diklasifikasikan ke dalam tujuh kelas semantik, yaitu human, animal, plant, building, artifact, decoration, dan Mythology (god/demon). Tahap pertama mengevaluasi lima arsitektur CNN (DeepLabV3+, PSPNet, SegNet, DenseUNet, U-Net) untuk segmentasi RGBD 4-kanal, dengan DeepLabV3+ terpilih sebagai *backbone* terbaik (test mIoU 0,4866). Tahap kedua membangun arsitektur dual-branch, RGBD Branch (4-kanal: R, G, B, Depth) mencapai test mIoU 0,5034 dan Edge-Depth Branch (2-kanal: *Soft-Edge* berbasis Difference of Gaussian, Depth) mencapai test mIoU 0,4251.

Hasil eksperimen pada 56 citra uji menunjukkan bahwa CSSF memperoleh kinerja terbaik dengan mIoU sebesar 0,5220. Studi ablasi membuktikan kontribusi setiap komponen: penambahan depth meningkatkan mIoU sebesar 0,22 poin persentase dibandingkan RGB saja, sedangkan FocalDiceLoss memberikan peningkatan tambahan sebesar 0,12 poin persentase dibandingkan Cross-Entropy Loss. Perbandingan dengan 11 metode *state-of-the-art* menunjukkan bahwa CSSF secara konsisten mengungguli seluruh metode pembanding pada dataset BRSD, sehingga membuktikan bahwa pendekatan complementary-stream sum fusion lebih unggul dibandingkan per-class adaptive fusion.

Kata Kunci: Segmentasi Semantik, Relief Candi, DeepLabV3+, Complementary-Stream Sum Fusion

UNIVERSITAS
GADJAH MADA
Pengembangan Metode Complementary-Stream Sum Fusion Untuk Segmentasi Semantik Pada Citra Relief Candi Borobudur
Lukman Awaludin, Prof. Drs. Agus Harjoko, M.Sc., Ph.D.; Wahyono, S.Kom., Ph.D.; Dr. Dra. Niken Wirasanti, M.Si.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

Development of the Complementary-Stream Sum Fusion Method for Semantic Segmentation of Borobudur Temple Bas-Relief Images

Lukman Awaludin
21/484857/SPA/00803

Temple relief is a cultural heritage of high historical value. Semantic segmentation of reliefs faces challenges due to background complexity, uneven illumination, material degradation, and visual boundary ambiguity between objects and background. This study proposes Complementary-Stream Sum Fusion (CSSF), a deep learning-based semantic segmentation method that combines surface texture information (RGB-D) with spatial and structural feature representations (Edge-Depth) through a per-class adaptive fusion mechanism using the Expectation-Maximization (EM) algorithm.

The dataset comprises 248 relief panels from Borobudur Temple classified into eight semantic classes, namely background, human, animal, plant, building, artifact, decoration, and god/demon. The first stage evaluates five CNN architectures (DeepLabV3+, PSPNet, SegNet, DenseUNet, U-Net) for 4-channel RGBD segmentation, with DeepLabV3+ selected as the best backbone (test mIoU 0.4866). The second stage constructs a dual-branch architecture: the RGBD Branch (4-channel: R, G, B, Depth) achieves a test mIoU of 0.5034, while the Edge-Depth Branch (2-channel: Soft-Edge based on Difference of Gaussian, Depth) achieves a test mIoU of 0.4251.

Experimental results on 56 test images demonstrate that CSSF achieves the highest performance, with an mIoU of 0.5220. Ablation studies confirm the effectiveness of each component: adding depth improves the mIoU by 0.22 percentage points over RGB alone, while FocalDiceLoss provides an additional gain of 0.12 percentage points compared with Cross-Entropy Loss. Comparisons with 11 state-of-the-art methods further demonstrate that CSSF consistently outperforms all baselines on the BRSD dataset, validating the performance advantage of complementary-stream sum fusion over per-class adaptive fusion.

Keywords: Semantic Segmentation, Temple Relief, DeepLabV3+, Complementary-Stream Sum Fusion