

Laporan Preprocessing Dataset Flood Area Segmentation

1. Pendahuluan

Dataset Flood Area Segmentation memiliki karakteristik khusus: variasi warna air, refleksi, bayangan, dan perbedaan resolusi citra drone. Preprocessing diperlukan agar model segmentasi memperoleh input konsisten dan mampu mengenali boundary banjir yang sering blur dan low-contrast.

2. Standardisasi Resolusi

Semua citra diubah ke resolusi 256x256 menggunakan interpolasi bilinear, sedangkan mask menggunakan nearest-neighbor. Hal ini diperlukan karena dataset memiliki resolusi beragam dan U-Net membutuhkan ukuran input tetap. Resize juga menjaga konsistensi spasial untuk boundary air.

3. Normalisasi Piksel

Nilai piksel dinormalisasi ke rentang 0–1 agar stabil dalam proses training CNN. Dataset ini memiliki perbedaan brightness signifikan antar lokasi, sehingga normalisasi membantu menstabilkan distribusi intensitas.

4. CLAHE (Contrast Limited Adaptive Histogram Equalization)

CLAHE diterapkan pada kanal luminance untuk mengurangi pengaruh shadow dan highlight pada permukaan air. Pada dataset ini, banyak citra menunjukkan pantulan cahaya kuat dan area gelap yang mengganggu deteksi boundary.

5. Gamma Correction

Gamma correction dengan rentang 0.8–1.2 membantu menormalkan citra yang terlalu terang atau terlalu gelap. Variasi cahaya pada dataset banjir sangat tinggi, sehingga teknik ini meningkatkan robustness model.

6. Edge-aware Sharpening

Unsharp masking ringan diterapkan untuk mempertegas boundary air–tanah. Dataset banjir sering memiliki batas air yang blur dan tidak jelas, sehingga sharpening membantu modul FBRM dalam menangkap edge dengan lebih baik.

7. Augmentasi Domain-Specific

Augmentasi yang digunakan mencerminkan kondisi banjir nyata: - Random horizontal flip: aman karena orientasi tidak memengaruhi struktur air. - Random brightness & contrast: meniru kondisi cuaca yang berubah-ubah. - Gaussian blur: meniru blur akibat pergerakan drone. - Random shadow: meniru bayangan dari bangunan atau pepohonan. - Color jitter: meniru variasi warna air (keruh, jernih, coklat, gelap). Augmentasi ini sangat penting karena dataset memperlihatkan variasi

kondisi sangat ekstrem.

8. Pemrosesan Mask

Mask hanya dinormalisasi ke 0 dan 1 serta di-resize menggunakan nearest-neighbor. Penting untuk menjaga boundary agar tidak blur karena tujuan utama adalah segmentasi tepi air secara presisi.

9. Kesimpulan

Pipeline preprocessing ini dirancang secara khusus untuk dataset banjir yang memiliki masalah utama berupa boundary blur, variasi warna air, shadow, dan highlight. Semua langkah disusun untuk mendukung modul FBRM yang sensitif terhadap edge. Pipeline ini meningkatkan konsistensi data dan performa model segmentasi banjir.