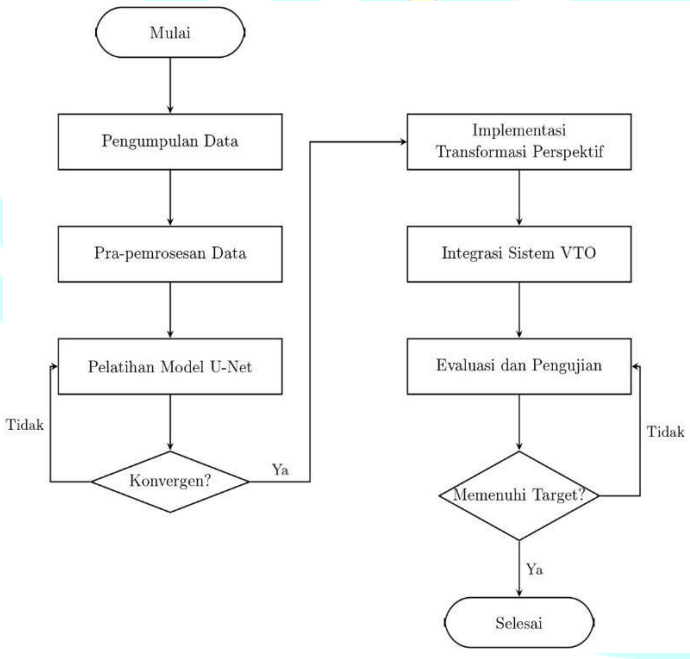
BAB III
METODE PENELITIAN

3.1 Objek Penelitian

Objek dalam penelitian ini adalah tampilan visual dari area lantai pada citra ruangan interior yang akan diintegrasikan dengan produk keramik lantai PT Concord Industry. Penelitian ini memfokuskan pengamatan pada karakteristik geometri lantai dan tekstur permukaan keramik, baik tipe *matte* maupun *glossy*. Fokus utama diletakkan pada pemetaan motif keramik yang presisi sesuai dengan perspektif ruang serta pemisahan area lantai dari elemen interior lainnya (seperti furnitur atau dinding) untuk menghasilkan visualisasi *Virtual Try-On* yang realistis.

3.2 Alur Pernelitian

Penelitian ini mengikuti alur metodologi yang terstruktur sebagaimana diilustrasikan pada Gambar 3.1.



Gambar 3.1 Alur Penelitian

3.2.1 Pengumpulan Data

Pengumpulan data dalam penelitian ini terdiri dari dua jenis, yaitu data primer dan data sekunder. Data primer dikumpulkan langsung oleh peneliti, sedangkan data sekunder diperoleh dari sumber yang sudah ada.

A. Data Primer

Data primer adalah data yang dikumpulkan langsung oleh peneliti untuk kebutuhan spesifik penelitian ini. Data primer yang digunakan meliputi citra ruangan interior lokal, aset motif keramik, dan data anotasi manual.

1. Citra Ruangan Interior Lokal

Pengambilan citra interior dilakukan secara langsung di lokasi yang relevan dengan konteks pengguna produk PT Concord Industry. Spesifikasi pengumpulan data citra ruangan disajikan pada Tabel 3.2, sedangkan contoh citra ruangan interior yang digunakan dalam penelitian ini ditunjukkan pada Gambar 3.2.

Tabel 3.1 Spesifikasi citra interiror

Parameter	**Spesifikasi**
Sumber	Citra interior
Jumlah Target	Minimal 300 citra
Resolusi Minimum	1280 × 720 piksel
Format File	JPG, PNG
Jenis Ruangan	Ruang tamu, kamar tidur
Kondisi Pencahayaan	alami
Perangkat	Kamera smartphone
Orientasi	*Landscape/Potrait*



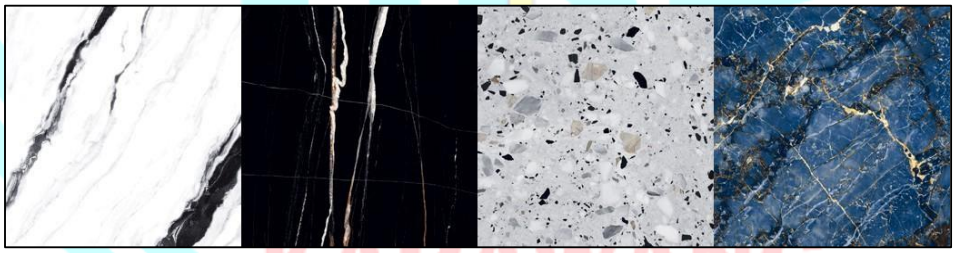
Gambar 3.2 Contoh Citra Interior

2. Aset Motif Produk Keramik Lantai

   Aset motif keramik diperoleh langsung dari katalog digital PT Concord Industry. Spesifikasi aset motif keramik disajikan pada Tabel 3.3, sedangkan contoh sampel aset motif keramik yang digunakan dalam penelitian ini ditunjukkan pada Gambar 3.3.

Tabel 3.2 Spesifikasi Aset motif Keramik Lantai

Parameter	**Spesifikasi**
Sumber	Katalog digital PT Concord Industry
Jumlah Variasi Motif	50
Resolusi	512 × 512 piksel
Format File	JPG, PNG
Kategori Motif	*Granite Tile*
Kedalaman Warna	24-bit RGB
Karakteristik	*Matte / Glossy*



Gambar 3.3 Sampel Aset Motif Keramik PT Concord Industry

3. Data Anotasi

   Untuk citra ruangan interior yang dikumpulkan, dilakukan proses anotasi manual untuk membuat *ground truth* segmentasi. Spesifikasi proses anotasi disajikan pada Tabel 3.4.

Tabel 3.3 Spesifikasi Anotasi

Parameter	**Spesifikasi**
Tools Anotasi	Roboflow
Jenis Anotasi	*Polygon segmentation*
Kelas Label	Lantai

Parameter	Spesifikasi
Format Output	PNG mask (*binary*)
Resolusi Mask	Sama dengan citra asli
Nilai Piksel	0 (bukan lantai), 255 (lantai)
Jumlah Anotator	1 orang (peneliti)
Validasi	*Cross-check* visual

B. Data Sekunder

Dataset sekunder yang digunakan dalam penelitian ini adalah SUN RGB-D yang diperoleh melalui repositori HuggingFace Datasets (wyrx/SUNRGBD_seg). Dataset ini dipilih sebagai pengganti NYU Depth V2 karena memiliki cakupan variasi ruangan *indoor* yang lebih luas dengan total 5.285 citra berlabel semantik. SUN RGB-D mencakup beragam jenis ruangan seperti kamar tidur, ruang tamu, dapur, kantor, dan ruang kelas yang diambil menggunakan berbagai sensor kedalaman. Dataset ini dimanfaatkan sebagai basis pelatihan awal (*pre-training*) model segmentasi sebelum dilakukan proses *fine-tuning* menggunakan data primer lokal. Spesifikasi dataset SUN RGB-D disajikan pada Tabel 3.5.

Tabel 3. 4 Spesifikasi Data Sekunder

Parameter	Spesifikasi
Sumber	HuggingFace (wyrx/SUNRGBD_seg)
Jumlah Citra Berlabel	5.285 citra RGB
Resolusi Asli	640 × 480 piksel
Format Citra	HuggingFace Datasets
Jenis Ruangan	Kamar tidur, ruang tamu, dapur, kantor, dll.
Anotasi	*Pixel-level semantic segmentation* (37 kelas)
Penggunaan	*Pre-training* model sebelum *fine-tuning*

C. Protokol Pengambilan Citra Ruangan

Untuk memastikan konsistensi dan kualitas data primer, diterapkan protokol pengambilan citra yang terstandarisasi. Protokol ini mencakup aspek posisi kamera, sudut pengambilan, pencahayaan, dan format penyimpanan sebagaimana disajikan pada Tabel 3.6

Tabel 3.5 Protokol Pengambilan Citra

Aspek	Ketentuan
Posisi Kamera	Setinggi dada (±120–150 cm dari lantai)
Sudut Pengambilan	Horizontal atau sedikit menunduk (0°–30°)
Jarak ke Lantai	2–5 meter dari titik terdekat lantai
Area Lantai Terlihat	Minimal 30% dari total frame
Pencahayaan	Cukup terang, hindari backlight ekstrem
Fokus	Area lantai harus tajam dan tidak *blur*
Obstruksi	Diperbolehkan furnitur menutupi sebagian lantai
Format Penyimpanan	JPG (kualitas 90%+) atau PNG

3.2.2 Pra-pemrosesan Data

Tahap pra-pemrosesan bertujuan untuk menyiapkan data agar sesuai dengan kebutuhan arsitektur U-Net dan meningkatkan generalisasi model. Proses ini meliputi normalisasi citra, penyeragaman dimensi, augmentasi data, dan pembagian dataset.

1. Normalisasi Citra

Seluruh citra input dinormalisasi ke rentang [0, 1] dengan membagi nilai piksel dengan 255. Normalisasi ini mempercepat konvergensi selama pelatihan dan menstabilkan gradien.

2. Penyeragaman Dimensi

   Citra dan mask di-*resize* ke dimensi tetap 512×512 piksel menggunakan interpolasi bilinear untuk citra RGB dan *interpolasi nearest-neighbor* untuk mask segmentasi guna mempertahankan integritas label.

3. Augmentasi Data

   Untuk meningkatkan robustness model terhadap variasi kondisi pengambilan gambar, diterapkan teknik augmentasi secara *on-the-fly* selama pelatihan. Teknik augmentasi yang digunakan beserta parameternya disajikan pada Tabel 3.7.

Tabel 3.6 Parameter Augmentasi

Teknik Augmentasi	**Parameter**	**Probabilitas**
Horizontal Flip	-	0.5
Affine Transform	*translate ±5%, scale 0.9-1.1, rotate ±10°*	0.4
Random Resized Crop	*scale 0.6-1.0, ratio 0.85-1.15*	0.4
Grid Distortion	*num_steps=5, distort_limit=0.2*	0.2
Color Jitter	*brightness=0.3, contrast=0.3, saturation=0.2, hue=0.05*	0.5
Random Brightness Contrast	*brightness ±0.2, contrast ±0.2*	0.4
Hue Saturation Value	*hue ±10, sat ±30, val ±20*	0.3

Probabilitas penerapan setiap teknik augmentasi pada Tabel 3.7 mengacu pada mekanisme *probabilistic execution* di mana setiap transformasi dijalankan secara stokastik berdasarkan parameter probabilitas (Buslaev et al., 2020). Pendekatan ini memungkinkan konstruksi pipeline augmentasi yang fleksibel dan terkontrol, sehingga tidak seluruh transformasi selalu diterapkan pada setiap sampel data, namun tetap meningkatkan keberagaman data latih tanpa merusak struktur semantik citra.

4. Pembagian Dataset

   Dataset dibagi menggunakan *stratified sampling* berdasarkan keberadaan area lantai pada setiap citra untuk menjaga distribusi kelas yang representatif. Pembagian dilakukan dengan proporsi 80% sebagai data pelatihan (*training set*) sejumlah 4.228 citra dan 20% sebagai data validasi (*validation set*) sejumlah 1.057 citra. Stratifikasi dilakukan menggunakan fungsi *train_test_split* dari pustaka scikit-learn dengan parameter random_state=42 untuk memastikan reproduktifitas hasil. Proporsi ini diterapkan pada tahap pre-training menggunakan dataset SUN RGB-D, sedangkan pada tahap *fine-tuning* menggunakan dataset lokal yang telah dibagi melalui platform Roboflow ke dalam set 70% pelatihan, 20% validasi, dan 10% pengujian.

3.2.3 Pelatihan Model Segmentasi

Model segmentasi dalam penelitian ini dibangun menggunakan arsitektur U-Net yang terdiri dari jalur *encoder*, *bottleneck*, jalur *decoder*, dan lapisan output. Arsitektur ini dipilih karena kemampuannya dalam menghasilkan segmentasi yang presisi pada level piksel.

A. Struktur U-Net

   Arsitektur U-Net yang digunakan dalam penelitian ini terdiri dari dua jalur simetris, yaitu jalur *encoder* (*contracting path*) dan jalur *decoder* (*expanding path*), yang saling terhubung melalui mekanisme *skip connections*. Struktur ini memungkinkan model mempertahankan informasi spasial penting selama proses ekstraksi dan rekonstruksi fitur.

   1. Jalur *Encoder* (*Contracting Path*)

Jalur *encoder* menggunakan model *ConvNeXt-V2-Huge* dari pustaka timm (*convnextv2_huge.fcmae_ft_in22k_in1k_512*) yang telah di-pre-train pada *ImageNet-22K* kemudian di-*fine-tune* pada *ImageNet-1K*. Model ini digunakan dalam mode *features_only* dengan empat level output yang menghasilkan *feature map* dengan jumlah channel: 352, 704, 1.408, dan 2.816. Penggunaan *encoder* pre-trained memungkinkan model memanfaatkan representasi fitur visual yang telah dipelajari dari jutaan citra, sehingga mempercepat konvergensi dan meningkatkan kemampuan generalisasi. *Drop path rate* sebesar 0.2 diterapkan sebagai mekanisme regularisasi pada *encoder*.

2. *Bottleneck*

   Lapisan terdalam dengan 1024 filter yang menangkap representasi fitur paling abstrak dari citra input.

3. Jalur *Decoder* (*Expanding Path*)

   Jalur *decoder* merekonstruksi resolusi spasial melalui tiga blok *decoder* yang masing-masing terdiri dari *Transposed Convolution* 2×2 untuk *upsampling*, modul *Attention Gate* untuk penyaringan fitur *skip connection*, konkatenasi dengan *feature map* dari *encoder*, dan dua lapisan Convolutional 3×3 dengan *Batch Normalization*, fungsi aktivasi GELU, serta Dropout2D (rate 0.1) sebagai regularisasi ringan. Konfigurasi channel pada setiap level *decoder*: 1.408, 704, dan 352.

4. Lapisan *Output*

   Lapisan *output* terdiri dari tiga tahap konvolusi bertingkat: Conv2d 3×3 (352→128 channel), Conv2d 3×3 (128→64 channel), dan Conv2d 1×1 (64→1 channel). Setiap lapisan konvolusi dilengkapi *Batch Normalization* dan fungsi aktivasi GELU, dengan Dropout2D (rate 0.2) setelah lapisan pertama. Output akhir di-interpolasi ke resolusi input asli menggunakan *bilinear interpolation*.

B. Fungsi Loss

Penelitian ini menggunakan *Compound Loss* yang terdiri dari tiga komponen untuk mengoptimalkan segmentasi pada berbagai aspek:

$$L_{total} = \alpha \cdot L_{BCE} + \beta \cdot L_{Dice} \quad (6)$$

Di mana:

1. $L_{BCE} = -[y \cdot \log(\hat{y}) + (1 - y) \cdot \log(1 - \hat{y})]$
2. $L_{Dice} = 1 - \frac{2 \cdot |y \cap \hat{y}| + s}{|y| + |\hat{y}| + s}$

Keterangan variabel dalam persamaan tersebut adalah y merupakan nilai *ground truth* pada setiap piksel, $\hat{y}$ merupakan nilai probabilitas hasil prediksi model, α dan β merupakan koefisien bobot masing-masing bernilai 0.5, serta ε merupakan *smoothing factor* dengan nilai 1×10^{-7}.

C. Konfigurasi Pelatihan

Konfigurasi parameter pelatihan model disusun untuk mengoptimalkan proses pembelajaran. Detail konfigurasi pelatihan disajikan pada Tabel 3.8.

Tabel 3.7 Konfigurasi Pelatihan Model

Parameter	***Pre-training***	***Fine-tuning***
Input Shape	*512 × 512 × 3*	*512 × 512 × 3*
Batch Size	*16*	*16*
Epochs	*34 (early stopping)*	*10*
Optimizer	*AdamW*	*AdamW*
Learning Rate	*1×10^{-4}*	*1×10^{-5}*
Weight Decay	*0.05*	*0.05*
LR Scheduler	*Cosine Annealing + Linear Warmup*	*Cosine Annealing (T_max=10, eta_min=1e-6)*
Loss Function	*Compound Loss*	*Compound Loss*
Mixed Precision	*BFloat16*	*BFloat16*

Gradient Clipping	*Max norm 1.0*	*Max norm 1.0*
Platform	*Google Colab (GPU A100)*	*Google Colab (GPU G4)*
Dataset	*SUN RGB-D (5.285 citra)*	*Data lokal Roboflow*

3.2.4 Transformasi Perspektif

Transformasi perspektif digunakan untuk memetakan motif keramik ke area lantai yang telah tersegmentasi dengan mempertahankan keselarasan sudut pandang. Proses ini melibatkan komputasi matriks homografi dan operasi *warping*.

A. Matriks Homografi

Transformasi perspektif diimplementasikan menggunakan matriks homografi H berukuran 3×3 yang memetakan koordinat piksel dari motif keramik ke area lantai yang tersegmentasi:

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (7)$$

Di mana:

1. x dan y adalah koordinat titik pada citra asli (*input image*).
2. Nilai 1 pada vektor koordinat menunjukkan bahwa titik dinyatakan dalam bentuk koordinat homogen.
3. x′ dan y'merupakan koordinat hasil transformasi perspektif pada citra tujuan.
4. w′ adalah faktor skala pada koordinat homogen yang digunakan untuk proses normalisasi.
5. Matriks H berukuran 3×3 merupakan matriks homografi yang menggambarkan hubungan transformasi perspektif antara citra sumber dan citra tujuan.
6. Elemen h_{11}hingga h_{33} adalah parameter homografi yang diperoleh dari proses estimasi berdasarkan titik-titik korespondensi.

Koordinat akhir dalam sistem kartesian dua dimensi diperoleh melalui Persamaan

$$x'' = \frac{x'}{w'}, \quad y'' = \frac{y'}{w'} \tag{8}$$

Di mana x' dan y'adalah koordinat homogen hasil transformasi perspektif, w'merupakan faktor skala, serta x'' dan y''adalah koordinat kartesian hasil normalisasi.

B. Estimasi Titik Korespondensi

Untuk menghitung matriks homografi, diperlukan minimal 4 pasang titik korespondensi antara motif sumber dan area target. Proses estimasi meliputi deteksi sudut area lantai menggunakan algoritma *convex hull* dan *aproksimasi poligonal*, pendefinisian titik sumber dari 4 sudut motif keramik, serta komputasi homografi menggunakan algoritma *Direct Linear Transform* (DLT) melalui fungsi *cv2.findHomography*() dari pustaka OpenCV.

C. Penerapan *Warping* dan Komposisi

Tahap akhir transformasi perspektif meliputi perspective *warping* menggunakan operasi *cv2.warpPerspective()* dan *alpha blending* untuk menghasilkan citra output *Virtual Try-On* melalui operasi komposisi:

$$I_{output} = mask \cdot I_{tekstur} + (1 - mask) \cdot I_{asli} \tag{9}$$

Di mana *mask* adalah hasil segmentasi yang telah dinormalisasi ke rentang [0, 1].

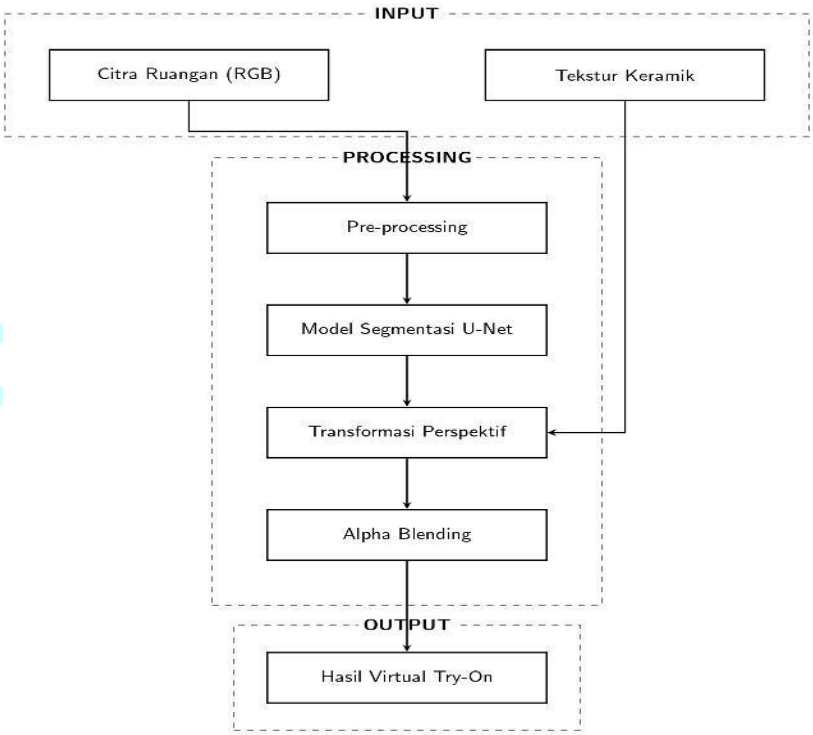
3.2.5 Implementasi Sistem

Tahap implementasi sistem merupakan proses realisasi rancangan ke dalam bentuk sistem yang dapat dioperasikan. Pada tahap ini, seluruh komponen yang telah dirancang diintegrasikan menjadi satu kesatuan sistem *Virtual Try-On* yang fungsional.

A. Arsitektur Sistem

Arsitektur sistem *Virtual Try-On* dirancang dengan pendekatan modular yang terdiri dari tiga layer utama yaitu *input layer*, *processing layer*, dan *output layer*. *Input layer* menerima dua jenis masukan berupa citra ruangan dan motif keramik. *Processing layer* melakukan serangkaian proses meliputi *pre-processing*, segmentasi menggunakan model U-Net, transformasi perspektif, dan *alpha blending*. *Output layer* menghasilkan

citra akhir berupa visualisasi ruangan dengan motif lantai baru. Diagram blok arsitektur sistem ditunjukkan pada Gambar 3.4.



Gambar 3.4 Diagram Blok Arsitektur sistem

B. Spesifikasi Perangkat Keras

Implementasi sistem memerlukan perangkat keras dengan kemampuan komputasi yang memadai, khususnya untuk proses pelatihan dan inferensi model *deep learning*. Spesifikasi perangkat keras yang digunakan dalam penelitian ini disajikan pada Tabel 3.9.

Tabel 3.8 Spesifikasi Perangkat Keras

Komponen	**Spesifikasi Minimal**	**Rekomendasi**
Prosesor	Intel Core i5 Gen 8	Intel Core i7 Gen 10+
RAM	8 GB	16 GB
GPU	NVIDIA GTX 1050 (4GB VRAM)	NVIDIA RTX 3060 (8GB VRAM)
Penyimpanan	50 GB SSD	100 GB SSD

C. Spesifikasi Perangkat Lunak

Pengembangan sistem dilakukan menggunakan bahasa pemrograman Python dengan dukungan berbagai library untuk pemrosesan citra dan *deep learning*. Spesifikasi perangkat lunak yang digunakan disajikan pada Tabel 3.10.

Tabel 3.9 Spesifikasi Perangkat Lunak

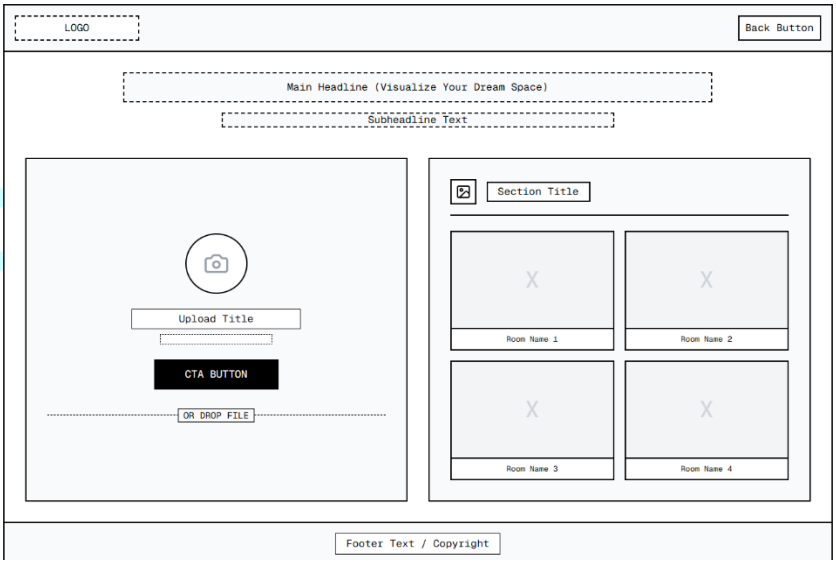
Komponen	**Versi**
Sistem Operasi	Ubuntu 20.04+
Python	3.8 - 3.10
TensorFlow/Keras	2.10+
OpenCV	4.5+
NumPy	1.21+
Matplotlib	3.5+
Next js	16+
React	19+
Tailwind CSS	4
Prisma	5+

D. Antarmuka Pengguna

Sistem diimplementasikan dalam bentuk aplikasi web dengan arsitektur *client-server* menggunakan Next.js sebagai *framework frontend* dan Python sebagai backend API. Antarmuka pengguna terdiri dari dua halaman utama dengan desain responsif dan intuitif.

1. Halaman Unggah

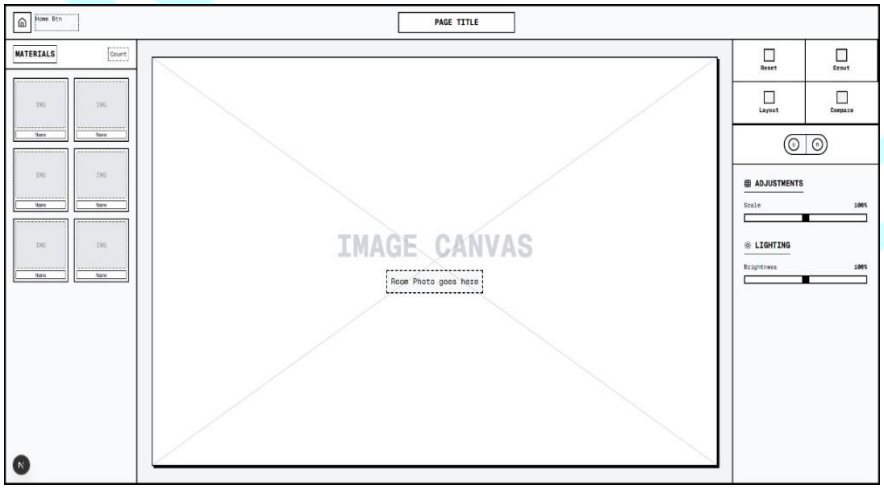
Halaman ini menyediakan area unggah citra ruangan dengan tombol pemilihan file (format JPG/PNG/WEBP) dan opsi ruangan contoh (*demo rooms*) yang dapat digunakan untuk uji coba sistem. Setelah citra dipilih, pengguna dapat melanjutkan ke tahap pemilihan motif keramik. *Wireframe* halaman unggah ditunjukkan pada Gambar 3.5.



Gambar 3.5 *Wireframe* Halaman Unggah

2. Halaman Visualisasi

Halaman ini menampilkan hasil visualisasi VTO dengan motif keramik yang telah dipetakan ke area lantai. Pengguna dapat mengganti motif keramik melalui galeri yang tersedia pada panel kiri dan melakukan penyesuaian visual melalui kontrol pada panel kanan. *Wireframe* halaman visualisasi ditunjukkan pada Gambar 3.6.



Gambar 3.6 *Wireframe* Halaman Visualisasi

3.2.6 Metode Evaluasi

Evaluasi sistem dilakukan untuk mengukur kinerja model segmentasi dan performa sistem secara keseluruhan. Metode evaluasi yang digunakan meliputi metrik kuantitatif untuk segmentasi dan pengujian performa sistem.

A. Metrik Evaluasi Segmentasi

Kinerja model segmentasi dievaluasi menggunakan beberapa metrik kuantitatif standar dalam bidang *computer vision*.

1. *Intersection over Union* (IoU)

IoU mengukur tingkat overlap antara area yang diprediksi dengan area sebenarnya. Nilai mendekati 1 menunjukkan segmentasi yang sangat akurat.

$$IoU = \frac{|\text{Prediksi} \cap \text{Ground Truth}|}{|\text{Prediksi} \cup \text{Ground Truth}|} \quad (9)$$

2. *Dice Coefficient (F1-Score)*

Dice coefficient memberikan bobot yang seimbang antara *precision* dan *recall* pada tingkat piksel.

$$Dice = \frac{2 \cdot |\text{Prediksi} \cap \text{Ground Truth}|}{|\text{Prediksi}| + |\text{Ground Truth}|} \quad (10)$$

3. *Pixel Accuracy*

Metrik dasar yang mengukur proporsi piksel yang diklasifikasikan dengan benar.

$$\text{Accuracy} = \frac{\text{Jumlah Piksel Benar}}{\text{Total Piksel}}$$

4. *Precision* dan *Recall*

Precision mengukur keakuratan prediksi positif.

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

Recall mengukur kemampuan model menemukan seluruh area positif.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (12)$$

Di mana :

1) *True Positive* (TP): Jumlah piksel lantai yang berhasil diklasifikasikan dengan benar sebagai lantai.

2) *True Negative* (TN): Jumlah piksel latar belakang (bukan lantai) yang berhasil diklasifikasikan dengan benar sebagai latar belakang.
3) *False Positive* (FP): Jumlah piksel latar belakang yang salah diklasifikasikan sebagai lantai.
4) *False Negative* (FN): Jumlah piksel lantai yang salah diklasifikasikan sebagai latar belakang.
5) ∩ (*Intersection*): Luas irisan atau area yang sama-sama dimiliki oleh prediksi dan label asli.
6) ∪ (*Union*): Luas gabungan antara area prediksi dan label asli.

B. Pengujian Performa Sistem

Pengujian performa sistem dilakukan untuk memastikan sistem dapat berjalan dengan baik. Parameter pengujian performa disajikan pada Tabel 3.11.

Tabel 3.10 Parameter Pengujian Performa

Parameter Uji	**Metode Pengukuran**
Waktu Inferensi	Rata-rata waktu pemrosesan per citra (ms)
Penggunaan Memori	*Peak memory usage* selama inferensi (MB)
Throughput	Jumlah citra yang dapat diproses per detik

3.3.7 Skenario Pengujian

Pengujian sistem dilakukan melalui beberapa skenario untuk memvalidasi fungsionalitas dan kinerja sistem secara menyeluruh. Skenario pengujian dirancang untuk mencakup berbagai kondisi *input* dan target *output* yang diharapkan sebagaimana disajikan pada Tabel 3.12.

Tabel 3.11 Sekenario Pengujian

ID	Skenario	Input	Target
UJ-01	Segmentasi ruangan standar	Citra ruangan dengan lantai terlihat jelas	IoU > 0.80
UJ-02	Segmentasi ruangan kompleks	Citra dengan furnitur menutupi sebagian lantai	IoU > 0.70
UJ-03	Transformasi perspektif	*Masking* segmentasi dan motif keramik	Pemetaan motif keramik terhadap mask segementasi.
UJ-04	Integrasi sistem	Citra ruangan dan pilihan motif keramik	Sistem berhasil Menghasilkan visualisasi sesuai motif terpilih

UBP KARAWANG