



Penerapan Multimodal Retrieval-Augmented Generation untuk Pencarian dan Generasi Motif Batik Nitik Menggunakan CLIP, BLIP-2, dan Stable Diffusion

Maylani Kusuma Wardhani Garin Muhammad Akbar Aulia Nadya Celviana
Informatics, University of Muhammadiyah Malang

Tantangan

- Sistem pencarian motif batik tradisional masih banyak bergantung pada metadata dan anotasi manual.
- Model generatif tanpa konteks berpotensi menghasilkan halusinasi visual dan motif yang tidak sesuai dengan pakem batik tradisional
- Metode pencarian konvensional sulit menghubungkan kemiripan visual dan makna filosofis secara bersamaan

Kontribusi

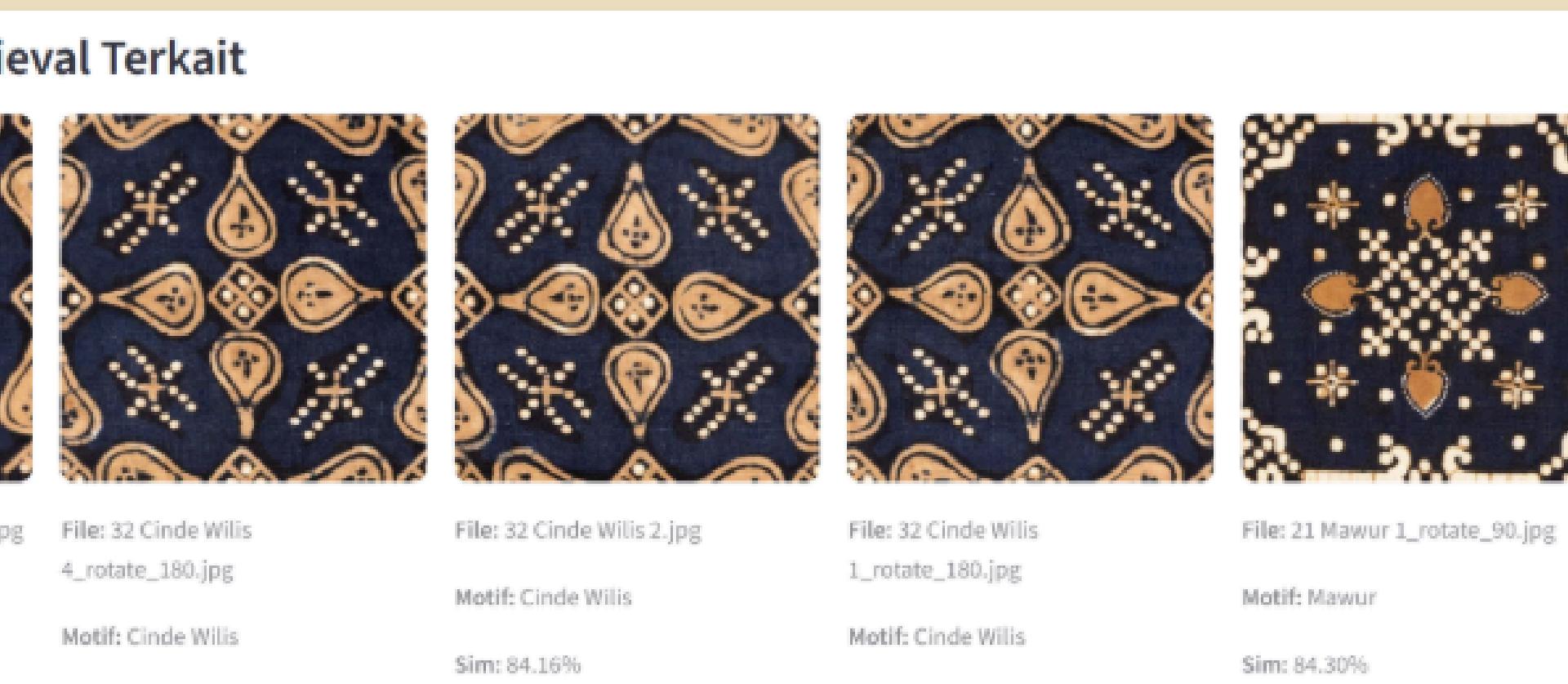
- Mengusulkan pendekatan Retrieval-Augmented Generation (RAG) multimodal untuk pencarian dan generasi motif batik nitik.
- Mengimplementasikan CLIP untuk pencarian citra berbasis teks dan gambar
- Memanfaatkan BLIP-2 untuk menghasilkan deskripsi visual sebagai pengayaan konteks
- Mengintegrasikan Stable Diffusion untuk menghasilkan motif batik nitik baru

Hasil Retrieval

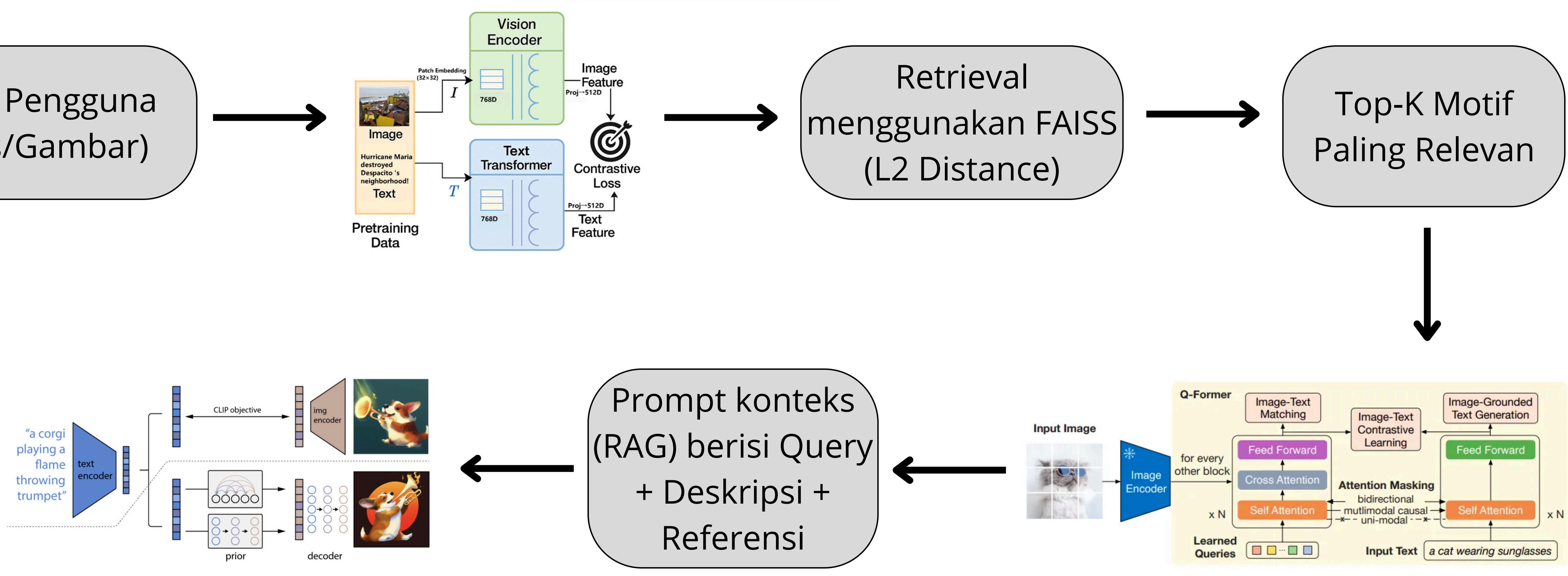
- Clip berhasil menemukan motif yang relevan secara visual dan akurat dengan query

Dataset

- Batik Nitik 960 dataset :
 - 960 images
 - 60 motif x 16 gambar
 - Augmentasi rotasi ($0^\circ, 90^\circ, 180^\circ, 270^\circ$)
 - Resolusi 512×512
 - Sumber : Mendeley Data



Metodologi



Hasil Generasi Motif

- Konsistensi Visual : Pola Nitik tetap Terjaga
- Kreativitas : Motif baru, tidak duplikat dataset



Kesimpulan

Sistem Multimodal RAG berhasil diterapkan untuk pencarian dan generasi motif Batik Nitik. Integrasi CLIP, BLIP-2, dan Stable Diffusion efektif menghasilkan motif baru yang tetap berakar pada nilai budaya