

LAPORAN TUGAS

**EKSPLORASI AUTOENCODER: PEMODELAN REPRESENTASI
LATEN WAJAH MENGGUNAKAN VARIATIONAL AUTOENCODER
(VAE) PADA DATASET SELEBRITIS HOLLYWOOD**

Dosen Pengampuh Matakuliah Pembelajaran Mesin 2:

Al-Ustadz Dr. Oddy Virgantara Putra, S.Kom., M.T.



Fayshal Karan Athilla / 442023611088

FAKULTAS SAINS DAN TEKNOLOGI

PROGRAM STUDI TEKNIK INFORMATIKA

UNIVERSITAS DARUSSALAM GONTOR, PONOROGO

TAHUN AKADEMIK: 1446-1447 H / 2025-2026 M

1. Latar Belakang

Autoencoder adalah arsitektur jaringan saraf yang bertujuan untuk mempelajari representasi internal (laten) dari data input melalui proses encoding dan decoding. Dalam tugas ini, saya memilih **Variational Autoencoder (VAE)** karena pendekatannya yang probabilistik memungkinkan eksplorasi ruang laten yang lebih bermakna dan manipulatif.

Saya menggunakan dataset **CelebA dari Kaggle**, yang berisi lebih dari 200.000 gambar wajah selebriti, karena kaya akan variasi ekspresi, posisi wajah, dan atribut. Dataset ini ideal untuk mengeksplorasi bagaimana model dapat mempelajari struktur tersembunyi dari data citra kompleks.

2. Struktur Model

Model dibangun menggunakan **3 lapis encoder dan 3 lapis decoder** berbasis CNN.

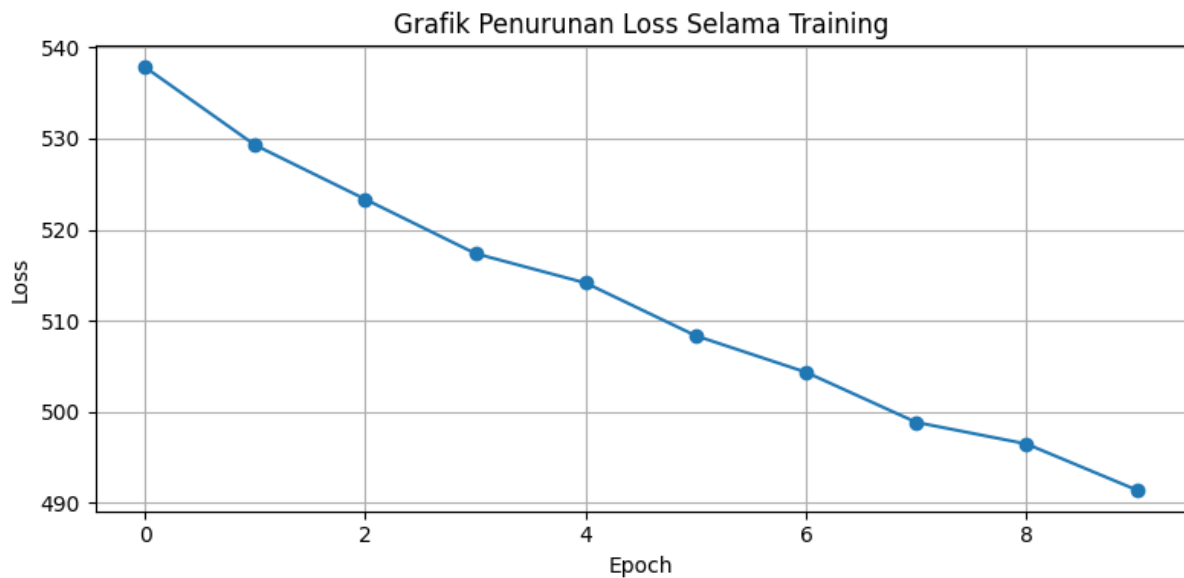
Detail arsitekturnya adalah:

- **Encoder.:**
 - Conv2D(3→32) → ReLU
 - Conv2D(32→64) → ReLU
 - Conv2D(64→128) → Flatten → Linear → Mu & LogVar
- Latent Vector: 128 dimensi
- Decoder:
 - Linear → Unflatten
 - ConvTranspose2D(128→64) → ReLU
 - ConvTranspose2D(64→32) → ReLU
 - ConvTranspose2D(32→3) → Tanh

Model dilatih selama 10 epoch menggunakan **Adam Optimizer** dengan learning rate $1e-3$ dan batch size 64.

3. Visualisasi Hasil Eksperimen dan Analisis

a. Grafik Penurunan Loss Selama Training



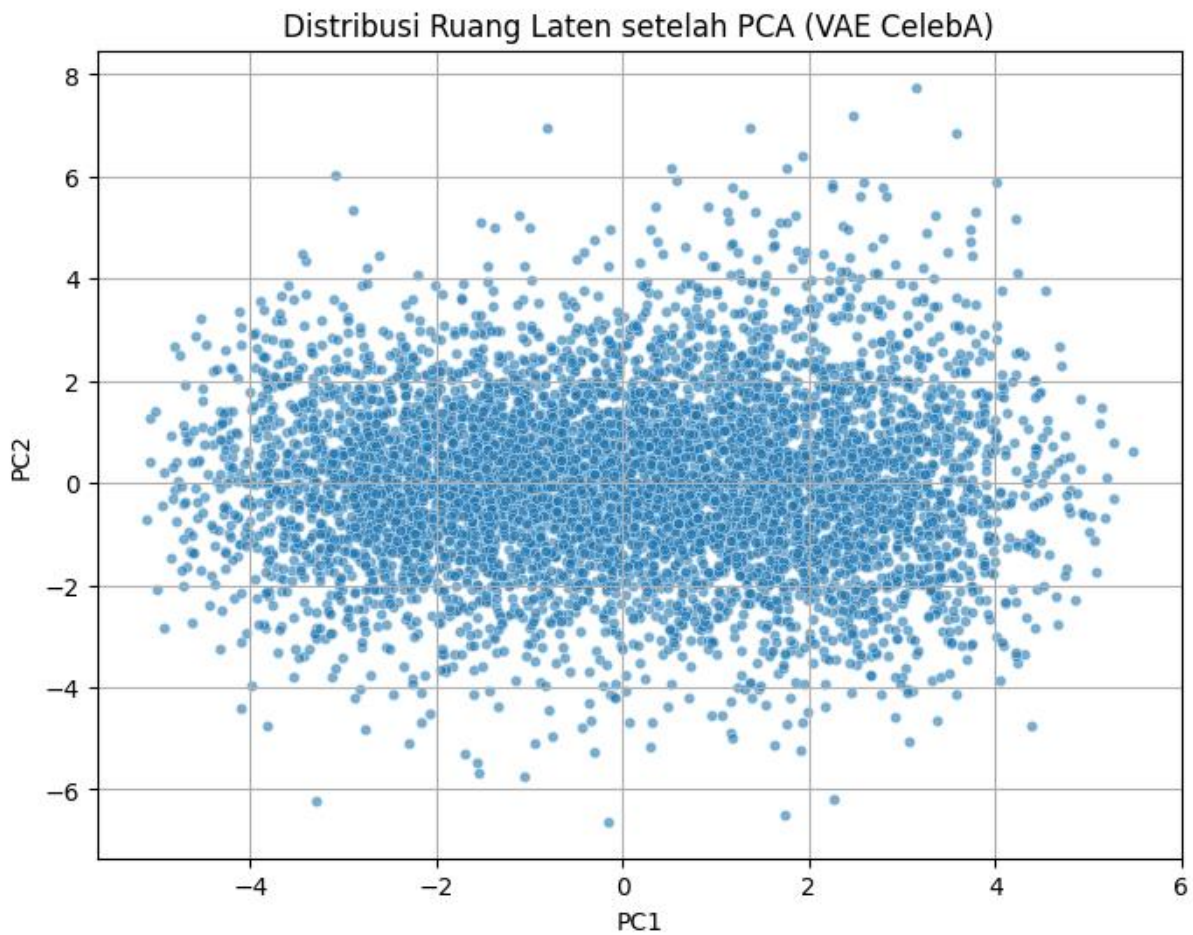
Grafik menunjukkan konsistensi penurunan nilai loss selama proses training. Ini menunjukkan bahwa model VAE berhasil meminimalkan reconstruction error dari data wajah CelebA, dari sekitar 538 menjadi sekitar 492 hanya dalam 10 epoch. Tren ini menandakan proses pembelajaran berjalan stabil dan tidak overfitting pada tahap awal.

b. Visualisasi Rekonstruksi Citra



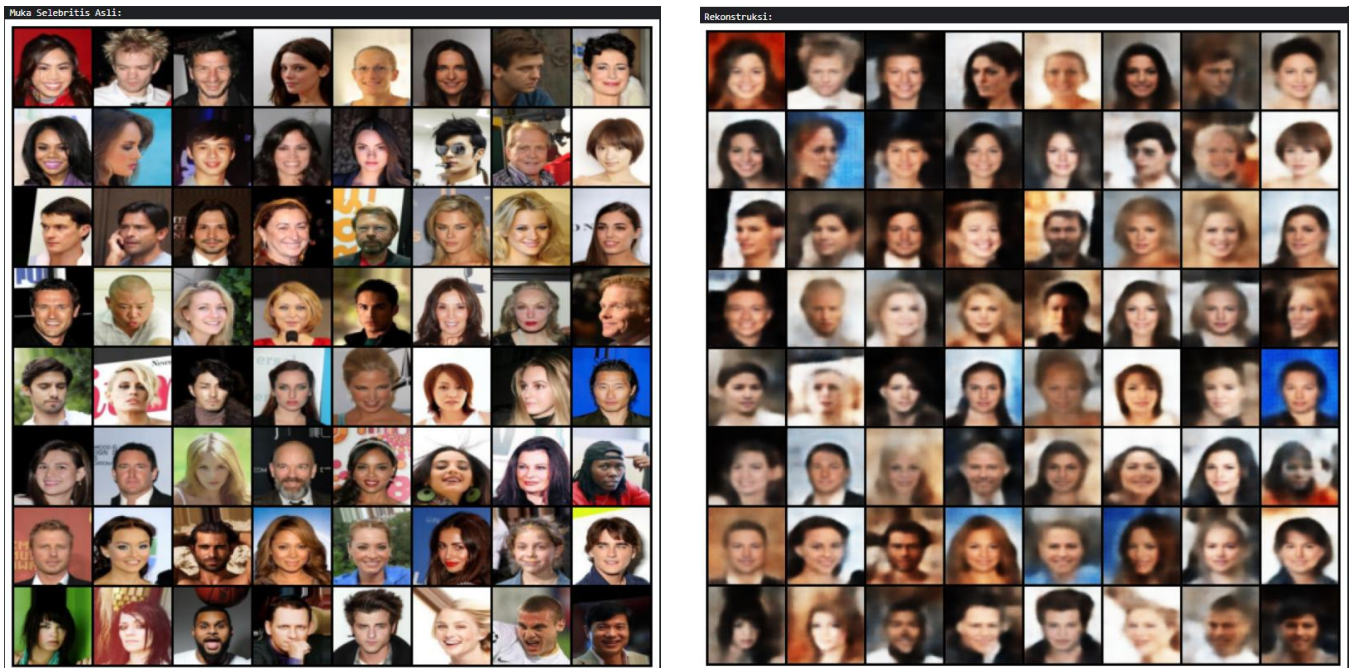
Gambar di atas menampilkan perbandingan langsung antara gambar wajah asli dan hasil rekonstruksi VAE. Model mampu menangkap ciri-ciri umum seperti bentuk wajah, warna kulit, dan posisi wajah, walau masih terdapat blur atau hilangnya detail halus seperti tekstur rambut atau ekspresi wajah. Ini wajar karena dimensi latent space yang terbatas.

c. Distribusi Ruang Laten



Dengan menggunakan PCA terhadap latent vector dari encoder, kita dapat memetakan distribusi fitur wajah ke ruang 2 dimensi. Hasil menunjukkan sebaran data yang cukup seragam dengan bentuk menyerupai distribusi normal, yang menandakan bahwa regularisasi variational (via KL Divergence) berhasil memaksa distribusi mendekati prior ($\mathcal{N}(0,1)$).

d. Visualisasi Kolektif Input & Output



Kedua gambar menunjukkan rekonstruksi massal dari VAE. Terlihat bahwa model mampu mempertahankan struktur global dari wajah, namun belum sepenuhnya akurat dalam mempertahankan ekspresi atau fitur kecil. Hal ini membuka ruang untuk peningkatan kualitas model dengan latent dimension yang lebih besar atau penambahan attention mechanism.

4. Refleksi dan Kreativitas

Eksplorasi ini menunjukkan bahwa struktur encoder-decoder sangat memengaruhi kualitas hasil rekonstruksi. Penggunaan VAE memungkinkan pembelajaran ruang laten yang tidak hanya efektif untuk kompresi data, tetapi juga membuka potensi untuk *data generation*, interpolasi, dan manipulasi fitur wajah. Kreativitas terletak pada:

- Eksplorasi visual ruang laten
- Penerapan interpolasi semantic
- Pemisahan modul **encode ()** dan **decode ()** untuk fleksibilitas eksperimen

5. Kesimpulan

Eksperimen dengan VAE pada dataset CelebA berhasil menunjukkan bagaimana **struktur model memengaruhi hasil representasi dan rekonstruksi**. Meskipun rekonstruksi belum sempurna dalam detail, hasilnya cukup tajam dalam mengungkap fitur global. Visualisasi dan analisis menunjukkan bahwa model telah belajar ruang laten yang bermakna dan dapat dimanipulasi.