BAB 1
PENDAHULUAN

1.1 Latar Belakang

Diabetes melitus merupakan penyakit yang bersifat kronis, dimana pankreas tidak bisa memproduksi insulin dalam jumlah yang cukup (Aminuddin dkk., 2023). Insulin dibutuhkan oleh tubuh sebagai alat pengontrol gula darah. Gula darah yang meningkat atau disebut dengan hiperglikemia menjadi akibat dari insulin yang tidak cukup dalam darah, dimana kondisi ini merusak banyak organ tubuh, khususnya di bagian saraf dan pembuluh darah.

Salah satu organ vital pada tubuh manusia yang dapat mengalami kerusakan akibat diabetes adalah mata. Terdapat bagian pada mata yang disebut retina, yaitu lapisan jaringan yang peka terhadap cahaya. Retina berisi jutaan sel fotoreseptor yang berfungsi menangkap cahaya dan mengubahnya menjadi sinyal visual yang diteruskan ke otak. Hiperglikemia menyebabkan kerusakan pada pembuluh darah kapiler di retina yang disebut retinopati diabetik. Penyakit ini menyebabkan penurunan kualitas penglihatan hingga kebutaan apabila tidak ditangani sejak dini (Steinmetz dkk., 2021).

Menurut data dari World Health Organization, secara global terdapat lebih dari 2,2 miliar orang mengalami gangguan penglihatan atau kebutaan, di mana sebagian besar kasus sebenarnya dapat dicegah melalui deteksi dan penanganan dini (Steinmetz dkk., 2021) Retinopati diabetika menjadi penyebab 4,8% dari total sebanyak 39 juta penduduk yang buta di seluruh dunia. Di Indonesia, sebanyak 43,1% penderita diabetes mengalami retinopati diabetik dengan penderita retinopati diabetik yang parah sebesar 26,1% (Kemenkes RI, 2023).

Retinopati diabetik umumnya tidak menimbulkan gejala kerusakan pada penglihatan karena media penglihatan atau makula masih belum mengalami kerusakan pada fase awal. Hal ini membuat penanganan dini seringkali telat dilakukan. Gejala penyakit ini meliputi peningkatan jumlah bayangan seperti bintik atau garis hitam pada penglihatan, penglihatan buram, kemampuan melihat di malam hari menurun, gangguan dalam melihat warna, hingga hilangnya kemampuan penglihatan (Purnama, 2023).

Pengecekan penyakit retinopati diabetik ini dapat dilakukan dengan beberapa metode, salah satu diantaranya yaitu funduskopi. Funduskopi adalah proses pemeriksaan retina secara visual menggunakan alat khusus yaitu lensa kamera fundus. Alat ini dapat memperbesar visual dalam mata sehingga struktur pada retina dapat terlihat jelas.

Visual yang didapatkan dari proses funduskopi yang disebut dengan citra fundus retina. Proses evaluasi citra fundus dilakukan oleh dokter mata atau oftalmologi. Melalui pengambilan citra ini, dokter mata bisa melakukan evaluasi terkait retinopati diabetik yang meliputi area makula, saraf optik dan pembuluh darah di retina.

Penggunaan kecerdasan buatan dalam melakukan deteksi retinopati diabetik saat ini sudah ada implementasinya. Terdapat beberapa aplikasi berbayar terkait dengan diagnosis penyakit ini. LumineticsCore menjadi aplikasi kecerdasan buatan yang paling banyak digunakan dalam deteksi retinopati diabetik, disusul dengan EyeArt (Teng dkk., 2026). Kedua aplikasi kecerdasan buatan ini menggunakan *Convolutional Neural Network* (CNN) sebagai arsitekturnya.

Perkembangan kecerdasan buatan memberikan metode pendekatan baru berbasis *Transformer*, seperti *Vision Transformer* (ViT) yang berfokus pada visual, yang bisa memahami fitur lebih baik pada citra (Dosovitskiy dkk., 2021). Model ViT lebih unggul dalam evaluasi citra dibandingkan dengan model-model terdahulu. Hal ini dikarenakan ViT dapat menangkap relasi keseluruhan bagian citra sekaligus melalui mekanisme *self-attention*, berbeda dengan model CNN biasa yang ekstraksi fiturnya bersifat lokal sehingga membutuhkan lebih banyak lapisan untuk memahami konteks citra secara keseluruhan.

Meski begitu, arsitektur ViT tersebut haus data sehingga perlu dataset yang besar agar ViT bisa bekerja secara optimal. Untuk mengatasi hal tersebut, dikembangkan pendekatan baru yaitu *Data-efficient Image Transformer* (DeiT) yang mengadaptasi pendekatan berbeda melalui konsep *distillation token* dengan bantuan model *teacher*. Pendekatan ini diciptakan agar arsitektur ViT mampu mencapai performa tinggi dengan data yang lebih sedikit. (Touvron dkk., 2021).

Perkembangan DeiT terbaru yang disebut dengan DeiT III melakukan adaptasi pada *training recipe* dan menghapus konsep *distillation token* yang

menjadi ciri khas DeiT generasi pertama (Touvron dkk., 2022). Penghapusan ini digantikan dengan *training recipe* yang lebih matang, sehingga model tetap dapat mencapai performa tinggi tanpa bergantung pada proses distilasi dari model *teacher* terpisah. Adaptasi ini menghasilkan model yang lebih efisien dalam penggunaan data terbatas, sekaligus mempertahankan konsep dasar mekanisme *self-attention* yang sama dengan ViT. Hingga saat ini, DeiT III menjadi andalan dari arsitektur DeiT dalam mengatasi keterbatasan data pada ViT.

Di sisi lain, terdapat penelitian baru terkait ViT pada awal tahun 2026, yaitu ViT-5, yang menghadirkan sejumlah perubahan pada tingkat arsitektur. Perubahan ini tidak mengubah konsep dasar mekanisme *self-attention* yang menjadi ciri khas ViT (F. Wang dkk., 2026). Meskipun demikian, ViT-5 masih tergolong arsitektur baru dan belum banyak diterapkan pada berbagai studi kasus. Oleh karena itu, ViT-5 perlu dibandingkan dengan DeiT III yang saat ini menjadi andalan pada arsitektur sejenis, untuk mengetahui apakah pembaruan arsitektural yang dibawa ViT-5 mampu menghasilkan performa yang setara atau bahkan lebih baik dibandingkan pendekatan efisiensi data yang sudah matang pada DeiT III, khususnya dalam konteks deteksi retinopati diabetik berdasarkan citra fundus retina.

Pada penelitian ini, akan dilatih model ViT dan DeiT berdasarkan citra fundus retina. Setiap metode memiliki varian yang membedakan tingkat performa dan penggunaan sumber daya pada saat melatih model dan memprediksi. Varian model ViT yang akan digunakan pada penelitian ini yaitu ViT-5, sejalan dengan pembahasan sebelumnya. Sedangkan untuk model DeiT, varian yang akan digunakan yaitu DeiT III. Pemilihan kedua varian ini didasarkan pada jumlah parameter yang relatif seimbang, yaitu sekitar 87 juta parameter, sehingga perbandingan performa yang dihasilkan berfokus pada perbedaan pendekatan yang dilakukan, bukan pada kesenjangan kapasitas parameter. Jumlah parameter tersebut juga memiliki kebutuhan sumber daya yang dapat ditolerir dengan spesifikasi sistem yang digunakan pada penelitian ini. Efisiensi sumber daya dalam penelitian ini ditinjau dari tiga aspek, yaitu jumlah parameter *trainable* masing-masing model, waktu yang dibutuhkan untuk satu kali pelatihan, serta waktu inferensi rata-rata per citra saat model digunakan untuk prediksi.

Penting untuk mengetahui model *Transformer* mana yang memiliki kinerja

terbaik, baik dari segi akurasi maupun efisiensi penggunaan sumber daya. Oleh karena itu, melalui penelitian ini, diharapkan dapat memberikan gambaran perbandingan kinerja model ViT-5 dan DeiT III dalam mengevaluasi penyakit retinopati diabetik berdasarkan citra fundus retina, dengan hasil produk diterapkan ke dalam aplikasi Android. Penerapan metode ini menggunakan konsep *Application Programming Interface* (API) untuk menjaga kemampuan model dalam melakukan deteksi, dengan performa yang konsisten di tiap perangkat yang berbeda. Sebagai tambahan, produk ini akan digabungkan dengan *Large Language Model* (LLM) untuk memberikan informasi lebih lanjut dan mendukung keputusan terkait dengan hasil deteksi.

1.2 Perumusan Masalah

Metode pendekatan kecerdasan buatan saat ini semakin maju berkat adanya arsitektur *Transformer*, khususnya ViT dan DeiT di bidang visual. Perkembangan terbaru menghadirkan ViT-5 sebagai penyempurnaan arsitektural ViT, serta DeiT III yang menjadi andalan dari pendekatan efisiensi data pada DeiT. Namun, ViT-5 sebagai arsitektur yang masih tergolong baru belum banyak diuji performanya dibandingkan dengan pendekatan DeiT III, khususnya dalam konteks deteksi retinopati diabetik menggunakan citra fundus retina. Oleh karena itu, diperlukan perbandingan yang jelas antara kedua model ini untuk mengetahui mana yang mampu memberikan performa lebih baik, baik dari segi akurasi maupun efisiensi penggunaan sumber daya. Penelitian ini akan melatih model ViT-5 dan DeiT III, kemudian melakukan evaluasi untuk mendapatkan model terbaik dalam konteks deteksi retinopati diabetik. Hasil penelitian diimplementasikan ke dalam aplikasi Android dengan konsep API dan tambahan *Large Language Model* (LLM) untuk membantu analisis dan mendukung keputusan berdasarkan hasil deteksi.

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah:

1) Bagaimana perbandingan performa antara model ViT-5 dan DeiT III dalam konteks deteksi retinopati diabetik?
2) Bagaimana implementasi metode *Transformer* dengan konsep API dan metode LLM tambahan ke dalam aplikasi Android?

1.3 Batasan Masalah

Agar ruang lingkup penelitian tidak terlalu luas, batasan masalah yang diterapkan adalah sebagai berikut:

1) Penyakit mata yang dideteksi terbatas pada retinopati diabetik yang terdiri dari 5 label, yaitu Normal, *Mild, Moderate, Severe,* dan *Proliverative*.
2) Data yang digunakan berupa citra fundus retina yang diperoleh dari APTOS yang dirilis pada tahun 2019 dalam kompetisi Kaggle.
3) Model diimplementasi dengan API dan digunakan pada aplikasi Android.

1.4 Tujuan dan Manfaat

1.4.1 Tujuan

Tujuan dari penelitian ini adalah mendeteksi penyakit retinopati diabetik menggunakan metode *Transformer* berdasarkan citra fundus retina.

1.4.2 Manfaat

Hasil dari penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1) Memberikan kontribusi dalam pengembangan ilmu di bidang *deep learning* dan *computer vision*,
2) Menjadi referensi bagi penelitian terkait perbandingan model *Transformer*.
3) Mendukung pengembangan aplikasi kesehatan berbasis kecerdasan buatan.
4) Membantu diagnosis penyakit mata menjadi lebih cepat, konsisten dan akurat.