

BAB I

PENDAHULUAN

1.1 Latar Belakang Penelitian

Kanker kulit merupakan ancaman kesehatan global yang semakin serius dan mendesak. Banyaknya jumlah kasus yang signifikan dalam beberapa tahun terakhir menunjukkan urgensi masalah ini. Kanker kulit secara umum terdapat dua jenis utama yaitu melanoma dan *Non-Melanoma Skin Cancer* (NMSC). Berdasarkan data yang dirilis oleh Global Cancer Observatory (2022), NMSC menempati peringkat kelima sebagai jenis kanker dengan jumlah kasus terbanyak secara global, dengan total 1.234.533 kasus baru dan 69.416 kematian tercatat pada tahun tersebut. Karsinoma sel basal dan karsinoma sel skuamosa adalah dua jenis NMSC yang berbahaya dan memiliki kasus yang cukup tinggi (Ascierto & Schadendorf, 2022). Sedangkan jenis kanker kulit melanoma menempati peringkat ke-17 dengan 331.722 kasus baru dan menyebabkan 58.667 kematian pada tahun tersebut. Meski jumlah kasusnya relatif lebih rendah dibanding jenis kanker lainnya, melanoma dikenal sebagai jenis kanker kulit yang lebih agresif dan memiliki tingkat fatalitas yang lebih tinggi. Selain itu, terdapat juga jenis tumor kulit jinak yang tidak berbahaya namun sering kali tampak mirip dengan melanoma yaitu *nevus* atau tahi lalat (Janowska dkk., 2021). Jenis ini umumnya tidak bersifat ganas dan tidak menyebabkan penyebaran sel kanker ke bagian tubuh lainnya (Zhang dkk., 2024).

Tingginya angka kasus kanker kulit jenis NMSC dan melanoma menunjukkan pentingnya deteksi dini. Deteksi dini kanker kulit sangat penting karena dapat menurunkan angka morbiditas (kondisi sakit), mortalitas (jumlah kematian), serta biaya pengobatan yang terkait, dengan memungkinkan penanganan yang lebih cepat dan efektif sebelum kanker mencapai stadium lanjut (Stătescu dkk., 2023). Semakin awal kanker kulit terdeteksi, seperti karsinoma sel basal, karsinoma sel skuamosa, atau melanoma, maka semakin besar pula kemungkinan pasien dapat menerima perawatan yang lebih efektif. Saat ini, diagnosis kanker kulit umumnya dilakukan melalui pemeriksaan visual oleh dokter atau ahli dermatologi, diikuti dengan biopsi kulit jika ditemukan lesi yang mencurigakan. Teknologi *dermoscopy* juga digunakan untuk memperbesar tampilan permukaan kulit guna

membantu identifikasi karakteristik kanker (Rey-Barroso dkk., 2021). Namun, metode tersebut memiliki keterbatasan seperti harus mengunjungi dokter spesialis atau fasilitas medis, yang mungkin sulit dijangkau oleh sebagian orang.

Salah satu solusi dari keterbatasan tersebut adalah dengan memanfaatkan teknologi *Artificial Intelligence* (AI). AI adalah cabang ilmu komputer yang memungkinkan mesin untuk meniru kemampuan manusia, seperti mengenali pola, membuat prediksi, dan mengambil keputusan berdasarkan data (Xu dkk., 2021). Seiring dengan pesatnya perkembangan AI, volume dan keragaman data yang tersedia terus meningkat secara signifikan. Data yang semakin besar dan kompleks ini membutuhkan pendekatan pengolahan yang lebih canggih dan arsitektur yang dirancang secara khusus untuk menangani tantangan tersebut. Salah satu solusi yang muncul adalah *deep learning* cabang dari *machine learning* yang menggunakan *deep neural networks* untuk meniru fungsi otak manusia (Dong dkk., 2021). Keunggulan *deep learning* terletak pada kemampuannya yang luar biasa dalam menangani data besar dan kompleks, serta mengidentifikasi pola yang rumit sehingga menjadi solusi yang efektif untuk berbagai aplikasi, seperti pengenalan gambar, klasifikasi gambar, analisis teks, pemrosesan suara, dan sebagainya. Salah satu jenis pendekatannya yaitu *supervised learning*. Di mana pada metode ini model dilatih menggunakan data berlabel, sehingga model belajar untuk memetakan *input* ke *output* yang benar berdasarkan contoh-contoh yang tersedia (Janiesch dkk., 2021). Pendekatan ini banyak digunakan khususnya untuk deteksi gambar atau klasifikasi gambar.

Salah satu arsitektur *deep learning* yang dinilai efektif untuk tugas deteksi gambar adalah *Convolutional Neural Network* (CNN) (Shete dkk., 2021). CNN dirancang khusus untuk pengenalan pola visual, seperti tepi, tekstur, dan bentuk, yang dapat diekstraksi dari gambar lesi kulit. Dengan kemampuan tersebut, arsitektur CNN memiliki potensi besar dalam mendeteksi dini kanker kulit, yang tidak hanya dapat meningkatkan peluang kesembuhan pasien, tetapi juga mengurangi risiko kematian akibat penyakit ini. Selain itu, dengan menggunakan arsitektur CNN, deteksi dini kanker kulit kini dapat dilakukan di mana saja, bahkan di lokasi yang jauh dari fasilitas kesehatan. Hal ini membuka peluang bagi

masyarakat di daerah terpencil untuk mendapatkan diagnosis yang lebih cepat tanpa harus mengunjungi fasilitas medis.

Salah satu pendekatan yang populer dalam pengembangan arsitektur CNN untuk tugas deteksi gambar adalah *transfer learning* (Salehi dkk., 2023). *Transfer learning* memungkinkan model untuk memanfaatkan pengetahuan dari model yang telah dilatih pada *dataset* besar, seperti ImageNet, kemudian menyesuaikan kemampuan tersebut untuk tugas spesifik, seperti deteksi kanker kulit. Dengan pendekatan ini, waktu pelatihan dapat dipersingkat dan performa model dapat lebih ditingkatkan, terutama ketika data yang tersedia terbatas.

Salah satu jenis arsitektur *transfer learning* yaitu EfficientNetB3 telah menunjukkan keberhasilannya dalam berbagai aplikasi di bidang medis, terutama dalam klasifikasi penyakit berbasis citra medis. Model ini terbukti efektif dalam mendeteksi kanker kulit dengan akurasi hingga 94% ketika dikombinasikan dengan teknik *data augmentation*, yang meningkatkan kinerja model dalam mengenali berbagai jenis kanker kulit (Al-Badrani dkk., 2024). Selain itu, dalam deteksi kanker paru-paru berbasis gambar CT scan, EfficientNetB3 mampu mencapai tingkat akurasi 93,33%, menunjukkan kemampuannya dalam membedakan berbagai jenis kanker paru-paru dengan presisi tinggi (Kaur & Sharma, 2024). Tidak hanya dalam bidang onkologi, model ini juga berhasil diterapkan dalam klasifikasi penyakit mata, dengan akurasi sebesar 93,8%, yang mengindikasikan potensi penggunaannya dalam lingkungan klinis untuk mendukung diagnosis oftalmologis (Soni, 2024). Keberhasilan ini menunjukkan bahwa EfficientNetB3 merupakan arsitektur *deep learning* yang andal dalam mendukung deteksi dan klasifikasi penyakit medis berbasis citra, yang dapat berkontribusi pada peningkatan efisiensi dan akurasi diagnosis medis berbasis kecerdasan buatan.

Kemajuan lebih lanjut dalam *deep learning* membawa inovasi baru yaitu *generative AI*. *Generative AI* merupakan sub bidang AI yang berfokus pada kemampuan mesin untuk menghasilkan data atau konten baru berdasarkan data yang sudah dipelajari sebelumnya (Feuerriegel dkk., 2024). Salah satu teknologi terdepan dalam *generative AI* adalah *Large Language Model* (LLM). LLM mampu memproses dan menghasilkan teks dalam skala besar. LLM menggunakan prinsip

deep learning untuk mempelajari pola dalam data bahasa alami dan menerapkan pembelajaran mendalam untuk menghasilkan teks yang bermakna dan kontekstual (Thirunavukarasu dkk., 2023).

Salah satu contoh penerapan teknologi LLM adalah *chatbot*. Memanfaatkan kemampuan untuk memahami dan menghasilkan bahasa alami, *chatbot* dapat berinteraksi secara dinamis dengan pengguna menggunakan LLM untuk memberikan tanggapan yang relevan dan akurat secara *real-time* (Tan dkk., 2024). Saat ini, ChatGPT merupakan *chatbot* paling populer, dan hal tersebut bukan tanpa alasan. berdasarkan penelitian yang dilakukan oleh Myrzakhan dkk. (2024), model GPT-4o yang digunakan dalam ChatGPT menunjukkan kinerja yang unggul secara konsisten di berbagai uji coba yang telah dilakukan. Model GPT-4o berhasil mengungguli beberapa model terkenal lainnya seperti Gemini 1.0 Pro dari Google, Claude-3 Opus dari Anthropic, dan Llama 3-70b-Instruct dari Meta.

Penelitian tersebut menggunakan berbagai parameter uji yang penting dalam mengukur keandalan LLM, dengan menggunakan beberapa *dataset* yang dirancang untuk mengukur performa pada tugas-tugas spesifik, seperti MMLU untuk mengukur pemahaman lintas disiplin, ARC untuk mengukur penalaran ilmiah, WG untuk mengukur pemahaman konteks, PIQA untuk mengukur pengetahuan fisik dan logika. CSQA mengukur pengetahuan umum, RACE untuk mengukur pemahaman bacaan, MedMCQA untuk mengukur pengetahuan medis, dan OBQA untuk mengukur pemahaman berdasarkan teks eksplisit. Berikut merupakan gambar 1.1 hasil uji perbandingan kemampuan LLM.

	Overall	MMLU	ARC	WG	PIQA	CSQA	Race	MedMCQA	OBQA	HS
GPT-4o 🏆	70.15	79.09	86.31	72.22	60.34	70.28	67.87	57.85	67.21	–
GPT-4-1106-preview 🥈	65.93	74.77	82.68	66.22	61.64	62.96	67.05	51.81	60.29	24.35
Claude-3 Opus 🥉	62.53	70.23	75.47	63.54	59.05	63.66	66.22	49.14	52.95	20.79
Mistral Large	60.84	68.76	72.32	56.83	61.21	55.35	70.17	43.44	58.66	24.47
GPT-3.5	60.32	65.38	78.42	64.56	54.89	67.89	60.11	41.42	49.90	29.99
Gemini 1.0 Pro	54.06	56.04	72.35	56.35	47.70	50.56	61.02	35.89	52.55	25.69
Llama3-70b-Instruct	52.92	59.67	67.09	57.14	43.10	55.49	58.21	41.67	40.94	–

Gambar 1.1 Hasil uji perbandingan kemampuan LLM (Myrzakhan, 2024)

Penelitian tersebut menunjukkan bahwa model GPT-4o unggul dalam perbandingan kemampuan LLM dengan mencapai rata-rata akurasi 70.15%, Mengungguli model lain seperti GPT-4-1106-preview dengan rata-rata akurasi

65.93% dan Claude-3 Opus dengan rata-rata akurasi 62.68%. Temuan tersebut menunjukkan bahwa GPT-4o memiliki performa yang lebih baik dalam tugas-tugas pemahaman bahasa alami dan menghasilkan *output* yang lebih relevan dan akurat dibandingkan dengan model-model lain yang diuji.

Selain kemampuan untuk memahami dan menghasilkan bahasa alami, GPT-4o memiliki kemampuan dalam menganalisis sebuah *input* gambar melalui fitur yang dinamakan *Vision*. Fitur ini memungkinkan GPT-4o untuk memproses dan memahami informasi visual yang terkandung dalam gambar, membuka peluang bagi aplikasi yang lebih luas dan kompleks. Menurut penelitian yang dilakukan oleh Shahriar dkk. (2024) GPT-4o *Vision* memiliki kemampuan yang sangat baik dalam menganalisis *input* visual yang berupa gambar. Gambar 1.2 menunjukkan bagaimana kemampuan GPT-4o dalam memahami dan mengolah informasi visual dengan sangat baik dibandingkan dengan versi GPT sebelumnya.

Model	Recognition	OCR	Knowledge	Language Generation	Spatial Awareness	Math	Total
GPT-3.5	22.3	31.4	15.6	16.6	32.9	24.0	27.6
GPT-4	34.3	66.3	25.6	36.6	60.6	72.0	48.1
GPT-4V	67.5	68.3	56.2	60.7	69.4	58.6	67.7
GPT-4o	80.6	87.5	75.7	77.1	84.7	88.1	83.9

Gambar 1.2 Perbandingan model GPT untuk pengenalan visual (Shahriar dkk. 2024)

Meskipun GPT-4o *Vision* menunjukkan kemampuan yang sangat baik dalam menganalisis *input* visual, terdapat beberapa keterbatasan ketika diminta untuk mendiagnosis jenis kanker kulit. Gambar 1.3 menunjukkan hasil pengujian GPT-4o *Vision* untuk mendiagnosis jenis kanker kulit dengan *input* gambar kanker kulit melanoma.



Tolong klasifikasi jenis kanker kulit ini dan berikan confidence scorenya

Gambar yang kamu kirimkan menunjukkan sebuah lesi kulit yang tidak beraturan bentuk dan warnanya — ciri-ciri ini memang khas pada **melanoma**, salah satu jenis kanker kulit yang paling berbahaya.

Namun, untuk memberikan **klasifikasi pasti** dan **confidence score**, saya membutuhkan model machine learning khusus, misalnya model CNN (Convolutional Neural Network) yang sudah dilatih pada dataset kanker kulit seperti **ISIC**.

Gambar 1.3 Pengujian GPT-4o *Vision* dalam mendiagnosis jenis kanker kulit

Hasilnya, GPT-4o *Vision* hanya dapat memberikan jawaban mengenai kemungkinan kondisi yang mungkin terjadi tanpa dapat memberikan *confidence score*. Model ini dapat mengenali beberapa karakteristik umum berdasarkan pengetahuannya, namun tidak mampu memberikan diagnosis yang spesifik. Selain itu, GPT-4o *Vision* juga tidak dapat menyajikan persentase kemungkinan jenis kanker kulit secara pasti.

Maka dari itu, penelitian ini akan mengimplementasikan model EfficientNetB3 dalam pengembangan aplikasi *chatbot* berbasis *website* untuk deteksi kanker kulit. Model EfficientNetB3 berperan utama untuk menganalisis gambar dan mengidentifikasi jenis kanker kulit berdasarkan pembelajaran dari *dataset* berlabel atau dengan pendekatan *supervised learning*. Untuk meningkatkan pengalaman pengguna, LLM GPT-4o akan diintegrasikan sebagai komponen tambahan dalam *chatbot* untuk memahami dan merespons bahasa alami pengguna, serta melakukan verifikasi awal terhadap gambar menggunakan kemampuan GPT-4o *Vision*. Verifikasi ini bertujuan menyaring gambar yang tidak relevan sebelum diproses lebih lanjut oleh model EfficientNetB3. Dengan pendekatan ini, sistem diharapkan mampu meningkatkan akurasi deteksi dini kanker kulit, sekaligus memperkaya interaksi pengguna melalui *chatbot* yang responsif dan cerdas.

1.2 Rumusan Masalah Penelitian

1. Bagaimana cara model *transfer learning* EfficientNetB3 dapat diimplementasikan untuk mengklasifikasikan jenis kanker kulit?
2. Bagaimana cara untuk membatasi ruang lingkup *chatbot* hanya pada topik terkait kanker kulit serta mengoptimalkan kemampuan *image analysis* GPT-4o *Vision*?
3. Bagaimana cara mengembangkan aplikasi *chatbot* berbasis *website* yang mengintegrasikan model *transfer learning* EfficientNetB3 kulit dan GPT-4o.

1.3 Tujuan Penelitian

1. Mengimplementasikan dan mengevaluasi model *transfer learning* EfficientNetB3 untuk mengklasifikasikan jenis kanker kulit.
2. Menganalisis dan mengimplementasikan teknik *prompt engineering* untuk membatasi ruang lingkup *chatbot* hanya pada topik terkait kanker kulit serta mengoptimalkan kemampuan *image analysis* GPT-4o *Vision*.
3. Merancang dan mengembangkan aplikasi *chatbot* berbasis *website* yang mengintegrasikan model *transfer learning* EfficientNetB3 dan GPT-4o.

1.4 Ruang Lingkup Penelitian

Berikut merupakan ruang lingkup dari penelitian ini:

1. Jenis kanker kulit yang akan diproses oleh model klasifikasi gambar adalah kanker kulit ganas yaitu melanoma, karsinoma sel basal, dan karsinoma sel skuamosa. Serta kanker kulit jinak yaitu *nevus*.
2. Sistem pada penelitian ini belum divalidasi secara langsung oleh pakar kanker kulit.
3. Akurasi model sangat bergantung pada kualitas dan representasi data pelatihan, sehingga kesalahan klasifikasi dapat terjadi jika gambar memiliki kualitas rendah atau kondisi pengambilan sudut gambar yang berbeda dengan data pelatihan.
4. GPT-4o *Vision* hanya digunakan untuk memvalidasi apakah gambar tersebut terindikasi kanker kulit atau tidak. Namun, kesalahan deteksi tetap mungkin terjadi.

5. *Output* yang dihasilkan oleh LLM-GPT-4o sepenuhnya merupakan pengetahuan LLM GPT-4o.
6. Format gambar yang diizinkan untuk deteksi gambar hanya format standar yaitu JPG, JPEG, PNG.

1.5 Manfaat Penelitian

1.5.1. Manfaat Teoritis

Penelitian ini diharapkan dapat berkontribusi dalam pengembangan literatur terkait penerapan kecerdasan buatan, khususnya dalam penggunaan arsitektur *transfer learning* EfficientNetB3 untuk mengklasifikasikan jenis kanker kulit serta integrasinya dengan LLM GPT-4o dalam memproses dan menghasilkan bahasa alami. Penelitian ini juga bertujuan untuk memberikan wawasan lebih mendalam mengenai potensi integrasi model *deep learning* untuk klasifikasi gambar dan LLM GPT-4o *Vision* untuk memvalidasi *input* gambar guna meningkatkan akurasi dan efisiensi aplikasi *chatbot* berbasis *website*. Selain itu, hasil penelitian ini dapat menjadi referensi bagi pengembangan aplikasi *chatbot* berbasis *website* yang lebih efektif dalam menyajikan informasi diagnostik dan rekomendasi interaktif berbasis AI.

1.5.2. Manfaat Praktis

Penelitian ini diharapkan dapat memberikan manfaat praktis untuk beberapa pihak, di antaranya:

1. Bagi masyarakat, menghadirkan aplikasi *chatbot* yang dapat diakses dan digunakan oleh masyarakat untuk mendapatkan informasi dan deteksi kanker kulit kapan saja dan di mana saja.
2. Bagi peneliti, penelitian ini memberikan kesempatan untuk mengembangkan pengetahuan dan keterampilan dalam pembuatan aplikasi *chatbot* dengan integrasi antara arsitektur *transfer learning* EfficientNetB3 dan LLM GPT-4o. Penelitian Ini juga membuka peluang untuk mencari metode baru dalam pengembangan *chatbot* berbasis AI yang didasarkan pada penelitian ini.

1.6 Struktur Organisasi Skripsi

Penulisan skripsi ini disusun secara sistematis dengan tujuan memberikan pemahaman yang menyeluruh terkait penelitian yang telah dilakukan. Adapun struktur organisasi skripsi ini meliputi:

BAB I: PENDAHULUAN

Bab ini berisi berbagai aspek pengantar yang menjadi landasan penelitian. Di dalamnya terdapat subbab latar belakang penelitian yang menjabarkan mengapa penelitian ini dilakukan. Subbab rumusan masalah yang merangkum pertanyaan-pertanyaan penelitian yang akan dijawab. Subbab tujuan penelitian yang menggambarkan hasil yang ingin dicapai. Subbab batasan masalah penelitian yang mendeskripsikan sejauh mana penelitian dilakukan, termasuk aspek-aspek yang akan dibahas dan yang dikecualikan, sehingga penelitian tetap fokus dan terarah. Subbab manfaat penelitian yang menunjukkan kontribusi teoritis maupun praktis yang diharapkan. Terakhir, subbab struktur organisasi skripsi yang memaparkan tata letak keseluruhan isi skripsi.

BAB II: TINJAUAN PUSTAKA

Bab ini menyajikan tinjauan pustaka yang berkaitan dengan penelitian. Bagian ini mencakup konsep-konsep dasar dan teori-teori yang menjadi landasan penelitian. Kajian pustaka berfungsi sebagai landasan teoretis yang kuat dan menempatkan penelitian ini dalam konteks penelitian yang lebih luas.

BAB III: METODE PENELITIAN

Bab ini menjelaskan secara rinci metode penelitian yang digunakan untuk mencapai tujuan penelitian, serta langkah-langkah yang dilakukan dalam proses penelitian.

BAB IV: HASIL DAN PEMBAHASAN

Bab ini berfokus pada pemaparan hasil penelitian dan analisisnya. Hasil yang diperoleh dari pengembangan *chatbot* deteksi kanker kulit dengan menggunakan model *transfer learning* EfficientNetB3 dan LLM GPT-4o berbasis *website* dijelaskan secara rinci. Analisis terhadap hasil dilakukan untuk menjawab rumusan masalah serta mengevaluasi pencapaian tujuan penelitian.

BAB V: SIMPULAN DAN SARAN

Bab terakhir ini menyajikan kesimpulan yang diambil berdasarkan analisis hasil penelitian yang telah dibahas sebelumnya. Selain itu, bab ini memberikan saran yang bermanfaat untuk penelitian selanjutnya atau penerapan praktis dari hasil penelitian, dengan harapan dapat mendukung pengembangan lebih lanjut di bidang terkait.