

Penambahan Fitur Ai *Text to Speech* dan AI *Food Recognition* Pada Aplikasi Fielthy Versi 2

1st Undang Maulana
Telecommunication Engineering
Faculty of Electrical Engineering
Bandung, Indonesia
maulana.undang1@gmail.com

2nd Rafi Hadi Muthi Wicaksono
Telecommunication Engineering
Faculty of Electrical Engineering
Bandung, Indonesia
rafihadimuthi2002@gmail.com

3rd Riski Dwi Putra
Telecommunication Engineering
Faculty of Electrical Engineering ()
Bandung, Indonesia
riskidputr8998@gmail.com

Abstrak— Perkembangan kecerdasan buatan (AI) telah mendorong sektor kesehatan untuk memanfaatkan teknologi ini dalam meningkatkan efisiensi pengawasan asupan makanan. Tujuan dari penelitian ini adalah mengembangkan sebuah aplikasi mobile yang menggabungkan teknologi AI *Food Recognition* dan *AI Text-to-Speech*, yang bertujuan membantu pengguna dalam mengidentifikasi makanan serta mendapatkan informasi mengenai kandungan gizi dan kalori secara otomatis melalui pemindaian gambar menggunakan kamera ponsel. Algoritma pembelajaran mesin yang diterapkan memastikan tingkat akurasi yang tinggi, sementara fitur *Text-to-Speech* memungkinkan informasi tersebut disampaikan secara lisan, memudahkan akses bagi pengguna dengan keterbatasan penglihatan atau yang memerlukan penggunaan tanpa tangan. Pendekatan yang digunakan dalam penelitian ini adalah kuantitatif dengan metode eksperimen untuk menguji efektivitas aplikasi dalam mengenali makanan serta memberikan rekomendasi berdasarkan data pengguna. Hasil pengujian menunjukkan akurasi yang sangat baik dan tingkat kepuasan pengguna yang tinggi, yang tercermin dalam *Net Promoter Score* (NPS) sebesar 40. Secara keseluruhan, aplikasi ini terbukti efektif dan praktis dalam membantu pengguna memantau asupan makanan secara lebih efisien dan interaktif.

Kata kunci— AI, *Food Recognition*, *Text-to-Speech*, Aplikasi Mobile, Pengenalan Gambar, Kesehatan, Pola Makan Sehat

I. PENDAHULUAN

Kemajuan teknologi digital telah membawa kemudahan dalam kehidupan sehari-hari, namun juga berdampak pada kebiasaan masyarakat dalam menjaga pola hidup sehat. Pola makan yang tidak terkontrol dan kurangnya aktivitas fisik menjadi faktor utama meningkatnya risiko penyakit seperti obesitas, diabetes, dan gangguan jantung [1]. Meskipun berbagai aplikasi kesehatan telah tersedia, banyak di antaranya belum sepenuhnya memenuhi kebutuhan pengguna, terutama dalam aspek pemantauan kesehatan yang komprehensif dan aksesibilitas informasi.

Seiring dengan perkembangan kecerdasan buatan (AI), integrasi teknologi ini dalam aplikasi kesehatan menjadi solusi potensial untuk meningkatkan efektivitas pemantauan kondisi tubuh. Pada pengembangan aplikasi Fielthy versi kedua, berbagai fitur berbasis AI ditambahkan, seperti *Food Recognition* untuk mengidentifikasi kandungan gizi makanan melalui foto, *AI Text-to-Speech* untuk meningkatkan aksesibilitas informasi, serta integrasi dengan *smartwatch* guna memantau aktivitas fisik secara *real-time*.

Penelitian ini bertujuan untuk mengatasi keterbatasan yang ada pada versi sebelumnya dengan menghadirkan fitur

yang lebih interaktif dan adaptif. Dengan pemanfaatan AI, aplikasi diharapkan dapat memberikan pengalaman yang lebih baik bagi pengguna dalam mengelola kesehatan mereka secara lebih efisien, akurat, dan mudah diakses.

II. KAJIAN TEORI

A. *AI Text to Speech*

Sistem TTS menggunakan algoritma AI untuk mengubah konten tertulis menjadi suara, membuat pembelajaran bahasa menjadi lebih interaktif dan mendukung berbagai kebutuhan pembelajaran[2]. Dalam aplikasi ini fitur *Text to Speech* diharapkan dapat membantu user yang kesusahan atau malas dalam membaca. Pada fitur ini akan digunakan untuk membacakan hasil dari menu-menu yang sudah ada maupun menu yang penulis tambahkan.

B. *AI Food Recognition*

Secara umum *Food recognition* adalah proses identifikasi dan klasifikasi jenis makanan menggunakan algoritma berbasis data visual[3]. Proses ini sering digunakan dalam aplikasi nutrisi, pemantauan kesehatan, atau sistem manajemen restoran berbasis AI. Untuk makanan sendiri, *Food Recognition* berfungsi untuk menganalisis karakteristik visual makanan, seperti bentuk, warna, tekstur, dan pola penyajian.

C. CNN (*Convolutional Neural Networks*)

CNN adalah jenis arsitektur deep learning yang dirancang khusus untuk pengolahan data yang memiliki struktur grid, seperti gambar. CNN menggunakan operasi konvolusi untuk mengekstraksi fitur penting dari data input, memungkinkan jaringan untuk mengenali pola secara otomatis. CNN sangat efektif untuk aplikasi pencitraan skala besar karena kemampuannya dalam mengadaptasi konvolusi diskrit dan memproses fitur yang beradaptasi dengan konten secara efisien[4].

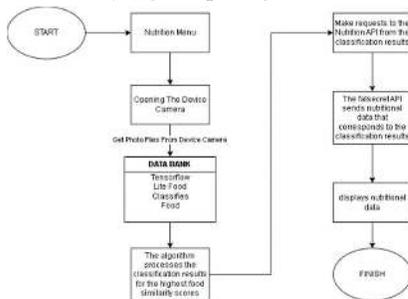
D. API (*Application Programming Interface*)

API (*Application Programming Interface*) adalah sebuah antarmuka yang memungkinkan dua aplikasi perangkat lunak untuk saling berkomunikasi. API menyediakan seperangkat aturan dan protokol yang mengatur bagaimana perangkat lunak saling bertukar data atau meminta layanan tertentu[5].

III. METODE

A. Food Recognition

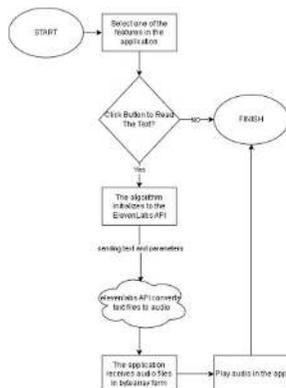
CNN adalah jenis arsitektur jaringan saraf tiruan yang sangat populer dalam pengenalan gambar, pengenalan pola, dan pemrosesan data visual. CNN dirancang khusus untuk menangani data berbentuk gambar atau data yang memiliki struktur spasial (data dengan dimensi tinggi seperti gambar, video, atau bahkan suara). Penulis memilih menggunakan Vision Classifier Food V1 sebagai model machine learning berbasis computer vision yang dirancang untuk mengenali dan mengklasifikasikan gambar makanan ke dalam berbagai kategori. Model ini merupakan salah satu implementasi dari teknik Deep Learning menggunakan Convolutional Neural Networks (CNNs), yang fokus pada pengenalan visual. Vision Classifier Food V1 berfungsi sebagai inti dari sistem pengenalan gambar. Ketika pengguna mengunggah gambar makanan, aplikasi akan mengirimkan gambar tersebut ke model untuk dianalisis dan diprediksi jenis makanannya. Model ini telah dilatih menggunakan dataset besar yang berisi berbagai gambar makanan yang telah diberi label dengan kategori yang sesuai. Dengan demikian, ketika gambar makanan diproses, model akan memberikan prediksi tentang jenis makanan yang ada pada gambar tersebut.



GAMBAR 1 Rancangan Food Recognition

B. Text to Speech

API FatSecret adalah layanan berbasis cloud yang menyediakan data dan informasi nutrisi tentang berbagai makanan, merek makanan, dan resep dari seluruh dunia[6]. API ini sering digunakan dalam aplikasi kesehatan dan kebugaran untuk membantu pengguna melacak asupan kalori dan gizi harian mereka. Dalam konteks AI Food Recognition, API FatSecret berperan sebagai sumber data untuk menghasilkan informasi nutrisi dari makanan yang dikenali oleh model AI



GAMBAR 2 Rancangan Text to Speech

IV. HASIL DAN PEMBAHASAN

A. Proses Pengujian Food Recognition

Pada pengujian ini penulis menggunakan 2 proses yaitu dengan pengujian manual dan pengujian ketangguhan. Untuk pengujian manual dilakukan dengan membandingkan nutrisi yang didapat oleh API FatSecret dengan TKPI (Tabel Komposisi Pangan Indonesia), dimana nutrisi makananyang dibandingkan akan berfokus kepada makanan yang umum di Indonesia. Untuk perbandingannya sendiri kami akan membandingkan jumlah protein, kalori, karbohidrat dan lemak. Untuk pengujian ketangguhan akan diuji coba ketangguhannya dalam beberapa skenario pencahayaan untuk memastikan model Food Recognition tetap bekerja dengan baik di berbagai kondisi pencahayaan.

Pada pengujian kali ini, proses pertama dilakukan dengan membandingkan nutrisi yang didapat oleh API FatSecret dengan TKPI (Tabel Komposisi Pangan Indonesia), dimana nutrisi makanan yang dibandingkan akan berfokus kepada makanan yang umum di Indonesia. Untuk perbandingannya sendiri kami akan membandingkan jumlah protein, kalori, karbohidrat dan lemak.

Pada Pengujian ini, model Food Recognition akan diuji coba ketangguhannya dalam beberapa skenario pencahayaan untuk memastikan model Food Recognition tetap bekerja dengan baik di berbagai kondisi pencahayaan. Untuk kondisinya sendiri sebagai berikut

- Pencahayaan Terang: Gambar dengan cahaya berlebih (overexposure).
- Pencahayaan Redup: Gambar dengan cahaya minim (underexposure).
- Pencahayaan Normal: Gambar dengan pencahayaan seimbang.
- Bayangan: Gambar dengan bayangan pada makanan

B. Proses Pengujian Text to Speech

Pada proses pengujian di Fitur AI Text to Speech ini menguji aspek – aspek yang sudah dijelaskan dan dirancang di bab 3 sebelumnya. Pengujian pertama dilakukan dengan cara mencatat latency saat aplikasi menggunakan fitur Text to Speech-nya. Setelah itu akan dibandingkan dengan banyaknya karakter huruf yang dikirim ke API-nya. Pengukuran latency disini dimulai saat pengguna menekan tombol untuk memunculkan suara pada salah satu menu.

C. Hasil Pengujian Food Recogniton

Hasil Perbandingan Kandungan Nutrisi Fielthy dengan TKPI menggunakan metode pengukuran sensitifitas serta spesifisitas mendapatkan hasil yang lumayan memuaskan, dengan hasil sebagai berikut :

- Sensitifitas: 0.67 (67%)
- Spesifistas : 1.0 (100%)
- F1-Score: 0.80 (80%)

Untuk hasil uji ketangguhan yang dilakukan untuk melihat apakah fitur Food Recognition tetap bekerja dengan baik di berbagai kondisi pencahayaan mendapatkan hasil sebagai berikut :

- Akurasi model sangat baik pada pencahayaan normal dan terang (100%).
- Akurasi menurun pada pencahayaan redup (90%) dan bayangan (80%).
- Tantangan utama: Model kurang optimal saat ada bayangan atau pencahayaan minim

D. Hasil Pengujian *Text to Speech*

Pengujian ini bertujuan untuk mengukur seberapa cepat fitur *Text to Speech* (TTS) merespons input pengguna. *Latency* diukur berdasarkan jumlah karakter yang dikirim ke API ElevenLabs, berikut adalah hasil temuan utama dari hasil pengujianya :

- *Latency* meningkat seiring bertambahnya jumlah karakter yang dikirimkan ke API
- Misalnya, untuk 78 karakter, *latency* adalah 893 ms, sedangkan untuk 507 karakter,
- *latency* naik menjadi 3764 ms.
- Tidak ada lonjakan yang ekstrem, tetapi peningkatan *latency* tetap perlu diperhatikan agar tidak mengganggu pengalaman pengguna

Selain uji *latency*, penulis juga menguji kepada pengguna dengan memberikan kuisioner yang bertujuan untuk mengetahui tingkat kepuasan pengguna yang menggunakan aplikasi. Pengujian ini diukur dari total 60 responden yang sudah mengisi via *Google Form*, berikut hasilnya :

- Mayoritas responden memberikan penilaian positif terhadap pengalaman penggunaan fitur *AI Text-to-Speech*.
- *Net Promoter Score* (NPS) sebesar 23.33, menunjukkan bahwa fitur ini cukup membantu, namun masih ada ruang untuk peningkatan.
- Tingkat kepuasan keseluruhan mencapai NPS 40, yang berarti mayoritas pengguna puas dengan performa aplikasi secara umum

V. KESIMPULAN

Hasil pengujian akurasi pengenalan makanan menunjukkan bahwa aplikasi ini mencapai nilai Sensitivitas sebesar 67%, yang mengindikasikan kemampuan sistem dalam mendeteksi makanan yang relevan dengan benar. Spesifisitas yang tinggi, sebesar 100%, juga menunjukkan bahwa sistem dapat menghindari identifikasi makanan yang tidak relevan. F1 Score, yang menggabungkan sensitivitas dan spesifisitas, mencapai nilai 80%, mencerminkan kinerja keseluruhan yang sangat baik dalam hal keseimbangan antara pengenalan makanan yang benar dan penghindaran kesalahan identifikasi.

Hasil dari pengujian *Latency Text-to-Speech*

menunjukkan bahwa aplikasi mampu merespons input pengguna dengan *latensi* yang cukup baik. *Latensi* meningkat seiring bertambahnya jumlah karakter yang dikirimkan ke API ElevenLabs, namun peningkatannya tidak ekstrem, dengan *latensi* tertinggi yang tercatat sebesar 3764 ms untuk 507 karakter. Hal ini menunjukkan bahwa meskipun responsivitas aplikasi masih dapat ditingkatkan, *latensi* tidak mengganggu pengalaman pengguna secara signifikan..

REFERENSI

- [1] D. Persetujuan Bersama, "PRESIDEN REPUBLIK INDONESIA 2-DEWAN PERWAKILAN RAKYAT REPUBLIK INDONESIA dan PRESIDEN REPUBLIK INDONESIA," 2022.
- [2] S. Bano, "INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH The Role of Artificial Intelligence-Based Technology in English Teaching and Learning," SCIENCE, ENGINEERING AND TECHNOLOGY, vol. 7, p. 18761, Dec. 2024, doi: 10.15680/IJMRSET.2024.0712197
- [3] A. ANHAR and R. A. PUTRA, "Perancangan dan Implementasi Self-Checkout System pada Toko Ritel menggunakan Convolutional Neural Network (CNN)," ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika, vol. 11, no. 2, p. 468, Apr. 2023, doi: 10.26760/elkomika.v11i2.466.
- [4] Udayana I Putu Agus Eka Darma and Nugraha Putu Gede Surya Cipta, "PREDIKSI CITRA MAKANAN MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK UNTUK MENENTUKAN BESARAN KALORI MAKANAN," p. 33, Jan. 2020 [Online]. Available: <http://www.depkes.go.id>
- [5] M. F. A. Muri, H. S. Utomo, and R. Sayyidati, "Search Engine Get Application Programming Interface," Jurnal Sains dan Informatika, vol. 5, no. 2, p. 90, Dec. 2019, doi: 10.34128/jsi.v5i2.175
- [6] Dhienalight and C. C. Lestari, "Rancang Bangun Sistem Rekomendasi Makanan Alternatif Berkalori Lebih Rendah Berbasis Konten Menggunakan Hierarchical Clustering," Teknika, vol. 9, no. 2, p. 89, Nov. 2020, doi: 10.34148/teknika.v9i2.280.