BAB II
TINJAUAN PUSTAKA

2.1.Tinjauan Pustaka

2.1.1. Digitalisasi Pencatatan Pengeluaran melalui Teknologi OCR

Dasar awal dalam pengembangan sistem expense management adalah digitalisasi proses-proses manual. Penelitian oleh Kumar et al. (2020) mengembangkan aplikasi Android yang memungkinkan pengguna untuk menyimpan expenses dengan cara memindai receipts dan mengekstrak informasi tekstual untuk pemrosesan lebih lanjut menggunakan teknologi OCR [17]. Penelitian ini menggunakan teknik preprocessing seperti image bifurcation dan shadow removal untuk menangani watermarks, invoice panjang, dan bayangan secara efektif. Meskipun berhasil memindahkan pencatatan dari manual ke digital dengan OCR, solusi semacam ini memperkenalkan masalah baru yaitu keharusan input data manual yang terulang untuk setiap transaksi dan rentan terhadap kesalahan karena akurasi OCR yang belum sempurna. Keterbatasan utama aplikasi pelacak biaya ini adalah tidak mendukung kebutuhan perjalanan kelompok dan tidak ada fitur split bill.

2.1.2. Integrasi OCR pada Sistem Manajemen Biaya Perjalanan

Untuk menghubungkan aplikasi perencanaan dan aplikasi keuangan adalah dengan otomatisasi yang lebih canggih. Hedge dkk. (2025) menjelaskan bahwa teknologi pemindaian struk (OCR) yang diintegrasikan dengan sistem manajemen seperti SAP Concur adalah kunci untuk menghilangkan proses input manual yang melelahkan [18]. Penelitian ini menunjukkan transformasi dari manual expense management menuju OCR-enabled travel expense claiming system. Dengan kemampuan untuk automate data extraction, reduce human error, enhance accuracy, dan accelerate processing times, OCR tidak hanya revolutionizing expense management tetapi juga memiliki potensi untuk transform berbagai business processes [18]. Namun, penelitian ini hanya fokus pada enterprise-level

implementation untuk travel expense management di korporasi besar dan belum mengeksplorasi aplikasi untuk individual travelers atau small groups.

2.1.3. Solusi Manajemen dan Pembagian Biaya pada Perjalanan Kelompok

Menjawab kebutuhan spesifik perjalanan kelompok, memunculkan penelitian yang berfokus pada solusi manajemen biaya kelompok. Penelitian oleh Hakim, dkk. merancang sebuah aplikasi split bill yang berfokus pada proses bisnis untuk transfer antar dompet digital[19]. Penelitian ini mengidentifikasi masalah dalam transfer antar dompet digital, terutama dalam kasus split bill dari sekelompok orang yang memiliki dompet digital yang berbeda. Biaya administrasi jika diakumulasikan bisa mencapai milyaran rupiah per bulan, fenomena yang dirasa tidak sepadan bagi masyarakat Indonesia. Aplikasi yang dirancang menggunakan metode pengembangan sistem prototyping dan responsive web untuk menawarkan solusi transfer saldo antar dompet digital secara gratis [19]. Namun, penelitian ini masih dalam tahap perancangan dan belum mengintegrasikan fitur OCR untuk automatic receipt scanning.

Adapun penelitian lainnya yang melakukan pengembangan "Military Split", sebuah aplikasi Flutter cross-platform (untuk Android & iOS) yang mampu membagi biaya hingga ke setiap barang yang dibeli (itemized bill splitting)[20]. Aplikasi ini dirancang untuk efficient bill splitting based on individual consumption, memungkinkan users untuk split expenses dengan akurat berdasarkan apa yang sebenarnya dikonsumsi oleh masing-masing person[20]. Meskipun aplikasi split bill ini canggih dalam menangani itemized splitting, aplikasi ini seperti "pulau" yang terpisah. Pengguna harus membuka aplikasi rencana perjalanannya, lalu membuka aplikasi split bill untuk mengetik ulang semua detail pengeluaran. Proses memasukkan data dua kali ini justru tidak efisien, padahal tujuan utama teknologi digital adalah untuk membuat segalanya lebih mudah.

2.1.4. Perencanaan Perjalanan Berbasis Data Real-Time

Di sisi lain, ada juga penelitian yang fokus pada pengalaman perencanaan perjalanan. Penelitian tentang Smart Trip Planner Based on Real-Time Data menunjukkan bahwa perencanaan liburan secara manual adalah proses yang cukup lama dengan banyaknya alat dan sumber untuk mencari informasi[21]. Travellers perlu mencari informasi dari berbagai sumber yang berbeda yang membuat prosesnya melelahkan dan tidak efisien. Penelitian ini mengusulkan smart trip planner yang menggunakan real-time data dan AI-based technologies untuk simplify planning process. Meskipun penelitian ini cukup baik dalam urusan trip planning dengan real-time data, akan tetapi terdapat kelemahan yang sama yakni (1) tidak mengurus pencatatan keuangan yang detail, seperti mengelola struk belanja atau tagihan makan secara langsung, dan (2) tidak ada fitur untuk split expenses dalam group travel.

2.2. Pengembangan Aplikasi *Trip Ledger*

Pemilihan model pengembangan perangkat lunak yang tepat menjadi faktor penting dalam keberhasilan suatu proyek rekayasa perangkat lunak. Studi komparatif yang dilakukan terhadap metodologi Agile dan Waterfall menunjukkan bahwa model Waterfall lebih sesuai diterapkan pada proyek dengan kebutuhan fungsional yang sudah jelas dan stabil sejak tahap awal, karena sifatnya yang sekuensial dan linear memastikan progres pengembangan berjalan disiplin melalui tahapan analisis kebutuhan, perancangan, implementasi, dan pengujian secara berurutan[22]. Karakteristik ini relevan dengan pengembangan aplikasi TripLedger, mengingat kebutuhan fungsional utama seperti pencatatan trip, manajemen anggota, split bill, dan ekstraksi struk telah dapat dirumuskan secara terstruktur sejak awal penelitian, sehingga pendekatan Waterfall dipandang sesuai untuk menjamin konsistensi antara rancangan dan hasil implementasi.

Selain aspek metodologi pengembangan, keberhasilan fitur ekstraksi struk turut ditentukan oleh bagaimana teknologi OCR diintegrasikan dengan kemampuan pemrosesan bahasa dari AI. Hasil pengujian benchmark terhadap kombinasi OCR dan Large Language Model (LLM) dalam mengekstraksi data dari struk belanja

menunjukkan tingkat keberhasilan rata-rata mencapai 97%, mengindikasikan bahwa integrasi OCR dengan LLM mampu menangani variasi tata letak struk yang kompleks serta meningkatkan akurasi dibandingkan penggunaan OCR konvensional saja [10]. Pendekatan inilah yang diadopsi dalam pengembangan TripLedger, di mana hasil ekstraksi teks mentah dari OCR diproses lebih lanjut menggunakan LLM melalui OpenAI API untuk menghasilkan data terstruktur berupa nama barang dan harga, sehingga proses pencatatan pengeluaran dapat dilakukan secara lebih akurat dan efisien.

2.3. Celah Penelitian

Dari tinjauan pustaka di atas, dapat diidentifikasi bahwa:

1) Pada aspek pencatatan pengeluaran berbasis OCR, aplikasi yang ada baru mampu mendigitalisasi proses pemindaian struk, namun masih mengharuskan input data manual berulang untuk setiap transaksi dan rentan terhadap kesalahan akibat akurasi OCR yang belum sempurna, serta belum mendukung kebutuhan perjalanan kelompok maupun fitur split bill [17].
2) Pada aspek integrasi OCR dengan sistem manajemen biaya, penerapan teknologi pemindaian struk yang telah terbukti mampu mengurangi human error dan mempercepat proses pencatatan masih terbatas pada implementasi tingkat enterprise untuk korporasi besar, dan belum dieksplorasi untuk kebutuhan individual traveler atau kelompok kecil[18].
3) Pada aspek manajemen dan pembagian biaya kelompok, solusi yang ada baru menjawab sebagian masalah secara terpisah, sebagian berfokus pada efisiensi transfer dana antar dompet digital namun belum mengintegrasikan OCR untuk pemindaian struk otomatis[23], sementara yang lain mampu melakukan itemized bill splitting secara akurat namun terisolasi dari aplikasi perencanaan perjalanan, sehingga pengguna harus menginput ulang data pengeluaran secara manual di aplikasi yang berbeda[24].

4) Pada aspek perencanaan perjalanan, solusi yang mengandalkan data real-time dan teknologi AI telah mampu menyederhanakan proses pencarian informasi perjalanan, namun belum mencakup pencatatan keuangan secara detail maupun fitur pembagian biaya untuk perjalanan kelompok [21].

Celah penelitian yang muncul adalah perlunya pengembangan aplikasi pencatatan keuangan liburan yang mengintegrasikan ekstraksi data struk berbasis OCR dan AI, pembagian biaya (split bill) secara otomatis, serta manajemen perjalanan kelompok dalam satu platform yang menyeluruh berbasis iOS. Berdasarkan tabel perbandingan penelitian yang dapat dilihat pada Tabel 2.1, terlihat bahwa aplikasi-aplikasi yang ada saat ini berkembang secara terpisah dan jarang diintegrasikan menjadi satu solusi yang utuh, sehingga pengguna harus berpindah-pindah aplikasi untuk mengelola satu perjalanan liburan secara lengkap. Hal ini menjadi landasan bagi penelitian untuk merancang solusi terintegrasi yang tidak hanya menyelesaikan pencatatan keuangan secara otomatis dan akurat, tetapi juga menyatukan seluruh kebutuhan pengelolaan keuangan liburan kelompok ke dalam satu platform yang menyeluruh.

Tabel 2. 1. Tabel Perbandingan Penelitian

	Teknologi	OCR	Metode	Fitur	Tujuan	Kelemahan
(Hegde et al. (2025))	Web/Mobile Application	Menggunakan OCR (Tesseract) + Deep Learning	Pengembangan Sistem/Aplikasi	Ekstraksi otomatis (tanggal, nominal, vendor), kategorisasi pengeluaran, analitik prediktif	Mengembangkan sistem pencatatan pengeluaran otomatis berbasis OCR untuk manajemen keuangan personal/organisasi	Tidak Menggunakan
(Kumar et al. (2020))	Android App	Menggunakan OCR Saja	Image Processing, Eksperimental	Ekstraksi data dari bill receipts	Mengembangkan aplikasi ekstraksi informasi dari struk menggunakan OCR	Tidak Menggunakan
(Hakim et al. (2023))	Web Responsif	Tidak Menggunakan OCR & AI	Prototyping, Perancangan Sistem	Split Bill, Transfer Dompet	Merancang aplikasi split bill dengan transfer antar dompet digital tanpa	Tidak Menggunakan

				Digital	biaya admin	
(Thomas et al. (2024))	Flutter (Cross-Platform)	Tidak Menggunakan OCR & AI	Mobile App Development	Itemized Bill Splitting	Mengembangkan aplikasi split bill berdasarkan individual consumption	Tidak Menggunakan
(Smart Trip Planner (2025))	Mobile App	Menggunakan AI saja	Real-Time Data Processing	Trip Planning, Itinerary	Mengembangkan smart trip planner dengan real-time data dan AI	Menggunakan AI
Margono *	iOS Native (Swift/SwiftUI)	Menggunakan OCR & AI	*Waterfall Development*	OCR Struk, Split Bill, Pencatatan Keuangan, laporan pengeluaran.	Mengembangkan aplikasi terintegrasi untuk pencatatan keuangan liburan kelompok dengan OCR dan AI, split bill, dan itinerary dalam satu platform iOS	Menggunakan AI

*) Penelitian yang dilakukan

2.4. Landasan Teori

2.4.1. Optical Character Recognition (OCR)

Optical Character Recognition (OCR) adalah teknologi yang berfungsi mengubah gambar berisi teks menjadi data digital yang dapat dibaca dan diolah oleh komputer. Proses OCR biasanya mencakup beberapa tahap, seperti pra-pemrosesan (reduksi noise dan koreksi kemiringan), segmentasi karakter, pengenalan pola, serta pascapemrosesan untuk memperbaiki kesalahan hasil pengenalan [25]. Perkembangan algoritma berbasis machine learning dan deep learning telah meningkatkan akurasi sistem OCR secara signifikan, terutama dalam menghadapi variasi font, ukuran teks, maupun kualitas gambar yang rendah [26]. Saat ini, model berbasis Convolutional Neural Network (CNN) dan Transformer banyak digunakan untuk meningkatkan ketepatan pengenalan karakter dan konteks teks[27].

Dalam implementasinya, OCR memiliki banyak aplikasi di berbagai bidang seperti perbankan, administrasi publik, logistik, dan kesehatan. Pada penelitian ini, OCR digunakan untuk mengekstraksi teks dari foto struk pengeluaran agar data transaksi dapat dicatat secara otomatis. Namun, tantangan utama OCR dalam konteks ini adalah beragamnya format dan struktur struk dari setiap toko. Oleh karena itu, hasil teks mentah dari OCR perlu diolah lebih lanjut menggunakan algoritma kecerdasan buatan seperti Large Language Model (LLM) agar dapat dikonversi menjadi data terstruktur (misalnya nama toko, total harga, dan tanggal transaksi)[14].

2.4.2. Large Language Model (LLM) dan Artificial Intelligence

Large Language Model (LLM) merupakan model kecerdasan buatan berbasis arsitektur Transformer yang dilatih menggunakan data teks dalam jumlah besar. Model ini mampu memahami konteks bahasa alami dan menghasilkan keluaran yang menyerupai tulisan manusia. Dalam beberapa tahun terakhir, LLM seperti GPT dan Claude telah menunjukkan kemampuan luar biasa dalam melakukan tugas seperti text summarization, question answering, dan data extraction[15]. Dengan kemampuan contextual understanding, LLM dapat

digunakan untuk mengekstraksi data terstruktur dari teks tidak teratur, seperti hasil keluaran OCR [28], [29].

Dalam penelitian ini, LLM digunakan untuk mengonversi hasil teks mentah dari OCR menjadi format terstruktur seperti JSON. Proses ini dilakukan dengan memberikan instruksi berupa skema keluaran yang diinginkan, misalnya: "ekstrak nama toko, total harga, dan tanggal pembelian." Pendekatan ini dikenal sebagai schema-based extraction, dan terbukti efektif tanpa memerlukan dataset pelatihan tambahan [15]. Dengan demikian, integrasi antara OCR dan LLM dapat mempercepat proses pencatatan transaksi secara otomatis, mengurangi kesalahan manusia, serta meningkatkan akurasi pencatatan keuangan pengguna.

2.4.3. iOS

iOS merupakan sistem operasi mobile yang dikembangkan oleh Apple Inc. untuk perangkat bergerak seperti iPhone, iPad, dan iPod Touch. Sistem ini pertama kali diperkenalkan pada tahun 2007 bersama dengan peluncuran iPhone generasi pertama. iOS dibangun di atas kernel XNU (X is Not Unix) yang merupakan bagian dari Darwin operating system. XNU adalah hybrid kernel yang menggabungkan Mach microkernel dengan komponen dari Berkeley Software Distribution (BSD), sehingga memberikan stabilitas tinggi dan keamanan yang kuat pada level sistem operasi [30]. Arsitektur iOS terdiri dari empat lapisan utama: Core OS layer yang mengelola interaksi langsung dengan hardware dan menyediakan layanan tingkat rendah seperti memory management dan security; Core Services layer yang menyediakan fundamental services untuk aplikasi; Media layer yang menangani grafis, audio, dan video; serta Cocoa Touch layer yang menyediakan framework untuk membangun user interface dan menangani interaksi pengguna [30]. Struktur berlapis ini memungkinkan pengembang membangun aplikasi dengan performa optimal tanpa perlu memahami kompleksitas sistem tingkat rendah secara menyeluruh.

Salah satu keunggulan utama iOS adalah fokusnya pada keamanan dan privasi pengguna. iOS mengimplementasikan berbagai mekanisme keamanan seperti application sandboxing yang membatasi akses aplikasi terhadap sistem dan

data pengguna lain, memory randomization (ASLR) untuk mencegah serangan berbasis memori, serta enkripsi pada level storage dan komunikasi antar-proses [30]. Lebih lanjut, Apple secara konsisten menekankan pentingnya on-device processing, di mana pemrosesan data sensitif seperti biometric authentication, text recognition, dan image analysis dilakukan secara lokal pada perangkat tanpa perlu mengunggah data ke server eksternal [12].

Dalam konteks penelitian ini, kebijakan privasi dan on-device processing tersebut menjadi keunggulan signifikan. Aplikasi pencatatan keuangan yang dikembangkan akan memproses data pribadi pengguna seperti foto struk dan rincian pengeluaran. Dengan memanfaatkan Apple Vision Framework yang berjalan secara lokal pada perangkat, aplikasi dapat melakukan optical character recognition (OCR) dan ekstraksi data dari struk tanpa harus mengunggah gambar ke server eksternal, sehingga menjaga keamanan dan privasi data keuangan pengguna [13]. Kombinasi antara arsitektur yang robust, mekanisme keamanan yang berlapis, dan kemampuan pemrosesan lokal menjadikan platform iOS sangat relevan untuk mendukung tujuan penelitian yang berorientasi pada keamanan, efisiensi, dan privasi pengguna.

2.4.4. Apple Vision Framework

Apple Vision Framework adalah salah satu teknologi bawaan pada iOS yang menyediakan kemampuan analisis citra berbasis computer vision. Framework ini dapat melakukan berbagai tugas seperti deteksi wajah, objek, barcode, serta teks melalui API VNRecognizeTextRequest[31]. Vision Framework mendukung dua mode utama, yaitu mode cepat (fast) untuk pemrosesan real-time dan mode akurat (accurate) yang menggunakan model deep learning untuk hasil pengenalan teks yang lebih presisi[16]. Keunggulan utama framework ini adalah pemrosesan dilakukan sepenuhnya di perangkat pengguna, sehingga privasi data tetap terjaga tanpa harus terkoneksi dengan internet.

Dalam konteks penelitian ini, Vision Framework berperan penting untuk mendeteksi dan mengenali teks dari foto struk yang diambil pengguna. Hasil pengenalan teks dari framework ini selanjutnya diproses menggunakan model LLM agar dapat diubah menjadi format data terstruktur. Pendekatan ini memberikan

keseimbangan antara efisiensi pemrosesan lokal dan fleksibilitas pemahaman konteks teks melalui kecerdasan buatan. Selain itu, Vision Framework juga memiliki dukungan integrasi penuh dengan SwiftUI dan Core ML, sehingga sangat sesuai untuk aplikasi iOS berbasis OCR yang memerlukan performa tinggi serta keamanan data pengguna [13][16].

2.4.5. Swift & Swift UI

Swift merupakan bahasa pemrograman utama yang dikembangkan oleh Apple untuk menggantikan Objective-C. Bahasa ini diperkenalkan pada tahun 2014 dan dirancang agar aman, cepat, serta mudah dipelajari. Swift memiliki sistem type-safety dan memory-safety yang mengurangi risiko kesalahan selama pengembangan, seperti kesalahan referensi memori atau null pointer [32]. Selain itu, Swift mendukung paradigma protocol-oriented programming, yang mendorong desain kode yang modular, mudah diuji, dan efisien untuk dikembangkan. Bahasa ini juga terus diperbarui dengan peningkatan performa dan sintaks yang lebih ekspresif, menjadikannya pilihan ideal untuk pengembangan aplikasi modern di seluruh ekosistem Apple[32].

SwiftUI adalah framework antarmuka pengguna (UI) berbasis deklaratif yang diperkenalkan pada tahun 2019. Framework ini memungkinkan pengembang membangun tampilan dengan cara yang lebih sederhana dan efisien dibandingkan pendekatan imperatif sebelumnya seperti UIKit. Dengan pendekatan deklaratif, tampilan antarmuka ditentukan berdasarkan state, sehingga setiap perubahan data secara otomatis memperbarui tampilan tanpa perlu menulis logika pembaruan secara manual[33]. SwiftUI juga terintegrasi erat dengan Xcode, menyediakan fitur live preview yang mempercepat proses desain antarmuka. Dalam konteks penelitian ini, penggunaan SwiftUI dianggap tepat karena memudahkan pengembangan antarmuka dinamis untuk fitur seperti input pengeluaran, tampilan histori transaksi, serta visualisasi hasil OCR secara real-time[32], [33].

2.4.6. Google Firebase

Firebase merupakan platform Backend-as-a-Service (BaaS) yang

dikembangkan oleh Google untuk mempermudah pengembangan aplikasi mobile dan web. Platform ini menyediakan berbagai layanan seperti autentikasi pengguna, basis data real-time, analitik, serta penyimpanan cloud[34]. Firebase dirancang agar pengembang dapat membangun aplikasi tanpa harus mengelola server sendiri, sehingga menghemat waktu dan biaya pengembangan. Salah satu fitur unggulan Firebase adalah Cloud Firestore, yaitu basis data NoSQL yang mendukung sinkronisasi real-time antarperangkat dan pengguna [35].

Dalam penelitian ini, Firebase digunakan untuk mengelola data hasil pemrosesan OCR dan LLM. Firebase Authentication digunakan untuk proses login pengguna secara aman, Firestore untuk menyimpan hasil ekstraksi data transaksi, dan Cloud Storage untuk menyimpan foto struk asli. Kelebihan lainnya, Firebase menyediakan Security Rules untuk membatasi akses data hanya kepada pengguna yang berhak, sesuai prinsip keamanan data [34],[35]. Dengan ekosistem yang terintegrasi dan mudah digunakan, Firebase menjadi solusi ideal untuk aplikasi pencatatan keuangan berbasis iOS yang memerlukan sinkronisasi cepat dan penyimpanan data yang andal.