

Implementasi Teknologi Optical Character Recognition (OCR) Untuk Simplifikasi Proses Pengajuan Reimbursement PPh 23 Berbasis Web

¹Fahrul Nurzaman, ²Endra Dwiyanto

¹Teknik Informatika, Universitas Persada Indonesia YAI, Jakarta

²Teknik Informatika, Universitas Persada Indonesia YAI, Jakarta

E-mail: ¹ fahrul.nurzaman@upi-yai.ac.id, ² endra.dwiyanto@gmail.com

ABSTRAK

Proses input data dari file dokumen ke form pengajuan dilakukan secara manual sehingga terdapat banyak kelemahan dan kendala dalam proses reimbursement Pph 23. Kendala yang sering terjadi adalah kesalahan input dari operator sehingga memperlambat proses pengajuan reimbursement dikarenakan perlu dilakukan revisi dan pengecekan ulang secara manual. Dengan lambatnya proses ini akan merugikan klien-klien yang harus menunggu lama hasil pengajuan reimbursement Pph 23. Dengan adanya hal tersebut maka menyebabkan banyak yang komplain yang ada datang. Hal menjadi perhatian khusus dikarenakan hal ini menjadi salah satu faktor keberlangsungan bisnis. Untuk mengurangi komplek tersebut maka perlu dibuatkan solusi yang tepat, yaitu pembacaan data secara langsung dari file dokumen dengan menerapkan teknologi Optical Character Recognition (OCR). Tujuan penelitian ini adalah membuat Aplikasi berbasis dengan menerapkan teknologi Optical Character Recognition (OCR) yaitu membaca data pada file dokumen hasil scanning dalam proses pengajuan reimbursement PPh 23. Dengan Teknologi OCR ini dapat membantu digitalisasi data secara otomatis dari dokumen pendukung seperti file scanning nota penjualan atau bukti bayar. Dengan adanya Aplikasi dengan menerapkan teknologi OCR, proses pengajuan reimbursement PPh 23 menjadi lebih mudah tanpa harus diinput datanya secara manual oleh operator. Metodologi penelitian yang digunakan adalah metode kualitatif dengan melakukan pengamatan, wawancara, dan studi literatur, serta menggunakan metode analisis deskriptif untuk menganalisa data yang diperoleh dari penelitian. Hasil penelitian menunjukkan bahwa aplikasi yang menerapkan teknologi OCR berbasis web dapat meningkatkan efektifitas dalam proses pengajuan reimbursement PPh 23 dengan tidak dilakukannya lagi input data secara manual dan dapat mengurangi kesalahan dalam menginput data.

Kata kunci : Optical Character Recognition (OCR), Reimbursement, Pph 23, Aplikasi Berbasis Web

ABSTRACT

The process of inputting data from the document file to the submission form is done manually so that there are many weaknesses and obstacles in the Pph 23 reimbursement process. The obstacle that often occurs is input errors from the operator so that it slows down the reimbursement submission process because it needs to be revised and rechecked manually. With the slow process, it will harm clients who have to wait a long time for the results of the Pph 23 reimbursement submission. With this, it causes many complaints to come. This is of particular concern because this is one of the factors of business continuity. To reduce these complaints, it is necessary to make the right solution, namely reading data directly from document files by applying Optical Character Recognition (OCR) technology. The purpose of this research is to create an application based on

applying Optical Character Recognition (OCR) technology, namely reading data on scanned document files in the process of applying for reimbursement of Income Tax 23. With this OCR technology can help digitize data automatically from supporting documents such as scanning files of sales notes or proof of payment. With the application by applying OCR technology, the process of applying for PPh 23 reimbursement becomes easier without having to input the data manually by the operator. The research methodology used is a qualitative method by conducting observations, interviews, and studies. The research methodology used is a qualitative method by conducting observations, interviews, and literature studies, and using descriptive analysis methods to analyze the data obtained from the research. The results showed that applications that apply web-based OCR technology can increase effectiveness in the process of submitting Income Tax reimbursement by no longer inputting data manually and can reduce errors in inputting data.

Keyword : Optical Character Recognition (OCR), Reimbursement, PPh 23, Web Based Application

1. PENDAHULUAN

Efektifitas dan otomatisasi merupakan hal yang sangat penting untuk diperhatikan dalam operasional bisnis. Hal ini menjadi penting karena dapat meningkatkan performansi dan kinerja dalam pelayanan bisnis dan klien, sehingga dapat menjaga kepercayaan dan loyalitas klien. Proses manual yang dilakukan dalam operasional dan administrasi dalam Proses pengajuan reimbursement PPh 23 yang dilakukan klien saat ini masih manual sudah tidak relevan lagi dikarenakan timbulnya kesalahan dalam input data. Dengan adanya kesalahan dalam input data dapat mengakibatkan keterlambatan penyelesaian proses pengajuan reimbursement dimana hal ini dapat menyebabkan complain dari klien. Untuk mengatasi permasalahan yang muncul maka perlu dilakukan solusi dalam input data dari dokumen lampiran pengajuan, yaitu dengan menerapkan teknologi OCR pada Aplikasi pengajuan reimbursement berbasis web. Penelitian ini bertujuan untuk mengembangkan Aplikasi berbasis web yang menjadi solusi inovatif dengan mengintegrasikan penerapan teknologi OCR dalam proses pengajuan reimbursement untuk pembayaran pajak PPh 23. Dengan adanya aplikasi ini menjadi sebuah Solusi yang tepat untuk

memberikan efektifitas kerja dalam operasional administrasi keuangan. Dari latar belakang di atas maka permasalahan yang dapat dirumuskan adalah sebagai berikut yaitu Bagaimana cara menerapkan teknologi OCR pada aplikasi berbasis web untuk proses pengajuan reimbursement PPh 23 agar dapat meningkatkan efektifitas dalam administrasi dan operasional. Penelitian ini memiliki batasan yaitu menggunakan Tesseract sebagai fitur OCR.

2. LANDASAN TEORI

Pajak Penghasilan Pasal 23

Pajak Penghasilan Pasal 23 (PPh 23) merupakan salah satu jenis pajak penghasilan yang ada di Indonesia. Secara singkat, Pajak Penghasilan Pasal 23 (PPh 23) merupakan pajak yang dikenakan pada penghasilan atas modal, penyertaan jasa, hadiah, bunga, deviden, royalti, atau hadiah dan penghargaan. Pemotongan PPh 23 dilakukan oleh pihak pemberi penghasilan sehubungan dengan pembayaran berupa deviden, bunga, royalti, sewa, dan jasa kepada Wajib Pajak, dan Bentuk Usaha Tetap (BUT). Terkait dalam proses bisnis, Reimbursement merupakan proses penggantian dana atau pengembalian sejumlah biaya yang telah dikeluarkan

oleh pegawai, unit organisasi atau perusahaan seperti pembelian barang atau jasa, serta pembayaran pajak, termasuk pembayaran Pajak Penghasilan Pasal 23 (PPh 23).

Web Aplikasi

Web Aplikasi yang dikembangkan menggunakan kerangka kerja web PHP Laravel. Kerangka kerja ini menggunakan arsitektur MVC (Model-View-Controller) untuk memisahkan logika, presentasi, dan pengelolaan data aplikasi. Model berfungsi dalam pengelolaan dan interaksi dengan data. View berfungsi dalam pengelolaan tampilan yang disajikan ke pengguna. Controller untuk mengelola logika aplikasi, alur informasi, dan perilaku pengguna. Dengan adanya pemisahan ini memudahkan pengembang untuk membangun aplikasi yang terstruktur, melakukan pemeliharaan aplikasi, dan penambahan perluasan lingkup aplikasi.

Optical Character Recognition (OCR)

Optical Character Recognition (OCR) adalah salah satu teknologi dimana komputer memungkinkan komputer mengenali dan mengekstrak kata atau kalimat dari sebuah file softcopy dokumen fisik atau gambar. Dengan Teknologi OCR data yang ada pada softcopy dokumen diubah menjadi format digital yang dapat disimpan, diolah, dan dicari secara mudah. Untuk mengidentifikasi karakter, kata dan kalimat teknologi OCR menggunakan algoritma pemrosesan gambar dan machine. Aplikasi diberbagai bidang menerapkan teknologi ini untuk mengolah data menjadi data digital, termasuk di bidang perpajakan. Tesseract salah satu mesin pengenalan karakter optik (OCR) yang terdapat pada Apache 2.0. Menurut Michael Everson (2018), Tesseract adalah mesin OCR yang kuat dan serbaguna yang dapat digunakan untuk mengenali teks dari berbagai jenis dokumen. Menurut Suresh Govindaraju (2020), Tesseract

adalah mesin OCR yang dapat digunakan untuk berbagai tujuan, seperti pengenalan teks, terjemahan, pencarian, dan pelacakan.

3. METODOLOGI

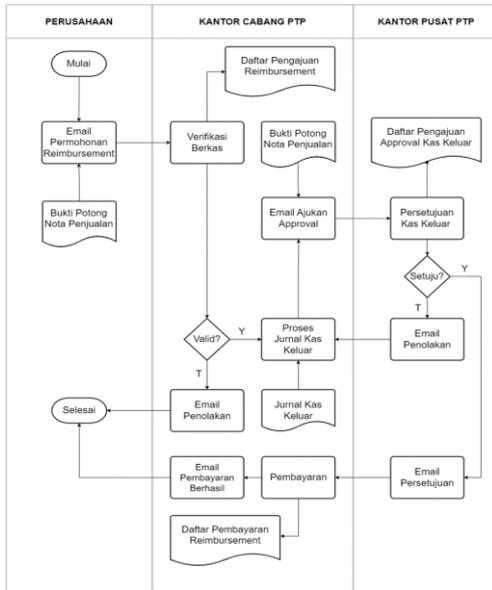
Penelitian ini menggunakan metodologi Software Development Life Cycle (SDLC) dengan model Waterfall yang terstruktur dan linier. Pada model waterfall Terdapat 4 tahap yaitu sebagai berikut:

1. Analisis Kebutuhan, pada tahap ini dilakukan pengumpulan data dan informasi terkait kebutuhan pengguna akan fungsi dan fitur aplikasi yang dikembangkan.
2. Analisa dan Perancangan Sistem, pada tahap ini dilakukan analisa kebutuhan sistem baik kebutuhan fungsional maupun non fungsional dengan membuat Arsitektur Aplikasi, alur proses Aplikasi, Skema basis data, dan perancangan antarmuka.
3. Implementasi, pada tahap ini dilakukan pembuatan kode program menggunakan bahasa pemrograman dan perangkat lunak pengembangan kode program.
4. Pengujian perangkat lunak, pada tahap ini dilakukan pemeriksaan terhadap fungsi aplikasi apakah sudah sesuai dengan kebutuhan dari pengguna.

4. HASIL DAN PEMBAHASAN

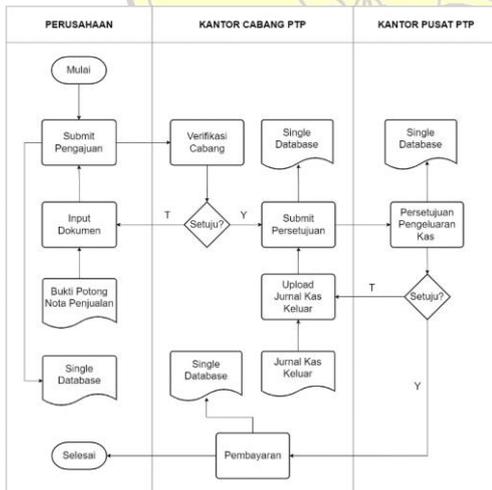
Analisa Kebutuhan

Proses yang berjalan saat ini masih manual yaitu mengisi data dengan melihat acuan dari *softcopy* dokumen yang dikirim melalui email oleh klien ke bagian administrasi. Berikut di bawah proses alur dokumen yang berjalan :



Gambar 1. Proses Bisnis yang berjalan saat ini

Untuk mengganti proses manual pada sistem berjalan saat ini maka dibuat sebuah aplikasi yang digunakan oleh klien. Berikut di bawah ini usulan proses bisnis dengan harapan penerapan aplikasi berbasis web dalam proses pengajuan reimbursement pajak yang diajukan klien.



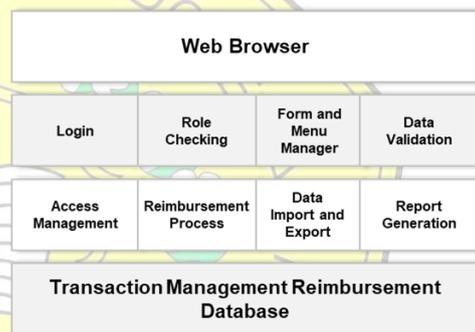
Gambar 2. Usulan Solusi Proses Bisnis yang diharapkan

Solusi di atas dapat menyederhanakan proses yang berjalan saat ini, dimana klien langsung mengakses aplikasi dengan meng-upload dokumen dan dokumen

yang diupload langsung terbaca data dengan menggunakan teknologi OCR. Data yang dibaca secara otomatis akan tampil dihalaman aplikasi dan tersimpan ke dalam database. Dengan adanya otomatis pembacaan data tersebut maka dapat mengurangi kesalahan input yang dilakukan user administrasi.

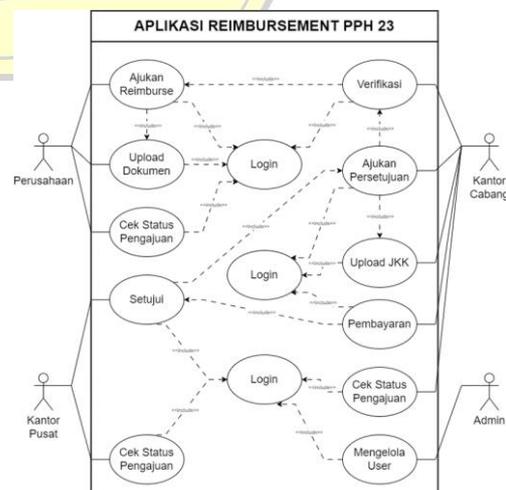
Perancangan Aplikasi

Berdasarkan usulan solusi proses pengajuan reimbursement PPh 23, maka dibutuhkan aplikasi berbasis web dengan menerapkan teknologi OCR. Berikut adalah gambaran arsitektur aplikasi tersebut.

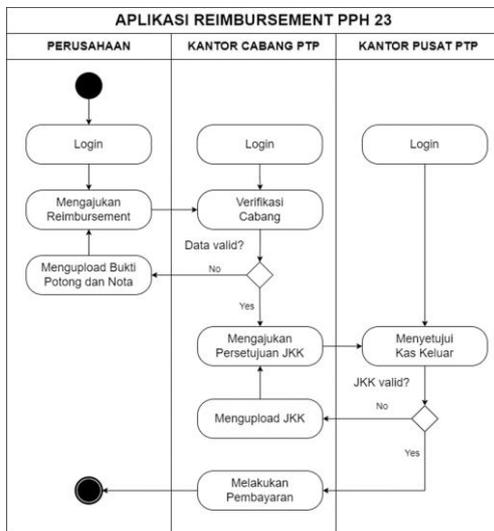


Gambar 3. Arsitektur Aplikasi

Untuk rancangan sistem dibuat use case diagram dan activity diagram untuk menggambarkan fungsi-fungsi apa saja yang terdapat dalam aplikasi.

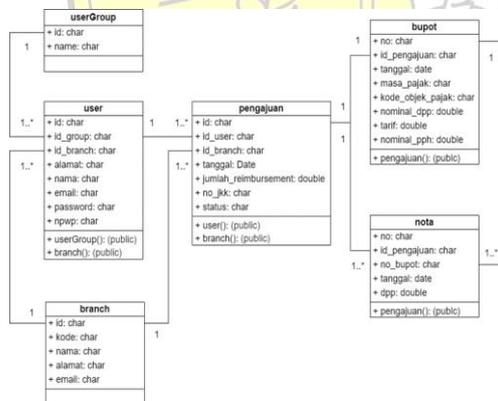


Gambar 4. Use Case Diagram Aplikasi

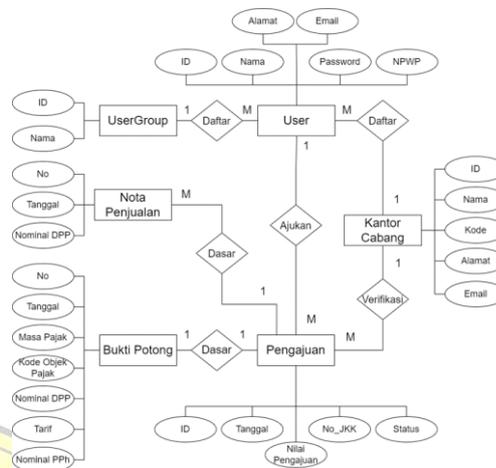


Gambar 5. Activity Diagram Aplikasi

Kemudian dilakukan perancangan class diagram untuk menggambarkan struktur dan hubungan antar kelas.



Gambar 6. Class Diagram Aplikasi



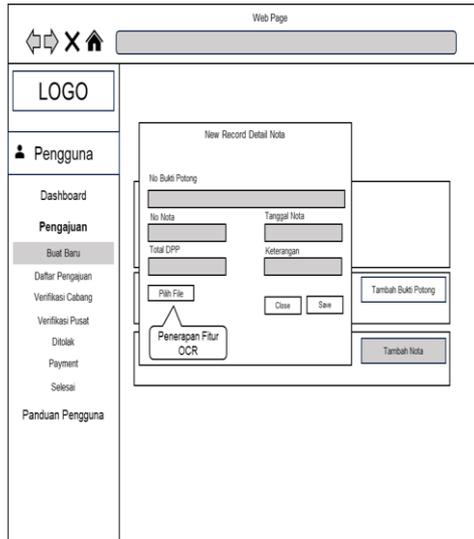
Gambar 3.7 ER-Diagram Aplikasi

Berdasarkan ER-Diagram yang sudah dibuat, berikut adalah stuktur data tabel untuk mengelola data aplikasi.

Tabel 1 Skema Tabel Database

No	Nama Tabel	Atribut	Tipe Data	Keterangan
1	users	id	char(36)	primary key
		id_kantor_cabang	char(36)	foreign key
		nama	varchar(191)	
		email	varchar(191)	
		alamat	varchar(191)	
		npwp	varchar(191)	
		nama_bank	varchar(191)	
		nomor_rekening	varchar(191)	
		password	varchar(191)	
		2	kantor_cabang	id
kode_kantor_cabang	varchar(191)			
nama	varchar(191)			
alamat	varchar(191)			
email	varchar(191)			
3	pengajuan_reimbursement	id	char(36)	primary key
		id_users	char(36)	foreign key
		id_kantor_cabang	char(36)	foreign key
		tanggal_permohonan	date	
		nilai_pengajuan	double (20,2)	
		nomor_ikk	varchar(191)	
		status	char(36)	
4	bukti_potong	nomor	char(36)	primary key
		id_pengajuan	char(36)	foreign key
		nomor_bukti_potong	varchar(191)	
		tanggal_bukti_potong	date	
		masa_pajak	varchar(191)	
		kode_objek_pajak	varchar(191)	
		nominal_dpp	double (20,2)	
5	nota_penjualan	nomor	char(36)	primary key
		id_pengajuan	char(36)	foreign key
		id_bukti_potong	char(36)	foreign key
		nomor_nota	varchar(191)	
		tanggal_nota	date	
		nominal_dpp	double (20,2)	

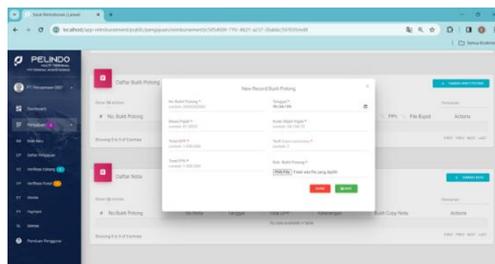
Setelah dilakukan perancangan basi data, selanjutnya dilakukan perancangan antar muka. Perancangan antar muka aplikasi bertujuan untuk membuat tampilan input dan output yang digunakan pengguna untuk menjalankan fungsi aplikasi.



Gambar 8. Rancangan Antar Muka Aplikasi

Implementasi

Pada tahap implementasi dilakukan persiapan dengan menentukan kebutuhan non fungsional seperti perangkat keras dan infrastruktur serta perangkat lunak pendukung aplikasi berbasis dan sistem manajemen database. Bahasa pemrograman dapat ditentukan untuk mendukung pembuatan kode-kode program dengan menggunakan framework Laravel. Berikut hasil implementasi berupa tampilan dari aplikasi berbasis web.



Gambar 9. Tampilan Aplikasi

Uji Coba Aplikasi

Uji coba aplikasi dilakukan untuk memastikan bahwa semua fungsi dan fitur aplikasi berjalan sesuai dengan hasil analisis dan perancangan sistem. Uji coba aplikasi juga dilakukan untuk menjawab permasalahan dan tujuan dari penelitian. Uji coba aplikasi meliputi pengujian fungsional dan pengujian akurasi. Pengujian fungsional bertujuan fungsi-fungsi pada aplikasi berjalan sesuai dengan kebutuhan pengguna. Pengujian dilakukan dengan memasukkan data dan memeriksa hasil keluaran dari proses pengolahan data.

Tabel 1 Hasil Pengujian Fungsional

Transisi	Awal	Aksi	Tujuan	Hasil Pengujian
T1	Halaman Login	Klik Login	Halaman Utama	Ok
T2	Halaman Utama	Klik Buat Baru	Halaman Buat Baru Pengajuan	Ok
T3	Halaman Utama	Klik Daftar Pengajuan	Halaman Daftar Pengajuan	Ok
T4	Halaman Informasi Surat Permohonan	Klik Tambah Bukti Potong	Halaman Tambah Bukti Potong	Ok
T5	Halaman Informasi Surat Permohonan	Klik Tambah Nota	Halaman Tambah Nota	Ok
T6	Halaman Informasi Surat Permohonan	Klik Submit Pengajuan	Halaman Notifikasi Submit Pengajuan	Ok
T7	Halaman Utama	Klik Verifikasi Cabang	Halaman Daftar Pengajuan Verifikasi Cabang	Ok
T8	Halaman Detail Surat Permohonan	Klik Setujui	Halaman Notifikasi Setujui Verifikasi Cabang	Ok
T9	Halaman Utama	Klik JKK Cabang	Halaman Daftar Pengajuan JKK Cabang	Ok
T10	Halaman Detail Surat Permohonan	Klik Upload File JKK	Halaman Notifikasi File JKK telah di-upload	Ok
T11	Halaman Detail Surat Permohonan	Klik Ajukan Pusat	Halaman Notifikasi Ajukan ke Pusat	Ok
T12	Halaman Utama	Klik Verifikasi Pusat	Halaman Daftar Pengajuan Verifikasi Pusat	Ok
T13	Halaman Detail Surat Permohonan	Klik Setujui	Halaman Notifikasi Approval Pusat	Ok
T14	Halaman Utama	Klik Payment	Halaman Daftar Pengajuan Pembayaran	Ok
T15	Halaman Detail Surat Permohonan	Klik Pembayaran Berhasil	Halaman Notifikasi Pembayaran Berhasil	Ok
T16	Halaman Utama	Klik User Pelaku Usaha	Halaman Daftar User PTP Pelaku Usaha	Ok
T17	Halaman Daftar User PTP Pelaku Usaha	Klik New Record	Halaman New Record User PTP Pelaku Usaha	Ok
T18	Halaman Utama	Klik Logout	Halaman Login	Ok

Pengujian akurasi bertujuan memvalidasi data yang diekstrak dari file dokumen lampiran. Pengujian dilakukan dengan membandingkan data yang diekstrak dengan data yang terdapat pada dokumen. Dilakukan tiga alternatif pilihan dalam pengujian akurasi untuk mengekstraksi data. Berikut di bawah hasil dari pengujian.

Tabel 2. Hasil Alternatif Pilihan 1

Item Data	Kesesuaian	Tidak Kesesuaian	Gagal Ekstraksi	Tingkat Akurasi
Nomor Nota	34	16	0	68%
Tanggal Nota	4	46	0	8%
Total DPP	41	6	3	82%
Keseluruhan	79	68	3	53%

Tabel 3. Hasil Alternatif Pilihan 2

Item Data	Kesesuaian	Tidak Kesesuaian	Gagal Ekstraksi	Tingkat Akurasi
Nomor Nota	24	11	15	48%
Tanggal Nota	44	5	1	88%
Total DPP	39	8	3	78%
Keseluruhan	107	24	19	71%

Tabel 4. Hasil Pengujian Pilihan 3

Item Data	Kesesuaian	Tidak Kesesuaian	Gagal Ekstraksi	Tingkat Akurasi
Nomor Nota	23	12	15	46%
Tanggal Nota	25	1	24	50%
Total DPP	39	6	5	78%
Keseluruhan	87	19	44	58%

Dari hasil ketiga pilihan pengujian di atas, maka pilihan yang kedua adalah pengujian yang mendekati keakurasian data, dengan tingkat akurasi keseluruhan mencapai 71%. Selanjutnya dilakukan dengan pengujian dengan mempertimbangkan ukuran piksel dari file *softcopy* dokumen untuk mendapatkan informasi terkait batasan ukuran file yang dapat diproses oleh aplikasi. Pengujian dilakukan menggunakan sample yang dapat diproses 100% pada alternatif pilihan 2, kemudian diuji berdasarkan ukuran file yang ditentukan pada lebar piksel.

Tabel 5. Hasil Pengujian Ukuran File

Lebar Piksel	Hasil Pengujian Ukuran File		
	Nomor Nota	Tanggal Nota	Total DPP
720 px	G	G	G
840 px	G	G	DS
960 px	DS	DTS	DS
1080 px	DS	DS	DS
1200 px	DS	DS	DS

Hasil pengujian dari tabel di atas bahwa batasan ukuran file berdasarkan lebar piksel yang dapat diproses oleh aplikasi adalah file dengan ukuran minimal 1080 px.

5. KESIMPULAN

Berdasarkan hasil implementasi dan pengujian aplikasi reimbursement Pph 23 berbasis web dengan menerapkan teknologi OCR dapat diambil kesimpulan sebagai berikut: Teknologi OCR dapat diimplementasikan pada aplikasi pengajuan reimbursement PPh 23 dengan menggunakan fitur Tesseract yang tersedia dalam framework Laravel. Implementasi Teknologi OCR dapat membaca data dari dokumen pendukung reimbursement secara otomatis, sehingga tidak diperlukan proses input data secara manual yang dilakukan oleh user. Tingkat keakurasian dalam membaca data dokumen pendukung adalah sebesar 71%, dengan rinciannya sebagai berikut untuk tanggal nota sebesar 88%, Total Dasar Pengenaan Pajak sebesar 78%, dan nomor nota sebesar 48% dengan ukuran lebar piksel minimal 1080 px.

6. UCAPAN TERIMA KASIH

Terima kasih kepada pihak-pihak yang sudah membantu dan mendukung penelitian ini, sehingga penelitian ini dapat diselesaikan.

DAFTAR PUSTAKA

- Aminudin. (2016). Cara Efektif Belajar Framework Laravel. Yogyakarta: Lokomedia.
- Date,C.J. (2003). An Introduction to Database System, 8th edition. Addison Wesley, Massachusets.
- Everson, M. (2018, August 20). Tesseract: A free, open-source optical character recognition engine. Retrieved from <https://tesseract-ocr.github.io/>
- F. Mohammad, J. Anarase, M. Shingote, and P. Ghanwat, "Optical character recognition implementation using pattern matching," International Journal of Computer Science and Information Technologies, vol. 5, no. 2, pp. 2088–2090, 2014.
- Fowler, Martin. (2002) Patterns of Enterprise Application Architecture. Martin Fowler.
- Pressman, Roger S. (2014) Software Engineering: A Practitioner's Approach, Edisi ke-8. McGraw-Hill Education.

