

PENGEMBANGAN APLIKASI TRANSLASI TEKS ARAB PEGON KE BAHASA INDONESIA MENGGUNAKAN INTEGRASI OCR DAN API AI BERBASIS ANDROID

Lutfiya Rahmatika Sari, Tri Listyorini, Endang Supriyati

Teknik Informatika, Universitas Muria Kudus
Gondangmanis, PO Box 53, Bae, Kudus 59352
lutfiyarahmatika@gmail.com

ABSTRAK

Arab Pegon merupakan sistem penulisan berbasis aksara Arab yang digunakan untuk menuliskan bahasa Jawa, namun hingga saat ini belum didukung secara langsung oleh layanan penerjemahan mesin populer. Kondisi tersebut menyebabkan proses pemahaman teks Arab Pegon masih dilakukan secara manual dan memerlukan kemampuan khusus. Penelitian ini bertujuan untuk mengembangkan sistem penerjemahan Arab Pegon ke bahasa Indonesia berbasis aplikasi Android dengan memanfaatkan *Optical Character Recognition* (OCR) dan proses transliterasi. Sistem dibangun dengan menggunakan Tesseract OCR untuk mengekstraksi teks Arab Pegon dari citra, kemudian hasil OCR ditransliterasikan ke dalam bentuk Jawa Latin menggunakan pendekatan berbasis kamus. Selanjutnya, teks hasil transliterasi diterjemahkan ke bahasa Indonesia menggunakan layanan penerjemahan mesin. Implementasi aplikasi dilakukan menggunakan Kodular sebagai antarmuka pengguna. Hasil penelitian menunjukkan bahwa Tesseract OCR mampu mengenali teks Arab Pegon dari citra, meskipun hasil pengenalan masih memerlukan penyesuaian sebelum digunakan pada tahap translasi. Oleh karena itu, sistem menyediakan mekanisme penyuntingan hasil OCR dan pendekatan transliterasi berbasis kamus untuk meningkatkan keterbacaan teks. Aplikasi ini diharapkan dapat membantu pengguna dalam memahami teks Arab Pegon secara lebih mudah melalui tahapan translasi yang terstruktur.

Kata kunci : *Android, Arab Pegon, OCR, Transliterasi, Penerjemahan Mesin*

1. PENDAHULUAN

Arab Pegon merupakan sistem penulisan bahasa Jawa menggunakan huruf Arab yang banyak digunakan dalam berbagai literatur keagamaan, khususnya di lingkungan pesantren [1]. Sistem penulisan ini sering dijumpai dalam berbagai kitab keagamaan, seperti tafsir, fiqh, nahwu, dan kajian akhlak, salah satunya *Tafsir al-Ibriz* karya KH. Bisri Musthofa [1] [2]. Penggunaan Arab Pegon awalnya bertujuan untuk memudahkan pemahaman masyarakat Jawa terhadap isi teks berbahasa Arab [3]. Namun, pemahaman teks Arab Pegon masih menjadi tantangan, terutama bagi pembaca yang tidak terbiasa dengan bentuk tulisan tersebut atau tidak menguasai kaidah bahasa Jawa yang digunakan di dalamnya [4].

Perkembangan teknologi penerjemahan mesin telah memberikan kemudahan dalam menerjemahkan berbagai bahasa secara otomatis [5]. Namun, hingga saat ini layanan penerjemahan yang tersedia secara umum, seperti Google Translate, belum menyediakan dukungan penerjemahan langsung untuk Arab Pegon ke Bahasa Indonesia. Opsi bahasa yang tersedia hanya mencakup bahasa Arab dan bahasa Jawa. Kondisi ini menyebabkan teks Arab Pegon tidak dapat langsung diterjemahkan secara otomatis dan sering kali diperlakukan sebagai bahasa Arab, sehingga hasil terjemahan menjadi kurang sesuai dengan makna yang dimaksud.

Permasalahan tersebut menunjukkan adanya kesenjangan antara kebutuhan penerjemahan teks Arab Pegon dan kemampuan teknologi penerjemahan mesin yang tersedia. Mengingat Arab Pegon pada dasarnya merupakan bahasa Jawa yang ditulis

menggunakan huruf Arab, diperlukan pendekatan khusus agar proses penerjemahan dapat dilakukan secara lebih tepat. Salah satu pendekatan yang dapat diterapkan adalah transliterasi, yaitu mengubah teks Arab Pegon ke dalam bentuk bahasa Jawa beraksara Latin sebelum dilakukan penerjemahan ke Bahasa Indonesia [6].

Dalam penelitian ini, transliterasi digunakan sebagai tahap perantara untuk menjembatani keterbatasan layanan penerjemahan mesin. Proses transliterasi memungkinkan teks Arab Pegon diperlakukan sebagai bahasa Jawa yang sesuai dengan dukungan bahasa pada layanan penerjemahan. Selanjutnya, teks hasil transliterasi diterjemahkan ke Bahasa Indonesia menggunakan layanan penerjemahan berbasis API.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk mengembangkan aplikasi penerjemahan teks Arab Pegon ke bahasa Indonesia berbasis Android. Penyelesaian masalah dilakukan melalui pendekatan bertahap, yaitu pengenalan teks Arab Pegon dari citra menggunakan *Optical Character Recognition* (OCR), dilanjutkan dengan proses transliterasi ke bahasa Jawa beraksara Latin, kemudian diterjemahkan ke bahasa Indonesia menggunakan layanan penerjemahan mesin. Pendekatan ini diharapkan mampu menjembatani keterbatasan layanan penerjemahan yang belum mendukung Arab Pegon secara langsung.

2. TINJAUAN PUSTAKA

Beberapa penelitian terdahulu telah membahas penggunaan Arab Pegon dalam konteks pendidikan

dan penerjemahan. Roidah et al. [1] mengkaji pemanfaatan terjemahan Arab Pegon sebagai media pendukung pembelajaran ilmu nahwu. Penelitian tersebut lebih menekankan peran Arab Pegon dalam membantu pemahaman struktur bahasa Arab, tanpa melibatkan pemanfaatan teknologi pengolahan teks otomatis.

Penelitian yang dilakukan oleh Anis [4] membahas proses penerjemahan teks Arab Pegon ke dalam bahasa Indonesia secara manual. Fokus penelitian ini terletak pada teknik transliterasi dan penafsiran makna yang dilakukan oleh penerjemah, sehingga belum mencakup pengembangan sistem atau aplikasi penerjemahan berbasis komputer.

Penelitian lainnya oleh Mardiansyah et al. [6] mengembangkan sistem transliterasi teks Arab Pegon ke dalam aksara Latin menggunakan pendekatan berbasis aturan. Penelitian ini hanya mencakup proses alih aksara tanpa melanjutkan ke tahap penerjemahan makna ke bahasa Indonesia.

Berbeda dengan penelitian-penelitian terdahulu tersebut, penelitian ini mengusulkan pengembangan aplikasi yang mengintegrasikan beberapa tahapan dalam satu sistem, meliputi pengenalan teks Arab Pegon dari citra menggunakan *Optical Character Recognition* (OCR), proses transliterasi ke bahasa Jawa beraksara Latin, serta translasi ke bahasa Indonesia. Pendekatan ini diharapkan dapat memberikan solusi yang lebih praktis dan mudah diakses oleh pengguna dalam memahami teks Arab Pegon.

2.1. Arab Pegon

Arab Pegon merupakan sistem penulisan bahasa Jawa yang menggunakan aksara Arab dengan penyesuaian tertentu untuk merepresentasikan bunyi bahasa Jawa [1]. Sistem ini berkembang di lingkungan pesantren dan banyak digunakan dalam penulisan kitab-kitab keagamaan sebagai sarana penjelasan terhadap teks berbahasa Arab [7] [8]. Meskipun menggunakan huruf Arab, Arab Pegon tidak dapat dipahami sebagai bahasa Arab karena struktur bahasa dan kosakatanya mengikuti kaidah bahasa Jawa. Perbedaan karakteristik ini menyebabkan Arab Pegon tidak dapat diperlakukan sama dengan bahasa Arab dalam proses penerjemahan mesin.

2.2. Optical Character Recognition (OCR)

Optical Character Recognition (OCR) memungkinkan sistem komputer mengenali karakter pada gambar dan mentransformasikannya ke dalam format teks digital [9] [10]. Dalam konteks penelitian ini, OCR berfungsi sebagai alat bantu untuk mengekstraksi teks Arab Pegon dari gambar sebelum dilakukan proses penerjemahan. Namun, akurasi OCR sangat dipengaruhi oleh jenis *font*, kualitas citra, serta karakteristik bahasa yang dikenali [11]. Karena Arab Pegon memiliki bentuk huruf dan penggunaan tanda baca yang berbeda dari bahasa Arab standar, hasil

OCR cenderung memerlukan koreksi manual sebelum digunakan pada tahap selanjutnya.

2.3. Tesseract OCR

Sebagai mesin OCR *open-source*, Tesseract banyak digunakan karena kemampuannya dalam mengenali berbagai bahasa dan fleksibilitas integrasinya pada berbagai platform [12]. Tesseract telah menyediakan model bahasa Arab, namun belum secara khusus mendukung Arab Pegon. Akibatnya, pengenalan teks Arab Pegon menggunakan Tesseract masih menghasilkan beberapa kesalahan karakter. Dalam penelitian ini, Tesseract OCR digunakan sebagai alat bantu awal untuk memperoleh teks yang kemudian dapat disunting dan diproses lebih lanjut dalam sistem penerjemahan.

2.4. Transliterasi

Transliterasi adalah proses pengalihan aksara dari satu sistem tulisan ke sistem tulisan lain tanpa mengubah bahasa yang digunakan [6]. Dalam penelitian ini, transliterasi berperan penting sebagai tahap perantara dalam proses penerjemahan Arab Pegon. Mengingat Arab Pegon merupakan bahasa Jawa yang ditulis dengan aksara Arab, transliterasi ke aksara Latin bahasa Jawa diperlukan agar teks dapat diproses oleh layanan penerjemahan mesin. Dengan adanya transliterasi, keterbatasan layanan penerjemahan yang tidak mendukung Arab Pegon dapat diatasi secara efektif.

2.5. Penerjemahan Mesin

Penerjemahan mesin merupakan proses penerjemahan bahasa secara otomatis dengan memanfaatkan teknologi komputasi [13]. Layanan penerjemahan mesin seperti Google Translate telah mendukung berbagai pasangan bahasa, termasuk bahasa Jawa dan Bahasa Indonesia. Namun, layanan tersebut belum menyediakan dukungan langsung untuk Arab Pegon. Oleh karena itu, pendekatan penerjemahan dalam penelitian ini dilakukan secara bertahap, yaitu dengan menerjemahkan teks hasil transliterasi bahasa Jawa ke Bahasa Indonesia. Pendekatan ini memungkinkan proses penerjemahan dilakukan dengan memanfaatkan dukungan bahasa yang tersedia secara teknis.

2.6. API AI

Application Programming Interface (API) *Artificial Intelligence* (AI) memungkinkan penulis untuk mengintegrasikan layanan kecerdasan buatan ke dalam aplikasi [14]. Dalam penelitian ini, proses penerjemahan dilakukan dengan memanfaatkan layanan penerjemahan berbasis AI yang diintegrasikan melalui ekstensi pihak ketiga, sehingga sistem dapat melakukan translasi teks secara otomatis tanpa membangun model penerjemahan secara mandiri.

2.7. Kodular

Kodular merupakan *platform* pengembangan aplikasi Android berbasis *visual programming* yang menggunakan metode *block-based programming* [15]. Platform ini memungkinkan pengembang untuk membangun aplikasi berbasis Android tanpa menuliskan kode secara langsung. Penggunaan Kodular dalam penelitian ini bertujuan untuk mempermudah pengembangan antarmuka pengguna serta integrasi aplikasi dengan layanan OCR dan API AI, sehingga proses pengembangan dapat dilakukan secara lebih cepat dan terstruktur.

3. METODE PENELITIAN

Penelitian ini merupakan penelitian terapan (*applied research*) yang bertujuan untuk mengembangkan dan menguji sistem penerjemahan teks Arab Pegon ke dalam bahasa Indonesia berbasis aplikasi Android. Pendekatan yang digunakan adalah pendekatan eksperimental, di mana sistem dirancang dan diimplementasikan untuk mengetahui kemampuannya dalam mengenali dan menerjemahkan teks Arab Pegon.

Pengembangan sistem difokuskan pada penerapan teknologi *Optical Character Recognition* (OCR) untuk mengekstraksi teks dari citra, transliterasi teks Arab Pegon ke dalam bahasa Jawa beraksara Latin, serta penerjemahan ke dalam bahasa Indonesia. Pendekatan ini dipilih karena layanan penerjemahan mesin yang tersedia belum menyediakan dukungan penerjemahan langsung dari Arab Pegon ke bahasa Indonesia.

3.1. Data dan Sumber Data

Data yang digunakan dalam penelitian ini berupa citra teks Arab Pegon yang diperoleh dari kitab cetak berbahasa Arab Pegon pada Kitab *Tafsir Al-Ibriz* [2]. Teks tersebut dipilih karena mencerminkan penggunaan Arab Pegon secara nyata dalam literatur keislaman dan masih banyak digunakan dalam kegiatan pembelajaran.

Citra teks digunakan sebagai masukan pada sistem OCR untuk menghasilkan teks digital. Selain data citra, penelitian ini juga memanfaatkan data teks hasil OCR yang dapat disunting secara manual oleh pengguna apabila ditemukan ketidaksesuaian hasil pengenalan karakter. Penggunaan data dibatasi pada teks Arab Pegon cetak dengan kualitas gambar yang baik untuk mendukung proses pengenalan karakter.

3.2. Tahapan Pengembangan Sistem

Tahapan pengembangan sistem dilakukan secara bertahap dan sistematis, mulai dari pengambilan citra hingga penampilan hasil terjemahan pada aplikasi Android. Tahapan tersebut meliputi

3.3. Pra-pemrosesan Citra

Setelah Citra teks Arab Pegon terlebih dahulu diproses untuk meningkatkan kualitas gambar sebelum dilakukan pengenalan karakter. Proses pra-

pemrosesan meliputi konversi citra ke dalam bentuk *grayscale* serta peningkatan kontras untuk membantu proses OCR.

3.4. Pengenalan Teks Menggunakan OCR

Proses pengenalan teks dilakukan menggunakan Tesseract OCR untuk mengonversi citra teks Arab Pegon menjadi teks digital. Mengingat karakter Arab Pegon memiliki perbedaan dengan bahasa Arab standar, hasil OCR pada tahap ini belum sepenuhnya akurat.

3.5. Penyuntingan Manual Hasil OCR

Sistem menyediakan fasilitas penyuntingan manual terhadap hasil OCR. Fitur ini bertujuan untuk memungkinkan pengguna memperbaiki kesalahan pengenalan karakter sebelum teks diproses lebih lanjut pada tahap transliterasi dan penerjemahan.

3.6. Transliterasi Arab Pegon ke Bahasa Jawa Latin

Teks Arab Pegon hasil OCR maupun *input manual* terlebih dahulu ditransliterasi ke dalam bahasa Jawa beraksara Latin. Tahapan ini dilakukan karena sistem penerjemahan mesin yang tersedia tidak menyediakan opsi penerjemahan langsung dari Arab Pegon ke bahasa Indonesia [16].

Tabel 1. Contoh pemetaan transliterasi Arab Pegon ke Bahasa Jawa Latin

Huruf Pegon	Posisi	Padanan Latin
ب	Awal	b
ن	Tengah	n
ي	Akhir	i
غ	Tengah	ng
ا	Awal	a
و	Awal	w
و	Tengah	u
او	Awal	u
و	harakat	a

Transliterasi dilakukan menggunakan pendekatan berbasis kamus (*dictionary-based transliteration*) yang disimpan dalam berkas JSON, di mana setiap karakter Arab Pegon dipetakan ke padanan huruf Latin bahasa Jawa seperti ditunjukkan pada Tabel 1. Berkas tersebut berisi pemetaan karakter Arab Pegon ke padanan huruf Latin bahasa Jawa berdasarkan posisi karakter dalam kata, yaitu posisi awal, tengah, dan akhir. Selain itu, kamus juga memuat pemetaan tanda harakat untuk menghasilkan vokal yang sesuai.

Teks Arab Pegon dipecah menjadi satuan karakter, kemudian setiap karakter diproses secara berurutan. Sistem menentukan posisi karakter dalam kata untuk memilih padanan transliterasi yang sesuai. Jika karakter memiliki tanda harakat, nilai vokal ditambahkan berdasarkan aturan transliterasi yang telah ditentukan. Hasil transliterasi setiap karakter kemudian digabungkan untuk membentuk kata dalam bahasa Jawa beraksara Latin.

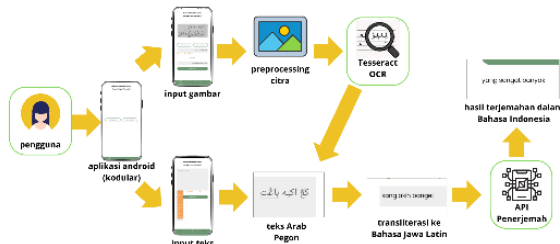
3.7. Penerjemahan ke Bahasa Indonesia

Teks hasil transliterasi bahasa Jawa Latin diterjemahkan ke dalam bahasa Indonesia menggunakan layanan penerjemahan berbasis API untuk menghasilkan terjemahan akhir yang dapat dipahami oleh pengguna.

3.8. Penampilan Hasil pada Aplikasi Android

Hasil OCR, transliterasi, dan terjemahan ditampilkan pada aplikasi Android berbasis Kodular sebagai antarmuka pengguna.

3.9. Arsitektur Sistem



Gambar 1. Arsitektur sistem

Arsitektur sistem dalam penelitian ini terdiri dari tiga komponen utama, yaitu aplikasi Android, server OCR, dan layanan penerjemahan. Aplikasi Android berfungsi sebagai antarmuka pengguna untuk mengambil gambar, menampilkan hasil OCR, serta menampilkan hasil translasi.

Server OCR dikembangkan menggunakan bahasa pemrograman Python dengan framework Flask. Server ini bertugas memproses citra teks menggunakan Tesseract OCR dan mengirimkan hasil pengenalan teks ke aplikasi Android. Proses komunikasi antara aplikasi dan server dilakukan melalui protokol HTTP.

3.10. Teknik Evaluasi

Evaluasi sistem pada penelitian ini dilakukan menggunakan metode pengujian *black-box*. Pengujian difokuskan pada fungsionalitas aplikasi berdasarkan masukan dan keluaran yang dihasilkan, tanpa menganalisis proses internal dari setiap komponen sistem.

Pengujian dilakukan dengan menguji fitur utama aplikasi, meliputi *input* teks Arab Pegon melalui citra (OCR), *input* teks Arab Pegon secara *manual*, proses transliterasi ke bahasa Jawa beraksara Latin, serta proses translasi ke bahasa Indonesia. Setiap fitur diuji untuk memastikan sistem dapat menghasilkan keluaran sesuai dengan tujuan yang diharapkan.

Hasil evaluasi digunakan untuk menilai apakah aplikasi telah berjalan sesuai dengan rancangan sistem dan mampu mendukung proses penerjemahan teks Arab Pegon ke bahasa Indonesia. Evaluasi ini tidak menitikberatkan pada pengukuran tingkat akurasi secara numerik, melainkan pada keberhasilan fungsi dan alur kerja sistem secara keseluruhan.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Pengembangan Sistem

Penelitian ini menghasilkan sebuah aplikasi Android yang berfungsi untuk membantu proses translasi teks Arab Pegon ke Bahasa Indonesia. Aplikasi menyediakan dua metode input, yaitu *input* gambar dan input teks manual. Pada metode *input* gambar, pengguna memilih citra teks Arab Pegon dari galeri perangkat, kemudian sistem melakukan proses pengenalan teks menggunakan teknologi OCR. Hasil OCR ditampilkan dalam bentuk teks yang dapat diedit secara manual oleh pengguna. Selanjutnya, teks yang telah diperbaiki dapat diproses ke tahap transliterasi dan penerjemahan.

Arsitektur sistem menggunakan pendekatan *client-server*, di mana aplikasi Android berperan sebagai *client*, sedangkan *server* digunakan untuk memproses OCR. Pendekatan ini memungkinkan pemrosesan citra dilakukan secara terpusat dan mempermudah integrasi dengan layanan pendukung lainnya.

4.2. Hasil Pengujian OCR Arab Pegon

Pengujian OCR dilakukan menggunakan beberapa sampel teks Arab Pegon yang diambil dari kitab *Tafsir al-Ibriz*. Pengujian difokuskan pada teks cetak dengan kualitas citra yang cukup baik.

Berdasarkan hasil pengujian, Tesseract OCR mampu mengenali sebagian besar karakter Arab Pegon, namun hasil teks yang dihasilkan belum sepenuhnya sesuai dengan teks asli pada citra. Beberapa karakter mengalami kesalahan pengenalan, seperti tertukarnya huruf, hilangnya harakat, serta munculnya karakter yang tidak relevan. Hal ini dipengaruhi oleh kualitas citra, ketajaman tulisan, serta keterbatasan model OCR dalam mengenali variasi penulisan Arab Pegon.



Gambar 2. Hasil OCR asli sebelum disunting

Gambar 2. menunjukkan contoh citra asli teks Arab Pegon yang digunakan sebagai masukan pada sistem, sedangkan Gambar 3. menampilkan hasil teks OCR setelah dilakukan penyuntingan secara manual oleh pengguna.



Gambar 3. Hasil OCR setelah disunting

Berdasarkan hasil pengujian, kesalahan pada tahap OCR berpengaruh langsung terhadap hasil transliterasi. Namun, dengan adanya fitur penyuntingan manual, pengguna dapat memperbaiki teks sebelum diproses lebih lanjut sehingga hasil transliterasi dan translasi menjadi lebih mudah dipahami.

4.3. Pengujian Proses Transliterasi dan Translasi pada Input OCR

Pengujian ini bertujuan untuk mengevaluasi hasil transliterasi dan translasi yang berasal dari teks Arab Pegon hasil *Optical Character Recognition* (OCR). Setelah proses OCR selesai, teks Arab Pegon ditampilkan pada kolom *input* aplikasi dan dapat disunting oleh pengguna apabila ditemukan kesalahan pengenalan karakter.

Transliterasi dari Arab Pegon ke Latin Jawa dilakukan secara otomatis ketika teks hasil OCR ditampilkan atau ketika pengguna melakukan perubahan pada teks tersebut. Hasil transliterasi kemudian ditampilkan pada kolom keluaran teks Latin. Selanjutnya, teks Latin Jawa tersebut diterjemahkan ke dalam bahasa Indonesia menggunakan layanan penerjemahan berbasis API.



Gambar 4. Hasil transliterasi dari teks Arab Pegon hasil OCR



Gambar 5. Hasil translasi Bahasa Indonesia dari teks hasil transliterasi OCR

Gambar 4. menunjukkan hasil transliterasi teks Arab Pegon ke Latin Jawa yang berasal dari input OCR, sedangkan Gambar 5. menampilkan hasil translasi ke dalam bahasa Indonesia.

Meskipun hasil OCR belum sepenuhnya akurat, sistem tetap memberikan keluaran teks yang dapat diedit oleh pengguna. Fitur penyuntingan manual ini memungkinkan pengguna untuk memperbaiki kesalahan pengenalan karakter sebelum proses selanjutnya dilakukan. Dengan demikian, keterbatasan akurasi OCR dapat diminimalkan melalui keterlibatan pengguna secara langsung.

4.4. Pengujian Transliterasi dan Translasi pada Input Manual

Selain menggunakan *input* OCR, sistem juga menyediakan fitur *input manual* berupa teks Arab Pegon yang diketik langsung oleh pengguna. Pengujian ini dilakukan untuk mengetahui kinerja transliterasi dan translasi tanpa dipengaruhi oleh kesalahan pengenalan karakter dari OCR.

Pada *input manual*, teks Arab Pegon yang dimasukkan akan langsung diproses oleh modul transliterasi secara otomatis. Setiap perubahan pada teks akan memicu proses transliterasi ulang sehingga hasil transliterasi selalu diperbarui secara *real-time*. Teks hasil transliterasi kemudian diterjemahkan ke dalam bahasa Indonesia melalui layanan

penerjemahan yang sama dengan pengujian sebelumnya.



Gambar 6. Hasil transliterasi dari input teks Arab Pegon manual



Gambar 7. Hasil translasi Bahasa Indonesia dari teks hasil transliterasi input manual

Gambar 6 menampilkan hasil transliterasi teks Arab Pegon yang dimasukkan secara manual, sedangkan Gambar 7 menunjukkan hasil translasi ke dalam bahasa Indonesia.

Hasil pengujian menunjukkan bahwa transliterasi dan translasi pada input manual memberikan hasil yang lebih konsisten dibandingkan input OCR. Hal ini disebabkan tidak adanya kesalahan pengenalan karakter, sehingga proses transliterasi berbasis kamus dapat bekerja secara optimal.

4.5. Perbandingan Hasil Input OCR dan Input Manual

Berdasarkan pengujian yang telah dilakukan, dapat disimpulkan bahwa penggunaan *input manual* menghasilkan tingkat keterbacaan transliterasi dan translasi yang lebih baik dibandingkan dengan input OCR. Meskipun demikian, fitur OCR tetap memiliki peran penting dalam membantu pengguna memperoleh teks awal dari citra kitab, terutama ketika teks Arab Pegon tidak tersedia dalam bentuk digital.

Dengan mengombinasikan OCR, penyuntingan manual, transliterasi otomatis, dan translasi bertahap, aplikasi PegonTranslate mampu memberikan solusi yang fleksibel bagi pengguna dalam memahami teks Arab Pegon

4.6. Hasil Pengujian Black-Box

Hasil pengujian aplikasi menggunakan metode *black-box* ditunjukkan pada Tabel 2. Pengujian dilakukan untuk memastikan setiap fungsi utama pada aplikasi dapat berjalan sesuai dengan kebutuhan

pengguna. Fokus pengujian diarahkan pada kesesuaian keluaran sistem terhadap masukan yang diberikan, tanpa memperhatikan proses *internal* dari sistem.

Tabel 2. Hasil Pengujian *Black-Box* Aplikasi PegonTranslate

No	Skenario Pengujian	Input	Output yang Diharapkan	Hasil
1	Input OCR	Citra Arab Pegon	Teks Arab Pegon tampil	Berhasil
2	Edit teks OCR	Teks disunting	Teks diperbarui	Berhasil
3	Transliterasi	Teks Arab Pegon	Teks Latin Jawa	Berhasil
4	Translasi	Teks Latin Jawa	Bahasa Indonesia	Berhasil
5	Input manual	Arab Pegon	Transliterasi dan Translasi	Berhasil

Secara keseluruhan, hasil pengujian black-box sebagaimana ditunjukkan pada Tabel 2 menunjukkan bahwa seluruh fungsi utama aplikasi telah berjalan sesuai dengan rancangan sistem dan mampu mendukung proses penerjemahan teks Arab Pegon ke bahasa Indonesia.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian dan pengujian yang telah dilakukan, dapat disimpulkan bahwa aplikasi PegonTranslate mampu menjalankan proses penerjemahan teks Arab Pegon ke bahasa Indonesia melalui tahapan pengenalan teks menggunakan *Optical Character Recognition* (OCR), transliterasi ke bahasa Jawa beraksara Latin, serta translasi ke bahasa Indonesia. Hasil pengujian *black-box* menunjukkan bahwa seluruh fungsi utama aplikasi dapat berjalan sesuai dengan rancangan sistem. Tesseract OCR mampu mengenali teks Arab Pegon dari citra, namun hasil pengenalan masih memiliki keterbatasan sehingga diperlukan penyuntingan *manual* sebelum proses transliterasi dilakukan. Pendekatan transliterasi berbasis kamus yang diterapkan dapat menghasilkan teks Latin Jawa yang dapat dipahami, terutama ketika *input* diperoleh secara manual atau setelah hasil OCR diperbaiki. Selanjutnya, proses translasi dari bahasa Jawa ke bahasa Indonesia dapat berjalan dengan baik menggunakan layanan penerjemahan yang tersedia. Untuk penelitian selanjutnya, disarankan dilakukan peningkatan kualitas pengenalan teks Arab Pegon, pengembangan aturan transliterasi yang lebih lengkap, serta pengembangan metode penerjemahan yang lebih sesuai dengan karakteristik bahasa Jawa agar hasil terjemahan menjadi lebih akurat dan alami.

DAFTAR PUSTAKA

- [1] S. Roidah, N. Izzah, and L. Munawaroh, "Analisis Penggunaan Terjemah Arab Pegon

- dalam Pembelajaran Nahwu Kitab Matan Al Jurumiyah Pondok Pesantren Insan Mulia Punggur Lampung Tengah,” *Al Maghazi Arab. Lang. High. Educ.*, vol. 2, no. 1, 2024.
- [2] E. Rachmawati, “Geneologi Tafsir Al- Ibrīz Karya KH. Bisri Musthofa,” *AYATUNA (Jurnal Ilmu Al-Qur’an dan Tafsir)*, no. 1, pp. 27–45, 2024.
- [3] J. I. Tofaynudin *et al.*, “Pendampingan Teras Belajar Pegon Bagi Santri untuk Membangun Literasi di Pondok Pesantren Mambaul Khoiriyatil Islamiyah Jember,” *J. Auladuna*, vol. 01, no. 01, pp. 193–202, 2022.
- [4] M. Y. Anis, “Teknik Amplifikasi dalam Penerjemahan Struktur Informasi Bahasa Jawa dengan Abjad Arab Pegon ke dalam Bahasa Indonesia,” *SALINGKA, Maj. Ilm. Bhs. dan Sastra*, vol. 20, no. 1, pp. 157–173, 2023.
- [5] Mas’udah, “Efektivitas Penggunaan Alat Instan dalam Penerjemahan Teks Bahasa Arab,” *Ilmuna J. Stud. Pendidik. Agama Islam*, vol. 7, no. 1, pp. 363–372, 2025.
- [6] D. Mardiansyah, Marlina, and L. Fitriyah, “Transliterasi Naskah Kitab Taudlihus Sholah,” *Titian Ilmu J. Ilm. Multi Sci.*, vol. 15, no. 1, 2023.
- [7] W. K. N. Fauzan, E. Supriyati, and T. Listyorini, “Inovasi Digital: Aplikasi Web untuk Pendataan dan Pelaporan Hafalan Kitab Kuning Santri,” *J. Digit*, vol. 15, no. 01, pp. 10–19, 2025.
- [8] I. Afifah and D. Sirojudin, “Efektivitas Arab Pegon dalam Pemahaman Kitab Kuning di Pesantren Darun Najah Malang,” *J. Educ. Manag. Stud.*, vol. 5, no. 6, pp. 41–45, 2022.
- [9] Y. Darmi, M. F. Sepriansyah, Y. Darnita, and Pahrizal, “Penerapan Metode Optical Character Recognition (OCR) untuk Mengidentifikasi Teks pada Identitas Dokumen Surat Izin Mengemudi (SIM),” vol. 9, no. 4, pp. 5992–5998, 2025.
- [10] S. M. Ritonga, D. Andini, I. Janurianty, and R. H. Nst, “Implementasi Optical Character Recognition dalam Pengolahan Data Statistik Survei Penduduk di Badan Pusat Statistik Kota Tebing Tinggi,” *J. Ris. Sist. Inf. dan Apl. Komputer(JRSIKOM)*, vol. 1, no. 4, pp. 28–33, 2025.
- [11] Z. A. Umar and M. Yunus, “Optimalisasi Teknologi OCR dalam Digitalisasi Manuskrip Hadis: Studi Akurasi dan Tantangan Linguistik Arab Klasik,” *Hamidah J. IlmuHadis*, vol. 1, no. 1, pp. 43–54, 2025.
- [12] A. Aprilino and I. H. Al Amin, “Implementasi Algoritma YOLO dan Tesseract OCR pada Sistem Deteksi Plat Nomor Otomatis,” *J. TEKNOINFO*, vol. 16, no. 1, pp. 54–59, 2022.
- [13] F. Razsiah, A. Josi, and S. Mubaroh, “Aplikasi Penerjemah Bahasa Bangka ke Bahasa Indonesia Menggunakan Neural Machine Translation Berbasis Website,” *J. Inov. Teknol. Terap.*, vol. 01, no. 1, 2023.
- [14] F. I. Muqsith, E. Supriyati, and T. Listyorini, “Klasifikasi Pengucapan Huruf Hijaiyah Berbasis Android Menggunakan CNN dengan Fitur Mel-Spectrogram,” *J. Inform. J. Pengemb. IT*, vol. 10, no. 1, pp. 67–78, 2025, doi: 10.30591/jpit.v10i1.8145.
- [15] Y. Kartika, Husnidar, and R. Hayati, “Pengembangan Aplikasi Pembelajaran Digital Berbasis Android Menggunakan Kodular pada Mata Kuliah Geometri,” *ASIMETRIS J. Pendidik. Mat. dan Sains*, vol. 04, no. 2, pp. 103–109, 2023.
- [16] T. Listyorini, Muljono, Purwanto, and R. S. Basuki, “Neural Approaches for Indonesian–Javanese Translation: A Comparison of RNN and Transformer Models,” *Int. Semin. Appl. Technol. Inf. Commun.*, pp. 575–584, Sep. 2025, doi: 10.1109/ISEMANTIC67418.2025.11291686.