

## Aplikasi Android untuk Alih Aksara Latin ke Arab Melayu dengan Pendekatan Berbasis Aturan

Yusra<sup>1</sup>, Muhammad Fikry<sup>2</sup>, Susmi Syahfrida Yani<sup>3</sup>, Muhammad Irsyad<sup>4</sup>, Suwanto Sanjaya<sup>5</sup>

<sup>1,2,3,4,5</sup>Teknik Informatika, UIN Sultan Syarif Kasim Riau

Jl. H.R. Soebrantas no. 155 KM. 18 Simpang Baru, Pekanbaru 28293

yusra@uin-suska.ac.id<sup>1</sup>, muhammad.fikry@uin-suska.ac.id<sup>2</sup>, susmi.syahfrida.yani@students.uin-suska.ac.id<sup>3</sup>, irsyadtech@uin-suska.ac.id<sup>4</sup>, suwantosanjaya@uin-suska.ac.id<sup>5</sup>

**Abstrak** – Bahasa Indonesia didasarkan dari bahasa Melayu yang dapat dituliskan dengan menggunakan aksara Latin dan aksara Arab Melayu. Saat ini, generasi muda di Riau lebih memahami penggunaan aksara Latin daripada aksara Arab Melayu. Meskipun mereka mendapatkan pelajaran menulis Arab Melayu yang terintegrasi di dalam kurikulum muatan lokal Budaya Melayu Riau (BMR), ketidaktahuan dan kurang pahamiannya aturan dalam menulis Arab Melayu dapat menyebabkan kesalahan penulisan. Untuk mencegah terjadinya kesalahan penulisan, digunakan aplikasi transliterasi (alih aksara). Transliterasi melakukan penyalinan dengan penggantian huruf dari abjad yang satu ke abjad yang lain. Aplikasi alih aksara dirancang berdasarkan aturan-aturan yang ada di buku Pedoman Umum Tulisan Arab Melayu, dan dibangun menggunakan bahasa pemrograman Java (Android). Pengujian dilakukan dengan membandingkan keluaran dari aplikasi terhadap hasil alih aksara oleh pakar aksara Arab Melayu. Akurasi dari hasil pengujian sebesar 95,5%. Persentase akurasi menunjukkan bahwa hasil validitasnya pada kriteria Sangat Valid.

**Kata Kunci** – algoritma, alih aksara, arab melayu, pemrosesan bahasa alami, pendekatan berbasis aturan

### PENDAHULUAN

Bahasa Indonesia tumbuh dan berkembang dari bahasa Melayu, salah satu bahasa dari rumpun bahasa Austronesia yang digunakan di Brunei, Indonesia, Malaysia, Singapura, dan sebagian dari Thailand. Di Indonesia, bahasa Melayu itu sendiri merupakan bahasa daerah yang memiliki banyak dialek dan tersebar di berbagai daerah. Sejak masuknya Islam ke Nusantara, orang Melayu menggunakan aksara Arab yang telah diubah untuk menuliskan bahasa Melayu, dikenal dengan aksara Arab Melayu atau Jawi. Aksara tersebut menggunakan huruf Arab asli dan huruf tambahan agar bunyi bacaannya sesuai dengan bahasa Melayu.

Pada masa penjajahan Belanda, penggunaan aksara Arab Melayu digantikan dengan aksara Latin.

Aksara Arab Melayu sangat penting keberadaannya dalam pengkajian khazanah intelektual Muslim Nusantara, sehingga sangat disayangkan aksara Arab Melayu berada dalam tahap yang mengkhawatirkan [1]. Disadari atau tidak, orang Melayu buta aksara miliknya sendiri. Penggunaan aksara Arab Melayu sudah jarang ditemukan dalam kegiatan sehari-hari. Generasi muda sebagai penerus keberadaan bahasa Melayu lebih memahami aksara Latin dibandingkan aksara Arab Melayu.

Sejumlah usaha telah dilakukan oleh Pemerintah Daerah di Provinsi Riau agar keberadaan aksara Arab Melayu sebagai salah satu identitas Melayu tetap terjaga secara turun temurun. Penulisan nama jalan dan nomenklatur di instansi pemerintah yang menyertakan tulisan Arab Melayu. Pengajaran baca tulis Arab Melayu dimasukkan ke dalam kurikulum pendidikan, menjadi bagian dari pelajaran muatan lokal Budaya Melayu Riau.

Namun demikian, usaha-usaha tersebut mengalami banyak kekurangan [2]. Adanya kesalahan penulisan pada papan nama jalan. Tidak semua kantor pemerintah yang nomenklaturnya menggunakan tulisan Arab Melayu, sementara kantor-kantor non pemerintah sama sekali tidak menggunakannya. Guru-guru muda yang mengajar tetapi tidak menguasai Arab Melayu dan belum mendapatkan pelatihan khusus. Masih sedikitnya tulisan dan karya sastra yang mengandung tulisan Arab Melayu, sehingga bahan bacaan siswa hanyalah buku pelajaran Budaya Melayu Riau. Berdasarkan hasil wawancara dengan salah seorang guru muatan lokal, diketahui bahwa siswa lebih lancar membaca tulisan Arab Melayu dibandingkan menuliskannya. Kesulitan yang dialami di antaranya adalah penulisan kata yang menggunakan huruf e lemah dan e keras, serta penggunaan huruf yang melambangkan vokal (huruf saksi).

Menurut Kamus Besar Bahasa Indonesia (KBBI), transliterasi (alih aksara) adalah penyalinan dengan penggantian huruf dari abjad yang satu ke abjad yang lain. Proses alih aksara dapat dikerjakan secara manual atau dibantu oleh komputer. Salah satu contohnya adalah Google Input Tools yang

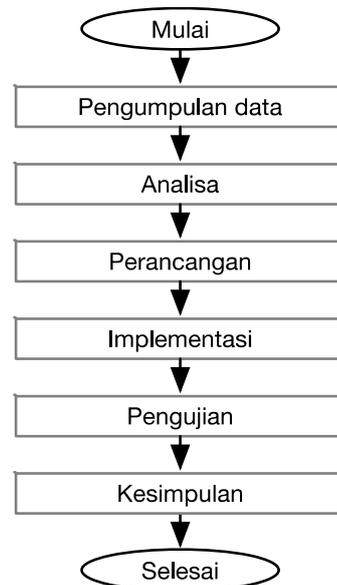
memungkinkan pengguna untuk mengetikkan teks dalam aksara Latin, kemudian secara otomatis memberikan rekomendasi transliterasi dalam aksara lain yang diinginkan berdasarkan kesamaan fonetik. Google Input Tools menyediakan berbagai pilihan aksara, tidak termasuk aksara Arab Melayu.

Oleh karena itu, dibutuhkan suatu perangkat lunak alih aksara yang memungkinkan penggantian teks bertulisan Latin ke dalam tulisan Arab Melayu. Dengan menggunakan perangkat lunak tersebut, dapat disediakan banyak bahan bacaan bertulisan Arab Melayu bagi masyarakat pada umumnya dan siswa pada khususnya. Selain itu, penggunaannya dapat mengetahui tulisan Arab Melayu yang benar, sehingga dapat dibandingkan dengan hasil tulisannya sendiri.

Alih aksara Latin ke Arab Melayu telah dilakukan oleh Mutmainnah [3] dan Sasongko [4]. Mutmainnah [3] melakukan alih aksara Latin ke Arab Melayu melalui pengenalan suku kata bahasa Indonesia dengan pendekatan Finite State Automata (FSA). Pemisahan kata menjadi suku kata didasarkan pada pola huruf yaitu Vokal (V) dan Konsonan (K). Suku kata diubah ke format Unicode sesuai kaidah dari aksara Arab Melayu, kemudian digabungkan kembali. Berdasarkan hasil pengujian terhadap 40 kata, diperoleh akurasi 92,5%. Sasongko [4] juga menggunakan pendekatan yang sama dengan menambahkan perbaikan terhadap kekurangan yang ditemukan ketika mengenali suku kata. Kaidah yang digunakan sebanyak 15 kaidah berdasarkan Rosadi dan Suhud [5] tahun 1960, sehingga belum mengikuti perkembangan bahasa Indonesia. Berdasarkan hasil pengujian terhadap 60 kata, diperoleh akurasi 94,74%.

Selain dua penelitian tersebut di atas, terdapat sejumlah penelitian terkait transliterasi (alih aksara) lainnya. Indrawan et al. [6] dan Kesiman et al. [7] [8] melakukan transliterasi Bali ke Latin, sementara Loekito et al. [9] melakukan kebalikannya. Quint [10] melakukan transliterasi bahasa Jepang ke Latin. Mathur dan Saxena [11] mengubah nama entitas dari bahasa Inggris ke fonem Hindi yang sebanding, sementara Kaur dan Goyal [12] mengubah nama entitas dari Punjabi ke bahasa Inggris. Ramakrishnan et al. [13] melakukan transliterasi dari tulisan Indic ke tulisan Kannada, sementara Raju et al. [14] dari Tamil ke Malayalam. Ahmadi [15] melakukan transliterasi tulisan Kurdish (dialek Sorani) berbasis Arab dan berbasis Latin dengan pendekatan berbasis aturan, Zahid et al. [16] melakukan transliterasi dari bahasa Inggris ke Urdu dengan algoritma Soundex, serta Mahdi dan Sfabakhsh [17] dari bahasa Inggris ke Persia dengan pendekatan *deep learning*.

Tahapan-tahapan yang dilakukan dapat dilihat pada Gambar 1.



Gambar 1. Tahapan penelitian

Pengumpulan data dilakukan untuk mendapatkan kata dasar bahasa Indonesia dari Kamus Besar Bahasa Indonesia (KBBI) IV dan kata serapan bahasa Arab yang diperoleh dari Pedoman Umum Ejaan Bahasa Indonesia (PUEBI). Selain itu, dipersiapkan kata uji sebanyak 200 kata yang akan dialih aksarakan oleh pakar aksara Arab Melayu.

Analisa alih aksara dilakukan berdasarkan aturan-aturan yang ada di buku Pedoman Umum Tulisan Arab Melayu [18]. Proses alih aksara menggunakan masukan berupa teks berbahasa Indonesia, melibatkan proses tokenisasi, *stemming*, pemenggalan kata menjadi suku kata, konversi ke Unicode, diakhiri dengan penggabungan suku kata menjadi kata.

Untuk keperluan pengujian, dirancang bangun aplikasi berbasis *web* beserta basis datanya. Selain itu, dirancang bangun aplikasi berbasis Android yang dapat digunakan untuk keperluan alih aksara. Pengujian dilakukan dengan melibatkan Drs. Fadillah Om, seorang pakar aksara Arab Melayu sekaligus penulis buku Pedoman Umum Tulisan Arab Melayu [18]. Setelah pengujian, dilakukan perhitungan akurasi alih aksara dan analisa terhadap hasil pengujian. Dari hasil persentase keberhasilan alih aksara, dapat ditentukan kriteria validitas sebagaimana diperlihatkan pada Tabel 1. Pada akhirnya, dilakukan pengambilan kesimpulan.

Tabel 1. Kategori nilai validitas

Rentang Nilai	Kriteria
80% < x ≤ 100%	Sangat valid

## METODE PENELITIAN

$60\% < x \leq 80\%$	Valid
$40\% < x \leq 60\%$	Cukup valid
$20\% < x \leq 40\%$	Kurang valid
$0\% < x \leq 20\%$	Tidak valid

### HASIL DAN PEMBAHASAN

Data yang berhasil dikumpulkan meliputi kata dasar bahasa Indonesia yang diperoleh dari KBBI IV sebanyak 29.934 kata, serta kata serapan bahasa Arab yang diperoleh dari Pedoman Umum Ejaan Bahasa Indonesia sebanyak 329 kata. Data kata dasar yang mengandung huruf e, ditandai apakah menggunakan e lemah atau e keras. Data kata serapan meliputi tulisan Latin dan Arab-nya. Data-data tersebut disimpan ke dalam basis data yang telah dibuat.

Analisa alih aksara dari Latin ke Arab Melayu didasarkan pada kaidah penulisan Pedoman Umum Tulisan Arab Melayu tahun 2011 [18]. Huruf Arab Melayu yang digunakan meliputi 40 huruf, meliputi a, b, c, d, e, f, g, h (ada dua), ch, i, j, k, kh, l, m, n, ng, ny, o, p, q, r, s, sy, t, u, v, w, x, y, z, a`, a (hamzah), dh, dz, sh, ts, th, dan zh. Setiap huruf, angka dan karakter lainnya memiliki representasi karakter Unicode yang bersesuaian. Kaidah penulisan sebanyak 11 aturan yang secara garis besar yaitu :

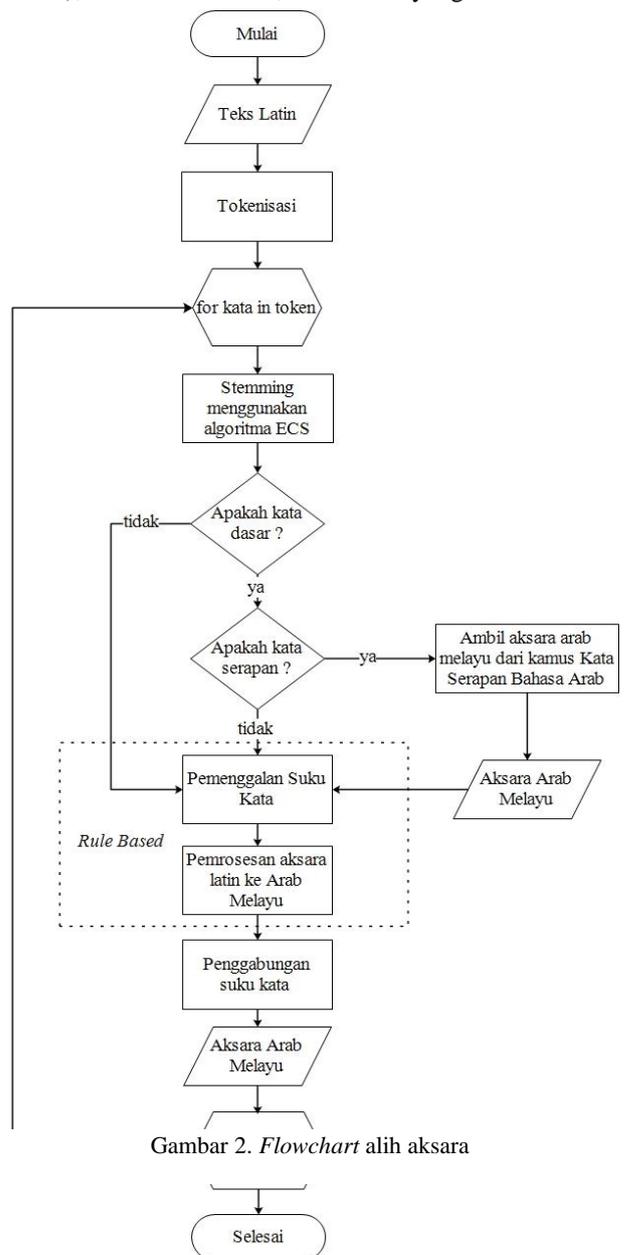
1. Tulisan Arab Melayu memakai huruf Arab, ditulis dari kanan ke kiri dan tidak memakai huruf kapital
2. Huruf tulisan Arab Melayu, meliputi huruf vokal dan huruf konsonan.
3. Perlambangan angka.
4. Penulisan huruf tulisan Arab Melayu
  - a. Huruf yang tidak dapat disambung
  - b. Huruf yang dapat disambung sebelah kanan saja
  - c. Huruf yang dapat disambung sebelah kanan dan kiri
5. Penulisan huruf saksi, Huruf saksi ditulis sesudah huruf konsonan, meliputi saksi a, i, u, e keras dan o.
6. Penulisan kata
  - a. Kata bersuku kata satu
  - b. Kata bersuku kata dua
  - c. Kata bersuku kata tiga, empat, lima dan seterusnya
7. Pemakaian ة (hamzah)
8. Pengecualian
  - a. Kata yang berasal dari bahasa Arab
  - b. Kata yang tidak ditulis menurut keadaan tulisan Arab Melayu
9. Menulis huruf k (ك , ق) pada kata
10. Huruf yang tidak dipakai pada penulisan tulisan Arab Melayu
11. Menulis kata berimbuhan
  - a. Kata dasar yang berimbuhan kata depan, awal, akhiran dan kata

ganti penulisan pada saksi a mendapat perubahan.

- b. Kata dasar berimbuhan dengan partikel dan kata ganti milik tidak mendapat perubahan penulisan kata dasar.

Flowchart alih aksara diperlihatkan pada Gambar 2.

Pengujian dilakukan terhadap 200 kata uji yang sudah dialih aksarakan oleh seorang pakar aksara Arab Melayu, yaitu Drs. Fadillah Om. Kata uji meliputi kata dasar (baik kata serapan maupun tidak), kata berimbuhan, serta kata yang memiliki



Gambar 2. Flowchart alih aksara

gabungan konsonan maupun diftong. Aplikasi pengujiannya dirancang bangun berbasis web, menggunakan bahasa pemrograman PHP dan basis

data MySQL. Tampilan aplikasi diperlihatkan pada Gambar 3.

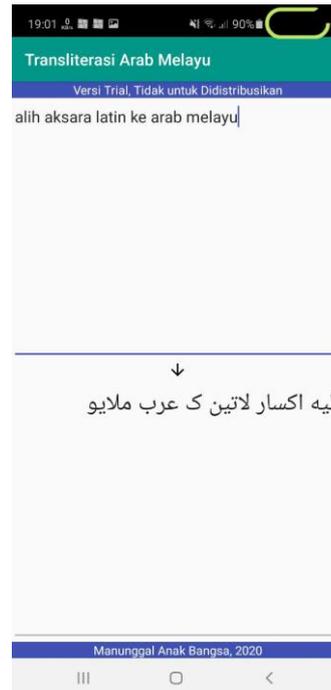
#	Kata	Suku Kata (oleh Manusia)	Suku Kata (oleh Algoritma)	Status	Transliterasi (oleh Manusia)	Transliterasi (oleh Algoritma)
1	allah	kata serapan	kata serapan	-	الله	الله
2	muhammad	kata serapan	kata serapan	-	محمد	محمد
3	susmi	sus-mi	sus-mi	berhasil	سوسمي	سوسمي
4	syahfrida	syah-fri-da	syah-fri-da	berhasil	شهفريد	شهفريد
5	yani	ya-ni	ya-ni	berhasil	ياني	ياني

Gambar 3. Tampilan aplikasi berbasis web untuk pengujian

Berdasarkan hasil pengujian, diperoleh 196 kata yang berhasil dilakukan pemenggalan suku katanya. Kata-kata yang tidak sesuai suku katanya gagal dialih aksarakan. Berdasarkan 196 kata tersebut, diperoleh 191 kata yang berhasil dialih aksarakan dengan benar. Dengan demikian diperoleh akurasi sebesar 95,5%. Persentase akurasi menunjukkan bahwa hasil validitasnya pada kriteria Sangat Valid.

Berdasarkan analisa terhadap hasil pengujian, pada proses alih aksara ditemukan kata yang bentuk tulisan Arab Melayu dan latinnya berbeda. Hal ini karena satu aksara Arab Melayu melambangkan beberapa aksara Latin, yaitu huruf ya ( ي ) yang melambangkan huruf e, i, y dan huruf wau ( و ) yang melambangkan huruf o, u, w. Contohnya aksara Arab Melayu نيتاق dapat dibaca huruf latinnya "patin" atau "paten".

Untuk memudahkan pengguna melakukan alih aksara, dirancang bangun aplikasi berbasis Android sebagaimana diperlihatkan pada Gambar 4.



Gambar 4. Tampilan aplikasi Android

## KESIMPULAN DAN SARAN

Dalam penelitian ini, diperlihatkan bahwa aplikasi alih aksara berhasil dirancang bangun berdasarkan Pedoman Umum Tulisan Arab Melayu. Akurasi dari hasil pengujian sebesar 95,5% dengan hasil validitas pada kriteria Sangat Valid. Alih aksara Latin ke Arab Melayu dengan pendekatan berbasis aturan yang telah diimplementasikan baik berbasis web maupun berbasis Android diharapkan dapat bermanfaat untuk penelitian selanjutnya.

Kedepannya, penelitian ini dapat dilanjutkan dengan melakukan alih aksara dari Arab Melayu ke Latin. Masukan ke aplikasi tidak hanya teks, tetapi berbentuk gambar (foto). Hal ini akan mempermudah masyarakat untuk membaca tulisan Arab Melayu. Pada akhirnya, aplikasi ini diharapkan bermanfaat dalam pengkajian sejarah dan budaya Melayu dimana sebagian sumbernya ditulis dengan aksara Arab Melayu.

## UCAPAN TERIMA KASIH

Kami mengucapkan terima kasih kepada Drs. Fadillah Om, penulis buku Pedoman Umum Tulisan Arab Melayu untuk Pendidikan Dasar, atas bantuan dan dukungannya dalam rancang bangun aplikasi Android ini.

REFERENSI

- [1] E. Roza, "Aksara Arab-Melayu di Nusantara dan Sumbangsihnya dalam Pengembangan Khazanah Intelektual," *Tsaqafah*, vol. 13, no. 1, 2017.
- [2] K. Masrohanti, "Arab Melayu Riwayatmu Kini," *Swarnadwipa*, Pekanbaru, Indonesia, pp. 6–18, 2016.
- [3] A. Mutmainnah, "Pengenalan Suku Kata Bahasa Indonesia Menggunakan Finite State Automata untuk Alih Aksara Indonesia – Aksara Jawi," *J. Sist. dan Teknol. Inf.*, vol. 3, no. 3, 2015.
- [4] A. Sasongko, "Meningkatkan Akurasi Alih Aksara Alfabet - Jawi Bahasa Indonesia Menggunakan Pendekatan FSA," *J. Inform. UBSI*, vol. 3, no. September, pp. 259–268, 2016.
- [5] A. Rosadi and M. Suhud, *Tjara menulis huruf Arab Melayu*. Bandung: Peladjar, 1960.
- [6] G. Indrawan, Sariyasa, and I. K. Paramarta, "A New Method of Latin-To-Balinese Script Transliteration based on Bali Simbar Font," in *2019 Fourth International Conference on Informatics and Computing (ICIC)*, 2019, pp. 1–6.
- [7] M. W. A. Kesiman, J. Burie, and J. Ogier, "A Complete Scheme of Spatially Categorized Glyph Recognition for the Transliteration of Balinese Palm Leaf Manuscripts," in *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, 2017, vol. 01, pp. 125–130.
- [8] M. W. A. Kesiman, J. Burie, J. Ogier, and P. Grangé, "Training Schemes for the Transliteration of the Balinese Script Into the Latin Script on Palm Leaf Manuscript Images," in *2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, 2018, pp. 297–302.
- [9] L. H. Loekito, G. Indrawan, Sariyasa, and I. K. Paramarta, "Error Analysis of Latin-to-Balinese Script Transliteration Method Based on Noto Sans Balinese Font BT - 3rd International Conference on Innovative Research Across Disciplines (ICIRAD 2019)," 2020, pp. 335–340.
- [10] J. Quint, "Automatic Japanese transliteration with a formalism for presyntactic analysis," in *International Conference on Natural Language Processing and Knowledge Engineering, 2003. Proceedings. 2003*, 2003, pp. 512–518.
- [11] S. Mathur and V. P. Saxena, "Hybrid approach to English-Hindi name entity transliteration," in *2014 IEEE Students' Conference on Electrical, Electronics and Computer Science*, 2014, pp. 1–5.
- [12] A. Kaur and V. Goyal, "Punjabi to English Machine Transliteration for Proper Nouns," in *2018 3rd International Conference On Internet of Things: Smart Innovation and Usages (IoT-SIU)*, 2018, pp. 1–7.
- [13] A. G. Ramakrishnan, R. D. Sequiera, S. S. Rao, and Shiva Kumar H R, "Transliteration of Indic languages to Kannada with a user-friendly interface," in *2015 IEEE International Advance Computing Conference (IACC)*, 2015, pp. 998–1001.
- [14] K. Raju, T. V Sreerekha, P. V Vidya, R. R. Rajeev, and P. C. Reghu Raj, "Tamil to Malayalam Transliteration," in *2015 Fifth International Conference on Advances in Computing and Communications (ICACC)*, 2015, pp. 12–15.
- [15] S. Ahmadi, "A Rule-Based Kurdish Text Transliteration System," *ACM Trans. Asian Low-Resour. Lang. Inf. Process.*, vol. 18, no. 2, Jan. 2019.
- [16] M. A. Zahid, N. I. Rao, and A. M. Siddiqui, "English to Urdu transliteration: An application of Soundex algorithm," in *2010 International Conference on Information and Emerging Technologies*, 2010, pp. 1–5.
- [17] M. Mahdi Mahsuli and R. Safabakhsh, "English to Persian transliteration using attention-based approach in deep learning," in *2017 Iranian Conference on Electrical Engineering (ICEE)*, 2017, pp. 174–178.
- [18] F. Om, *Pedoman Umum Tulisan Arab Melayu*, Revisi. Pekanbaru: Usaha Putra Riau, 2011.