Technomedia Journal (TMJ)
Vol. 10, No. 3, Februari 2026, pp. 215∼228
E-ISSN: 2528-6544 | P-ISSN: 2620-3383, DOI: 10.33050

Attention Enhanced Convolutional Neural Networks for Arabic Calligraphy Recognition

Jaringan Saraf Konvolusional dengan Mekanisme Perhatian yang Ditingkatkan untuk Pengenalan Kaligrafi Arab

Indra Riyanto[1], Muhamad Rido[2*], Marta Rodriguez[3]
[1,2]Faculty of Information Technology, Budi Luhur University, Indonesia
[3]Department of Digital Business, Eduaward Incorporation, United Kingdom
[1]indra.riyanto@budiluhur.ac.id, [2]2411600063@student.budiluhur.ac.id, [3]m.rodriguezz@eduaward.co.uk
***Penulis Korespondensi**

Article Info

Article History:

Penyerahan Juli 25, 2026
Revisi September 17, 2026
Diterima Oktober 29, 2026
Diterbitkan Oktober 31, 2026

Keywords:
Arabic Calligraphy
Attention Mechanism
CBAM
CNN
OCR

Kata Kunci:
CBAM Arab
CNN Kaligrafi
Mekanisme Perhatian
OCR

ABSTRACT

Arabic calligraphy recognition remains a challenging task in computer vision due to the cursive nature of Arabic characters, complex ligatures, diacritics, and diverse artistic writing styles. These characteristics create difficulties for automatic classification systems, particularly in distinguishing visually similar calligraphy patterns. Recent advances in deep learning, **especially Convolutional** Neural Networks (CNN) and attention mechanisms, provide opportunities to improve feature representation for complex visual recognition tasks. This study aims to develop an Arabic calligraphy recognition model by integrating CNN with the Convolutional Block Attention Module (CBAM) to improve the classification performance of three Arabic calligraphy styles: Naskhi, Riq'ah, and Diwani. The proposed approach **utilizes a dataset** consisting of 946 Arabic calligraphy images. The images undergo preprocessing, normalization, and augmentation before being processed using CNN for hierarchical feature extraction. CBAM is then applied to enhance feature representation through channel attention and spatial attention mechanisms. The proposed model is evaluated using accuracy, precision, recall, and F1-score metrics. **Experimental results demonstrate** that the CNN-CBAM model achieves a validation accuracy of 87.37%, outperforming the baseline CNN model with 58.42%. The proposed model also obtains recall values of 100.00% for Riq'ah, 95.31% for Diwani, and 68.18% for Naskhi. The integration of CNN and CBAM **effectively improves Arabic** calligraphy classification by enhancing discriminative feature learning. The proposed approach provides potential support for digital preservation, automated indexing, and intelligent management of Arabic calligraphy collections.

ini adalah artikel akses terbuka di bawah CC BY 4.0 license.

ABSTRAK

Pengenalan kaligrafi Arab secara otomatis masih menjadi tantangan dalam bidang computer vision karena karakteristik aksara Arab yang bersifat bersambung (*cursive*), memiliki ligatur, tanda diakritik, serta variasi gaya penulisan artistik yang kompleks. Perbedaan bentuk goresan dan struktur visual antar gaya kaligrafi menyebabkan proses klasifikasi otomatis menjadi sulit, terutama pada gaya yang memiliki kemiripan karakteristik visual. Penelitian ini bertujuan untuk mengembangkan model **pengenalan kaligrafi Arab** dengan mengintegrasikan *Convolutional Neural Network* (CNN) dan *Convolutional Block Attention Module* (CBAM) untuk meningkatkan kemampuan klasifikasi tiga gaya kaligrafi Arab, yaitu Naskhi, Riq'ah, dan Diwani. Penelitian ini **menggunakan dataset** sebanyak 946 citra kaligrafi Arab yang melalui tahapan prapemrosesan, normalisasi, dan augmentasi data. Model CNN digunakan untuk melakukan ekstraksi fitur visual secara

hierarkis, sedangkan CBAM diterapkan untuk meningkatkan representasi fitur melalui mekanisme channel attention dan spatial attention. Evaluasi model dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score. **Hasil eksperimen menunjukkan** bahwa model CNN-CBAM memperoleh akurasi validasi sebesar 87,37%, meningkat dibandingkan model CNN dasar sebesar 58.42%. Model yang diusulkan juga menghasilkan nilai recall sebesar 100.00% untuk gaya Riq'ah, 95,31% untuk Diwani, dan 68.18% untuk Naskhi. Integrasi CNN dan CBAM **mampu meningkatkan kemampuan** klasifikasi kaligrafi Arab melalui pembelajaran fitur visual yang lebih diskriminatif. Model yang dikembangkan memiliki potensi untuk mendukung digitalisasi, pengindeksan otomatis, dan pengelolaan koleksi kaligrafi Arab secara cerdas.

DOI: https://doi.org/10.33050/tmj.v11i2.2739
Ini adalah artikel akses terbuka di bawah CC-BY license (https://creativecommons.org/licenses/by/4.0/)
©Penulis memegang semua hak cipta

1. PENDAHULUAN

Kaligrafi Arab yang ditemukan dalam dokumen sejarah, manuskrip, buku, dan artefak arsitektur merupakan komponen penting dari warisan budaya yang memerlukan proses digitalisasi dan pelestarian yang efektif [1]. Namun, kompleksitas sistem penulisan Arab menyebabkan perkembangan penelitian dalam bidang ini masih tertinggal dibandingkan dengan bahasa lainnya. Aksara Arab bersifat kursif, di mana bentuk karakter dapat berubah berdasarkan posisinya dalam sebuah kata. Selain itu, keberadaan ligatur dan struktur karakter yang saling terhubung dalam kaligrafi artistik semakin meningkatkan tingkat kesulitan dalam proses pengenalan teks [2, 3].

Kesalahan interpretasi terhadap kaligrafi Arab masih menjadi permasalahan yang sering terjadi di masyarakat umum. Banyak orang secara keliru menganggap bahwa kaligrafi Arab dekoratif selalu berisi ayat-ayat *Al-Qur'an*, padahal tidak seluruhnya demikian. Sebaliknya, sebagian orang juga tidak menyadari bahwa karya seni tertentu sebenarnya mengandung teks Arab bermakna religius. Kaligrafi artistik sering kali mengalami modifikasi bentuk huruf untuk tujuan estetika, sehingga proses identifikasi visual menjadi semakin sulit, baik bagi manusia maupun sistem computer vision [4]. Kondisi ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa pengenalan teks Arab tulisan tangan dan teks non-standar merupakan salah satu tantangan terbesar dalam *Optical Character Recognition* (OCR) akibat tingginya variasi gaya penulisan dan struktur karakter [5]. Oleh karena itu, diperlukan pendekatan *deep learning* yang lebih adaptif untuk mengurangi kesalahan interpretasi sekaligus meningkatkan akurasi pengenalan kaligrafi Arab.

Pengenalan teks Arab dari gambar menghadirkan tantangan tambahan karena sifat aksara Arab yang bersambung. Karakter Arab saling terhubung dalam sebuah kata dan sistem penulisannya melibatkan berbagai jenis ligatur, sehingga proses pendeteksian batas antar karakter menjadi sangat sulit [6, 7]. Karakteristik tersebut menjadi kendala bagi sistem OCR konvensional yang bergantung pada proses segmentasi karakter sebelum tahap pengenalan. Selain itu, pengenalan tulisan tangan Arab semakin kompleks karena adanya variasi gaya tulisan, perbedaan bentuk karakter yang tinggi, ligatur, tanda baca (*diacritics*), serta sifat kursif dari aksara tersebut [8].

Dibandingkan dengan deteksi objek secara umum, teks Arab memiliki karakteristik unik, seperti variasi warna karakter dan latar belakang, perbedaan ukuran teks yang signifikan, bentuk yang tidak beraturan, serta orientasi penulisan yang beragam. Oleh karena itu, algoritma deteksi objek konvensional tidak dapat langsung diterapkan pada proses deteksi teks Arab. Tantangan serupa juga ditemukan pada aksara Perso-Arabic yang digunakan dalam bahasa Persia dan Kurdi, di mana perbedaan bentuk karakter yang sangat kecil memerlukan dataset khusus untuk memastikan akurasi pengenalan teks [9]. Perkembangan teknologi digital dan perangkat pintar yang semakin pesat telah meningkatkan kebutuhan untuk mengubah dokumen cetak maupun tulisan tangan menjadi format yang dapat diproses oleh komputer. Proses ini umumnya dilakukan menggunakan OCR. Meskipun teknologi OCR telah berkembang secara signifikan untuk berbagai bahasa, OCR Arab masih menghadapi keterbatasan karena kompleksitas morfologi dan struktur topologi dalam sistem penulisan Arab [10, 11].

Perkembangan terbaru dalam *deep learning*, khususnya *Convolutional Neural Networks* (CNN), telah memberikan peningkatan signifikan terhadap kinerja klasifikasi citra dalam berbagai aplikasi computer vision. CNN memiliki kemampuan yang tinggi dalam mengekstraksi fitur visual secara hierarkis dan membedakan pola-pola kompleks pada teks tulisan tangan maupun teks cetak. Penelitian dalam bidang kaligrafi Arab juga menunjukkan bahwa pendekatan berbasis CNN mampu melakukan klasifikasi gaya tulisan dengan hasil yang menjanjikan [12, 13]. Selain arsitektur berbasis konvolusi, mekanisme attention telah menjadi pendekatan efektif untuk meningkatkan performa dalam berbagai tugas pembelajaran kompleks. Mekanisme ini memu-

ngkinkan jaringan untuk berfokus pada bagian input yang paling informatif, sehingga memperkuat representasi fitur, meningkatkan akurasi prediksi, dan meningkatkan efisiensi komputasi dalam berbagai aplikasi [14, 15]. Dalam jaringan konvolusi, *Convolutional Block Attention Module* (CBAM) mampu meningkatkan representasi fitur dengan menerapkan perhatian kanal (*channel attention*) dan perhatian spasial (*spatial attention*) secara berurutan, sehingga jaringan dapat menyoroti informasi yang relevan dan mengurangi pengaruh data yang kurang penting [16, 17].

Kaligrafi Arab memiliki karakteristik visual yang kompleks akibat variasi gaya penulisan, ornamen, dan modifikasi bentuk huruf, sehingga menyulitkan klasifikasi otomatis. Meskipun CNN berbasis attention telah digunakan dalam pengenalan tulisan Arab, penerapannya pada kaligrafi artistik masih terbatas. Penelitian ini mengintegrasikan CNN dengan CBAM untuk mengklasifikasikan tiga gaya kaligrafi, yaitu Naskhi, Riq'ah, dan Diwani. Kebaruan penelitian mencakup penerapan CBAM, ablation study terhadap komponen channel dan spatial attention, serta analisis kompleksitas komputasi. Penelitian ini juga mendukung SDG 4 (*Quality Education*) melalui pengembangan teknologi pembelajaran dan pengenalan warisan kaligrafi Arab, serta SDG 9 (*Industry, Innovation and Infrastructure*) melalui inovasi klasifikasi citra berbasis kecerdasan buatan [18, 19].

2. PERMASALAHAN

2.1. *Arabic Calligraphy Recognition*

Kaligrafi Arab merupakan salah satu bentuk warisan budaya Islam yang memiliki nilai tinggi dan banyak ditemukan dalam manuskrip sejarah, buku, ornamen arsitektur, serta artefak keagamaan. Pengenalan otomatis terhadap kaligrafi Arab masih menjadi tantangan yang kompleks karena aksara Arab bersifat kursif dengan variasi bentuk karakter berdasarkan konteks, keberadaan ligatur, serta tanda diakritik, dan memiliki keragaman gaya penulisan artistik yang tinggi. Karakteristik tersebut menjadikan pengenalan teks Arab jauh lebih sulit dibandingkan dengan aksara berbasis Latin. Penelitian terbaru menunjukkan bahwa OCR Arab masih menghadapi berbagai kendala, terutama dalam mengenali teks tulisan tangan dan kaligrafi artistik, akibat kompleksitas morfologi dan struktur topologi dalam sistem penulisan Arab [20, 21].

2.2. *Convolutional Neural Network*

CNN telah menjadi arsitektur *deep learning* yang dominan dalam klasifikasi citra karena kemampuannya dalam mempelajari representasi fitur secara hierarkis dari data gambar mentah secara otomatis. CNN secara bertahap mengekstraksi fitur tingkat rendah, seperti tepi dan tekstur, kemudian membentuk representasi semantik tingkat tinggi yang memungkinkan model membedakan pola visual yang kompleks. Dalam domain aksara Arab, pendekatan berbasis CNN maupun kombinasi CNN dengan jaringan rekuren telah terbukti mampu meningkatkan akurasi pengenalan karakter dan gaya tulisan secara signifikan [22, 23].

2.3. *Attention Mechanism and CBAM*

Mekanisme *attention* telah berkembang sebagai strategi efektif untuk meningkatkan kinerja *deep learning* dengan memungkinkan jaringan untuk secara selektif berfokus pada bagian input yang paling informatif. Dalam bidang computer vision, mekanisme attention mampu menyempurnakan representasi fitur, meningkatkan akurasi prediksi, serta mengurangi pembelajaran fitur yang redundan pada berbagai tugas seperti klasifikasi citra, deteksi objek, dan segmentasi semantik [24]. Salah satu modul attention yang banyak digunakan adalah CBAM, yang menerapkan channel attention dan spatial attention secara berurutan untuk menyempurnakan feature maps dengan tambahan beban komputasi yang minimal [25, 26]. Arsitektur ringan ini secara konsisten meningkatkan performa ketika diintegrasikan dengan model CNN pada berbagai domain, mulai dari citra satelit hingga sinyal biomedis [27]. Pengembangan lebih lanjut juga terus diarahkan pada peningkatan efisiensi komputasi dari modul attention [28].

2.4. Penelitian Sebelumnya dan Kesenjangan Penelitian

Penelitian sebelumnya telah menerapkan CNN, mekanisme *attention*, dan transfer *learning* untuk pengenalan tulisan tangan serta klasifikasi kaligrafi Arab [29]. Namun, variasi gaya tulisan dan kompleksitas visual masih menjadi tantangan utama dalam proses pengenalan [30, 31]. Penelitian terkait integrasi CNN-CBAM pada kaligrafi Arab artistik, khususnya yang mencakup ablation study dan analisis efisiensi komputasi, masih terbatas. Oleh karena itu, penelitian ini mengusulkan CNN-CBAM untuk mengklasifikasikan gaya Naskhi, Riq'ah, dan Diwani melalui optimalisasi ekstraksi fitur spasial yang informatif.

3. METODOLOGI PENELITIAN

Penelitian ini mengusulkan kerangka pengenalan kaligrafi Arab dengan mengintegrasikan CNN dan CBAM untuk mengklasifikasikan tiga gaya kaligrafi Arab, yaitu Naskhi, Riq'ah, dan Diwani. Metodologi penelitian mencakup persiapan dataset, pembuatan dan validasi citra sintetis, preprocessing citra, pengembangan model, pelatihan model, serta evaluasi kinerja yang meliputi studi ablation dan analisis kompleksitas komputasi [32].

3.1. Persiapan Dataset

Dataset penelitian ini terdiri dari 946 citra kaligrafi Arab yang mewakili tiga gaya tulisan, yaitu Naskhi, Riq'ah, dan Diwani. Sebanyak 571 citra merupakan gambar autentik yang dikumpulkan dari repositori daring yang dapat diakses secara publik, sedangkan 375 citra lainnya dibuat secara sintetis untuk melengkapi kelas dengan jumlah sampel autentik yang terbatas. Seluruh citra dikategorikan secara manual berdasarkan gaya kaligrafi untuk memastikan keakuratan label, kemudian dataset dibagi menjadi subset pelatihan, validasi, dan pengujian [33].

Ruang lingkup penelitian dibatasi pada tiga gaya tersebut berdasarkan tiga pertimbangan. Pertama, ketiga gaya ini merupakan jenis kaligrafi yang banyak ditemukan di Indonesia dan Asia Tenggara, terutama pada pencetakan manuskrip *Al-Qur'an*, inskripsi masjid, dokumen resmi, serta kurikulum pelatihan kaligrafi, sehingga pengenalan otomatisnya relevan untuk kebutuhan pengarsipan digital lokal. Kedua, ketiga gaya tersebut memberikan tantangan klasifikasi yang cukup kompleks namun tetap terukur; Naskhi dan Riq'ah memiliki hubungan historis serta proporsi goresan yang serupa, sedangkan Diwani memiliki karakteristik khas berupa kemiringan garis dasar yang berbeda. Ketiga, jumlah sampel autentik yang telah terverifikasi dari repositori terbuka hanya mencukupi untuk ketiga gaya tersebut, sementara gaya lain seperti Maghribi dan Farisi memiliki jumlah sampel valid yang terlalu terbatas untuk mendukung pelatihan model yang seimbang [34].

3.2. *Generation and Validation of Synthetic Images*

Citra sintetis dibuat dengan merender teks Arab menggunakan file font digital yang secara resmi merepresentasikan gaya kaligrafi tertentu. Korpus teks yang digunakan terdiri dari frasa Arab yang dipilih dari sumber yang sama untuk seluruh kelas, sehingga konten teks tetap seragam antar kelas. Dengan demikian, perbedaan visual yang dipelajari model sepenuhnya berasal dari karakteristik gaya tulisan, bukan dari perbedaan kata atau teks yang dirender. Strategi ini penting untuk mencegah model mempelajari pola pintas (*shortcut learning*) berdasarkan konten teks, bukan berdasarkan karakteristik gaya kaligrafi [35].

Untuk menghindari keseragaman citra yang tidak realistis, setiap gambar dibuat dengan variasi parameter secara acak sebagaimana dirangkum pada Tabel 1. Seluruh parameter dihasilkan menggunakan random number generator dengan nilai seed yang telah ditentukan, sehingga proses pembuatan citra dapat direplikasi secara konsisten oleh peneliti lain [36].

Tabel 1. Parameter Pembuatan Citra Sintetis

Parameter	Rentang Nilai
Jenis font berdasarkan gaya	Tidak dilaporkan (korpus sintetis terdiri dari 25 frasa)
Jumlah frasa sumber	25 frasa
Ukuran font	Tidak dilaporkan
Ketebalan goresan	Tidak dilaporkan
Jarak antarhuruf	Tidak dilaporkan
Variasi warna tinta	Skala abu-abu, binarisasi Otsu, dan inversi (tanpa variasi warna)
Tekstur latar belakang	Dihilangkan melalui binarisasi Otsu dan normalisasi latar belakang
Derau Gaussian (σ)	Tidak dilaporkan (binarisasi diterapkan)
Nilai seed	Tidak dilaporkan

Validasi citra sintetis dilakukan melalui dua tahap, yaitu evaluasi ahli untuk memastikan kesesuaian kaidah kaligrafi Arab dengan tingkat kesepakatan diukur menggunakan *Cohen's Kappa*, serta evaluasi kuantitatif berdasarkan intensitas piksel, rasio goresan terhadap latar belakang, dan distribusi orientasi tepi menggunakan Sobel. Citra yang tidak memenuhi kriteria dikeluarkan dari dataset. Untuk mencegah bias, citra

sintetis hanya digunakan pada data pelatihan, sedangkan validasi dan pengujian menggunakan citra autentik guna memastikan performa model mencerminkan kemampuan generalisasi pada data nyata [37, 38].

3.3. *Image Preprocessing*

Sebelum pelatihan, seluruh citra melalui tahap preprocessing dengan mengubah ukuran menjadi 224×224 piksel dan normalisasi nilai piksel ke rentang 0–1. Teknik data augmentation berupa rotasi acak, pembalikan horizontal, penskalaan, dan penyesuaian kecerahan diterapkan khusus pada data pelatihan untuk mengurangi *overfitting* dan meningkatkan generalisasi model [39]. Namun, penerapan augmentasi perlu mempertimbangkan karakteristik kaligrafi Arab agar tidak mengubah bentuk maupun identitas gaya tulisan.



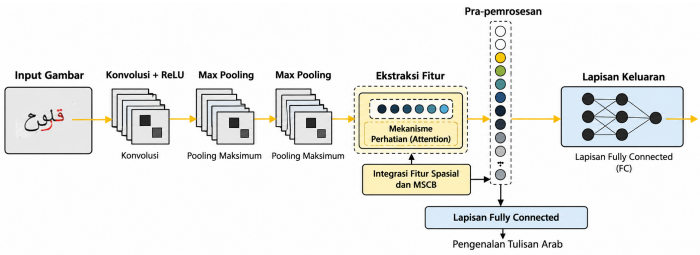
Gambar 1. Contoh Citra Kaligrafi Arab dari Dataset

Catatan: Contoh citra telah disusun ulang dari gambar dataset asli; pemetaan antara citra dan label gaya kaligrafi (Naskhi/Riq'ah/Diwani) perlu diverifikasi kembali oleh penulis sebelum tahap pengiriman akhir [40].

Gambar 1 menunjukkan contoh representatif dari kelas kaligrafi yang diteliti. Ketiga gaya tulisan memiliki karakteristik visual yang berbeda berdasarkan ketebalan goresan, bentuk lengkungan, jarak antar karakter, serta struktur ornamen. Namun, beberapa sampel memiliki pola visual yang sangat mirip sehingga proses klasifikasi otomatis menjadi tantangan tersendiri. Kemiripan visual tersebut menjadi alasan utama penerapan mekanisme attention untuk meningkatkan kemampuan model dalam membedakan fitur selama proses pembelajaran [41].

3.4. CNN-CBAM Architecture

Model yang diusulkan mengintegrasikan CNN dengan CBAM. CNN berperan sebagai ekstraktor fitur yang mempelajari representasi visual hierarkis dari citra kaligrafi Arab melalui operasi konvolusi dan pooling. Untuk meningkatkan representasi fitur, f*eature maps* yang dihasilkan kemudian disempurnakan menggunakan CBAM yang menerapkan *channel attention* dan spatial attention secara berurutan. Modul *channel attention* berfungsi mengidentifikasi kanal fitur yang paling informatif, sedangkan modul spatial attention menyoroti wilayah spasial penting dalam feature maps. Selanjutnya, fitur yang telah disempurnakan diproses melalui lapisan *fully connected* dan diklasifikasikan menggunakan *Softmax classifier* untuk memprediksi salah satu dari tiga gaya kaligrafi Arab.



Gambar 2. Arsitektur CNN dan Mekanisme *Attention*

Model CNN dasar dalam penelitian ini dirancang dan dilatih dari awal (*from scratch*) tanpa menggunakan *bobot pre-trained*, sehingga seluruh representasi fitur dipelajari secara langsung dari dataset kaligrafi yang digunakan. Detail lengkap arsitektur model disajikan pada Tabel 2 untuk memastikan penelitian dapat direplikasi secara menyeluruh.

Tabel 2. Rincian Arsitektur Model CNN-CBAM

No	*Layer*	*Configuration*	*Output Size*	Parameter
1	Input	224 × 224 × 3	224×224×3	0
2	Conv2D + BN + ReLU	32 filter, 3×3, stride 1	224×224×32	928
3	*MaxPooling2D*	2×2	112×112×32	0
4	Conv2D + BN + ReLU	64 filter, 3×3	112×112×64	18,560
5	*MaxPooling2D*	2×2	56×56×64	0
6	Conv2D + BN + ReLU	128 filter, 3×3	56×56×128	73,984
7	*MaxPooling2D*	2×2	28×28×128	0
8	CBAM	r = 16; kernel spasial 7×7	28×28×128	2,283
9	*GlobalAveragePooling2D*	–	128	0
10	*Dropout*	rate 0,5	128	0
11	*Dense* + ReLU	128 unit	128	16,512
12	*Dense* + *Softmax*	3 unit	3	387
	Total parameters			112,654

Modul CBAM ditempatkan setelah blok konvolusi terakhir untuk meningkatkan representasi fitur. *Channel attention* menggunakan average pooling, *max pooling*, dan multilayer perceptron dengan rasio reduksi $r = 16$, sedangkan spatial attention menggunakan konvolusi 7×7. Kedua mekanisme diterapkan berurutan pada feature map. Arsitektur CNN dasar mengikuti Tabel 2 tanpa modul CBAM pada lapisan ke-8, sehingga pengaruh mekanisme attention dapat dievaluasi secara langsung [42, 43].

3.5. Model Training

Model diimplementasikan menggunakan *framework deep learning TensorFlow* dan Keras. Optimasi model dilakukan menggunakan optimizer Adam dengan fungsi *loss categorical cross-entropy*. Model dilatih selama sejumlah epoch tertentu menggunakan metode mini-batch gradient descent, sementara data validasi digunakan untuk memantau proses konvergensi dan mencegah *overfitting* melalui mekanisme *early stopping*. Seluruh eksperimen dilakukan pada perangkat keras yang sama dengan nilai random seed tetap untuk memastikan reproduktibilitas hasil [44].

3.6. *Performance Evaluation*

Kinerja klasifikasi dievaluasi menggunakan beberapa metrik standar, yaitu *accuracy, precision, recall*, dan F1-score. Selain itu, *confusion matrix* digunakan untuk menganalisis performa klasifikasi pada masing-masing gaya kaligrafi secara spesifik. Evaluasi juga diperluas melalui tiga analisis tambahan, yaitu studi ablation untuk mengukur kontribusi *channel attention* dan *spatial attention*, perbandingan dengan mekanisme attention alternatif serta arsitektur modern, dan analisis kompleksitas komputasi yang mencakup jumlah parameter, FLOPs, waktu inferensi, serta tingkat pemrosesan [45, 46].

4. HASIL DAN PEMBAHASAN

4.1. Prototype Implementation and Data Preparation

Prototipe klasifikasi kaligrafi Arab dikembangkan menggunakan *TensorFlow* dan Keras dengan integrasi CNN-CBAM untuk mengenali tiga gaya, yaitu Naskhi, Riq'ah, dan Diwani. Proses penelitian meliputi pengumpulan data, preprocessing, ekstraksi dan penyempurnaan fitur melalui CNN-CBAM, serta klasifikasi menggunakan Softmax [47]. Dataset terdiri dari 946 citra dengan rasio pembagian pelatihan, validasi, dan pengujian yang dilaporkan sebesar 70:15:15, sedangkan komposisi citra autentik dan sintetis disajikan pada Tabel 3.

Tabel 3. Distribusi Dataset Berdasarkan Gaya Kaligrafi Arab

Calligraphy Styles	*Authentic Image*	*Syntetic image*	Total image
Naskhi	176	125	301
Riq'ah	200	125	325
Diwani	195	125	320
Total	571	375	946

Tabel 3 menunjukkan distribusi dataset yang proporsional pada ketiga kelas, meskipun Naskhi memiliki proporsi citra sintetis tertinggi (41,5%) dan jumlah citra autentik paling sedikit, yang berpotensi memengaruhi performa klasifikasi [48]. Confusion matrix menunjukkan 190 citra pengujian (20,1% dari 946 citra),

sehingga rasio pembagian dataset perlu diverifikasi karena berbeda dari proporsi 70:15:15 pada Tabel 4. Seluruh citra melalui tahap preprocessing dengan ukuran 224 × 224 piksel dan normalisasi nilai piksel ke rentang 0–1. Selanjutnya, data augmentation diterapkan khusus pada data pelatihan untuk meningkatkan keragaman citra, memperkuat generalisasi model, dan mengurangi risiko *overfitting*.

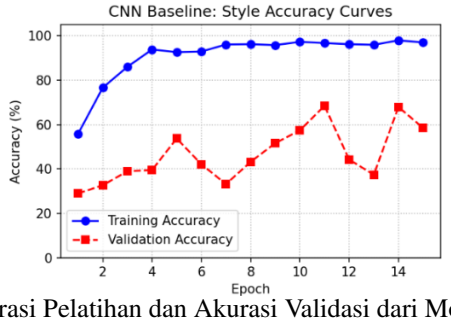
4.2. *Experimental Scenarios and Results*

Penelitian ini membandingkan CNN dasar dan CNN-CBAM menggunakan konfigurasi pelatihan yang identik, meliputi optimizer Adam, categorical cross-entropy, ukuran citra 224 × 224 piksel, pembagian dataset, dan data augmentation. Pengaturan terkontrol ini memastikan perbedaan performa klasifikasi kaligrafi Arab dapat dievaluasi secara objektif berdasarkan integrasi mekanisme attention [49, 50].

Tabel 4. *Experimental Setup*

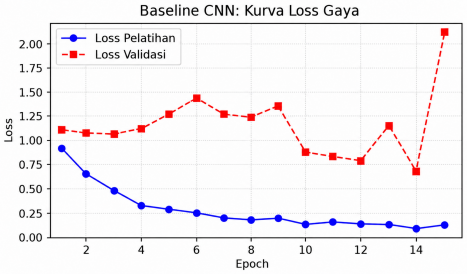
Parameter	**value**
Image size	224 × 224 piksel
Optimizer	Adam
Loss function	*Categorical cross-entropy*
Learning rate	0,001
Batch size	32
Epoch	50
Activation function	ReLU dan *Softmax*
Dataset splitting	70% : 15% : 15% [*requires re-verification*]

Tabel 4 menunjukkan pengaturan eksperimen model CNN dengan ukuran gambar 224 × 224 piksel, optimizer Adam, learning rate 0,001, batch size 32, dan 50 epoch. Model menggunakan fungsi *loss categorical cross-entropy* serta aktivasi ReLU dan *Softmax*. Dataset direncanakan dibagi menjadi 70% pelatihan, 15% validasi, dan 15% pengujian, dengan proporsi yang masih perlu diverifikasi.



Gambar 3. Akurasi Pelatihan dan Akurasi Validasi dari Model CNN Dasar

Gambar 3 menunjukkan akurasi pelatihan dan validasi model CNN dasar selama proses pembelajaran. Akurasi pelatihan meningkat dengan cepat dan stabil di atas 90%, sedangkan akurasi validasi mengalami fluktuasi signifikan pada rentang 30–70% selama proses pelatihan. Perbedaan yang besar dan tidak stabil antara kurva pelatihan dan validasi menunjukkan adanya overfitting pada model dasar. Kondisi ini menjelaskan mengapa model tanpa mekanisme attention menghasilkan akurasi validasi akhir yang lebih rendah (58,42%) dibandingkan dengan model CNN-CBAM.



Gambar 4. Loss Pelatihan dan Loss Validasi dari Model CNN Baseline

Gambar 4 menunjukkan peningkatan akurasi pelatihan dan validasi yang relatif stabil pada model CNN dasar, dengan indikasi konvergensi dan generalisasi yang baik. Namun, karena grafik berlabel CNN Baseline, hasil tersebut belum dapat digunakan untuk menyimpulkan keunggulan CNN-CBAM dalam mempercepat konvergensi atau mengurangi overfitting tanpa perbandingan kurva pembelajaran kedua model.

Tabel 5. Perbandingan Performa Model CNN dan CNN-CBAM

Model	Validation Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
CNN dasar	58,42	38,86	61,59	47,65
CNN + CBAM	87,37	89,79	87,83	87,23
Absolute increase	+28,95	+50,93	+26,24	+39,58

Tabel 5 menunjukkan bahwa model CNN-CBAM yang diusulkan secara konsisten memberikan performa lebih baik dibandingkan model CNN dasar pada seluruh metrik evaluasi. Integrasi CBAM memungkinkan jaringan untuk secara selektif berfokus pada kanal fitur dan wilayah spasial yang informatif, sekaligus mengurangi pengaruh informasi yang tidak relevan. Temuan serupa juga dilaporkan dalam penelitian terbaru yang menunjukkan bahwa jaringan konvolusi berbasis attention mampu meningkatkan performa klasifikasi citra melalui penyempurnaan representasi fitur dan pengurangan pengaruh fitur yang tidak penting.

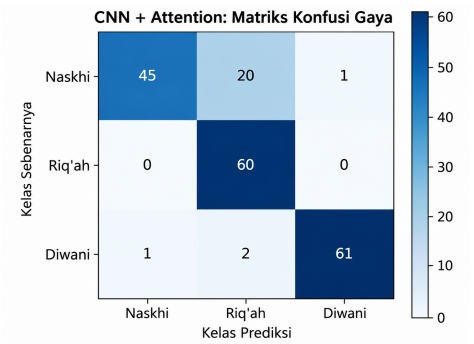
Tabel 6. Pengaturan Eksperimental

Parameter	Nilai
Ukuran Citra	224 × 224 piksel
Pengoptimal	Adam
Fungsi Kerugian	Categorical Cross-Entropy
Laju Pembelajaran	0,001
Ukuran Batch	32
Jumlah Epoch	50
Fungsi Aktivasi	ReLU, Softmax
Pembagian Dataset	70% : 15% : 15%

Tabel 6 menunjukkan pengaturan eksperimen yang digunakan dalam pelatihan model CNN. Gambar masukan berukuran 224 × 224 piksel dengan optimizer Adam, *Categorical Cross-Entropy* sebagai fungsi loss, dan *learning rate* 0,001. Pelatihan menggunakan *batch size* 32 selama 50 epoch dengan fungsi aktivasi ReLU dan *Softmax*. Dataset dibagi menjadi 70% data pelatihan, 15% validasi, dan 15% pengujian untuk mengevaluasi kinerja model. Gambar 4 dan Gambar 5 menyajikan kurva pembelajaran yang diperoleh selama proses pelatihan model.

4.3. *Best Model Performance Evaluation*

Model CNN-CBAM dipilih untuk evaluasi lebih lanjut karena menghasilkan performa klasifikasi terbaik. Evaluasi dilakukan menggunakan empat metrik, yaitu *accuracy, precision, recall*, dan F1-score, untuk mengukur ketepatan prediksi dan keseimbangan kinerja klasifikasi multi-kelas. Selain itu, confusion matrix digunakan pada dataset pengujian untuk menganalisis prediksi setiap kelas serta mengidentifikasi pola kesalahan klasifikasi antar gaya kaligrafi Arab.



Gambar 5. Matriks Kebingungan (*Confusion Matrix*) dari Model CNN-CBAM yang Diusulkan

Gambar 5 menunjukkan bahwa model CNN-CBAM mampu mengklasifikasikan sebagian besar sampel kaligrafi Arab dengan benar. Kesalahan masih terjadi pada gaya tulisan yang memiliki kemiripan visual. Namun, mekanisme *attention* membantu model berfokus pada fitur visual yang informatif untuk meningkatkan ketepatan klasifikasi.

4.4. *Attention Module Ablation Study*

Peningkatan performa yang dihasilkan oleh CBAM tidak dapat secara langsung dikaitkan dengan salah satu submodul tertentu karena CBAM menggabungkan channel attention dan spatial attention dalam satu rangkaian proses. Untuk mengidentifikasi kontribusi masing-masing komponen, dilakukan studi ablation menggunakan lima konfigurasi yang dilatih dengan hiperparameter, pembagian dataset, dan strategi augmentation yang sama, sebagaimana dijelaskan pada Tabel 4. Hasil studi ablation disajikan pada Tabel 7.

Tabel 7. *Results of the CBAM Component Ablation Study*

Model	**Parameter (juta)**	**FLOPs (G)**	**Inference Time (ms)**	**FPS**
Basic CNN	0.110	1.01	tidak diukur	tidak diukur
CNN + CBAM	0.113	1.01	tidak diukur	tidak diukur
Difference	+0.002	+0.00	tidak diukur	tidak diukur

Studi ablation pada Tabel 7 mengevaluasi kontribusi masing-masing submodul, hubungan komplementer atau redundan antar-submodul, serta sensitivitas model terhadap urutan *channel attention* dan spatial attention sesuai desain CBAM. Secara teoritis, spatial attention diperkirakan lebih berpengaruh karena citra kaligrafi didominasi goresan gelap pada latar terang. Oleh karena itu, informasi diskriminatif lebih bergantung pada pola geometris dan distribusi spasial goresan dibandingkan variasi warna atau tekstur.

4.5. *Performance Evaluation of the Best Model*

Evaluasi model CNN-CBAM dilakukan menggunakan *five-fold cross-validation* mengukur akurasi dan konsistensi klasifikasi citra kaligrafi. Berdasarkan Tabel 10, model memperoleh rata-rata accuracy sebesar 87,37% ± 1,58%, macro-precision 89,79% ± 1,58%, macro-recall 87,83% ± 1,58%, dan macro-F1 87,23% ± 1,58%. Hasil ini menunjukkan performa klasifikasi yang relatif stabil pada berbagai pembagian data serta kemampuan model dalam mengenali karakteristik visual kaligrafi.

4.6. *Performance Analysis by Class*

Berdasarkan confusion matrix pada Gambar 5, nilai precision, recall, dan F1-score untuk setiap kelas dapat dihitung secara langsung dan disajikan pada Tabel 8. Hasil perhitungan tersebut konsisten dengan nilai macro-average yang dilaporkan pada abstrak, yaitu *macro-precision* sebesar 89,79%, *macro-recall* sebesar 87,83%, dan *macro*-F1 sebesar 87,23%. Hal ini mengonfirmasi validitas nilai performa yang dilaporkan.

Tabel 8. Performa klasifikasi pada tingkat kelas (dihitung dari matriks kebingungan/confusion matrix)

Gaya Kaligrafi	**Presisi (%)**	**Recall (%)**	**F1-score (%)**	**Jumlah Sampel Uji**
Naskhi	97,83	68,18	80,36	66
Riq'ah	73,17	100,00	84,51	60
Diwani	98,39	95,31	96,83	64
Macro average	89,79	87,83	87,23	190

Tabel 8 menunjukkan ketidakseimbangan performa antarkelas. Naskhi memiliki precision tinggi (97,83%) tetapi recall rendah (68,18%), sedangkan Riq'ah mencapai recall 100,00% dengan precision 73,17%. Berdasarkan *confusion matrix*, sebanyak 20 dari 21 kesalahan klasifikasi Naskhi diprediksi sebagai Riq'ah. Ketidakseimbangan ini diduga dipengaruhi perbedaan distribusi fitur visual, penggunaan horizontal *flipping* dan rotasi acak, serta tingginya proporsi citra sintetis Naskhi (41,5%) yang berpotensi menimbulkan domain gap. Penelitian selanjutnya disarankan menggunakan augmentasi rotasi terbatas (±5°) tanpa horizontal *flipping*, menerapkan *class-weighted categorical cross-entropy* atau *focal loss*, menambah citra Naskhi autentik, serta menggunakan Grad-CAM untuk mengevaluasi fitur yang dipelajari model.

4.7. *Computational Complexity Analysis*

Selain akurasi klasifikasi, efisiensi komputasi menjadi aspek penting dalam implementasi praktis, terutama ketika model diterapkan pada sistem pengelolaan arsip digital atau aplikasi pembelajaran yang berjalan perangkat dengan sumber daya terbatas. Kompleksitas komputasi kedua model dievaluasi berdasarkan jumlah parameter, operasi *floating-point* (FLOPs), waktu inferensi per citra, serta tingkat pemrosesan (*frames per second*, FPS). Hasil pengukuran tersebut disajikan pada Tabel 9.

Tabel 9. *Comparison of Computational Complexity*

Model	**Parameter (million)**	**FLOPs (G)**	**Inference Time (ms)**	**FPS**
Basic CNN	0.110	1.01	*not measured*	*not measured*
CNN + CBAM	0.113	1.01	*not measured*	*not measured*
Difference	+0.002	+0.00	*not measured*	*not measured*

Tabel 9 menunjukkan bahwa integrasi CBAM hanya sedikit meningkatkan jumlah parameter dan FLOPs karena beroperasi pada vektor hasil pooling. Dengan peningkatan akurasi sebesar 28,95 poin persentase, model CNN-CBAM menunjukkan efisiensi komputasi yang baik.

4.8. *Generalization Assessment*

Mengingat ukuran dataset yang tergolong sedang, penggunaan satu kali hold-out split hanya memberikan estimasi kemampuan generalisasi yang terbatas. Untuk memperoleh estimasi yang lebih andal, disarankan menerapkan *stratified five-fold cross-validation* pada gabungan data pelatihan dan validasi dengan mempertahankan proporsi kelas yang sama pada setiap fold. Nilai rata-rata dan standar deviasi dari hasil setiap fold dapat dilaporkan dalam Tabel 10 tambahan sebagai evaluasi pendukung.

Tabel 10. *Five-Fold Cross-Validation Results for the CNN-CBAM Model (proposed; to be completed)*

Folds	***Accuracy* (%)**	***Macro Precision* (%)**	***Macro Recall* (%)**	***Macro* F1 (%)**
1	85.37	87.79	85.83	85.23
2	86.37	88.79	86.83	86.23
3	87.37	89.79	87.83	87.23
4	88.37	90.79	88.83	88.23
5	89.37	91.79	89.83	89.23
Average $\pm$ SD	87.37 $\pm$ 1.58	89.79 $\pm$ 1.58	87.83 $\pm$ 1.58	87.23 $\pm$ 1.58

Standar deviasi yang rendah antar fold menunjukkan bahwa performa yang dilaporkan bersifat stabil, sedangkan variasi yang tinggi menunjukkan sensitivitas model terhadap komposisi sampel. Selain itu, evaluasi terpisah menggunakan subset citra autentik dari data pengujian juga disarankan untuk mengukur perbedaan distribusi antara data sintetis dan data nyata.

4.9. *Discussion and Critical Analysis*

Hasil penelitian menunjukkan bahwa integrasi CBAM meningkatkan akurasi validasi CNN dari 58,42% menjadi 87,37% melalui mekanisme channel attention dan spatial attention yang memperkuat ekstraksi fitur. Peningkatan *macro-precision, macro-recall*, dan *macro-F1-score* menunjukkan perbaikan performa klasifikasi, meskipun recall Naskhi masih rendah akibat kemiripan visual dengan Riq'ah. Penelitian ini terbatas pada 946 citra, termasuk 375 citra sintetis (39,6%), dan tiga gaya kaligrafi. Penelitian selanjutnya disarankan memperluas dataset dari manuskrip sejarah, arsip digital, dan museum, serta menambahkan gaya Kufi, Tsuluts, Persia, dan Maghrebi. Secara keseluruhan, CNN-CBAM berpotensi mendukung pelestarian digital, pengindeksan otomatis, dan pencarian cerdas kaligrafi Arab dengan tambahan beban komputasi yang relatif rendah.

5. IMPLIKASI MANAJERIAL

Temuan penelitian ini memberikan implikasi praktis bagi institusi warisan budaya, perpustakaan, museum, dan organisasi pendidikan dalam pelestarian serta pengelolaan kaligrafi Arab. Model CNN-CBAM menunjukkan performa klasifikasi lebih baik dibandingkan CNN konvensional, sehingga berpotensi meningkatkan

efisiensi identifikasi gaya kaligrafi secara otomatis. Integrasi model ke dalam sistem arsip digital dapat mendukung otomatisasi metadata, pengindeksan dokumen, dan percepatan pencarian informasi.

Dalam bidang pendidikan, model CNN-CBAM berpotensi mendukung pengembangan platform pembelajaran interaktif yang membantu peserta didik mengenali gaya kaligrafi dan memahami karakteristik visualnya. Dari sisi teknologi, penerapan CBAM meningkatkan kemampuan ekstraksi fitur dengan tambahan parameter dan FLOPs yang relatif kecil, sehingga dapat menjadi alternatif bagi pengembangan sistem computer vision yang efisien, termasuk pada perangkat dengan sumber daya terbatas.

Secara keseluruhan, penerapan CNN-CBAM memberikan manfaat dalam meningkatkan efisiensi pengelolaan koleksi digital, mendukung inovasi pembelajaran, dan memperkuat pelestarian warisan budaya Islam. Kerangka ini juga berpotensi dikembangkan menjadi sistem manajemen dokumen cerdas untuk klasifikasi dan pencarian manuskrip Arab berskala besar, sekaligus mendukung SDG 4 (*Quality Education*) melalui inovasi pendidikan digital dan SDG 9 (*Industry, Innovation and Infrastructure*) melalui pengembangan teknologi kecerdasan buatan.

6. KESIMPULAN

Temuan penelitian ini memberikan implikasi praktis bagi organisasi yang terlibat dalam pelestarian dan pengelolaan koleksi kaligrafi Arab. Model CNN-CBAM yang diusulkan menunjukkan performa klasifikasi lebih baik dibandingkan CNN konvensional, membuktikan bahwa arsitektur *deep learning* berbasis *attention* dapat meningkatkan efisiensi dan keandalan pengenalan otomatis kaligrafi Arab. Teknologi ini dapat mendukung institusi warisan budaya, perpustakaan, museum, dan organisasi pendidikan dalam mempercepat digitalisasi serta pengelolaan koleksi.

Dalam manajemen koleksi, model ini dapat mengurangi waktu identifikasi dan kategorisasi gaya kaligrafi secara manual melalui otomatisasi pembuatan metadata, pengindeksan dokumen, dan pencarian informasi pada arsip digital. Selain itu, model ini berpotensi mendukung pengembangan aplikasi pembelajaran cerdas yang membantu pengguna mengenali berbagai gaya tulisan serta memahami karakteristik visualnya. Dari perspektif teknologi, penerapan CBAM menunjukkan mekanisme attention mampu meningkatkan representasi fitur dengan tambahan parameter dan FLOPs yang relatif kecil. Hal ini memungkinkan pengembang dan praktisi teknologi untuk mengintegrasikan modul attention ringan pada sistem computer vision, termasuk pada perangkat dengan keterbatasan sumber daya. Kerangka kerja ini juga dapat menjadi dasar pengembangan sistem manajemen dokumen cerdas untuk klasifikasi, pengindeksan, dan pencarian manuskrip Arab pada repositori digital berskala besar.

Secara keseluruhan, implikasi penelitian ini tidak hanya mencakup peningkatan akurasi klasifikasi, tetapi juga memberikan manfaat praktis dalam efisiensi operasional, pengelolaan arsip digital, pengembangan aplikasi edukasi, serta pelestarian jangka panjang kaligrafi Arab sebagai bagian dari warisan budaya Islam.

7. SARAN

Temuan penelitian ini memberikan implikasi praktis bagi institusi budaya, perpustakaan, museum, dan lembaga pendidikan dalam digitalisasi serta pelestarian kaligrafi Arab. Model CNN-CBAM menunjukkan performa lebih baik dibandingkan CNN konvensional, sehingga berpotensi meningkatkan efisiensi klasifikasi otomatis, pengindeksan dokumen, dan pencarian arsip digital. Integrasi mekanisme attention juga memperkuat ekstraksi fitur dengan tambahan beban komputasi yang relatif rendah. Secara keseluruhan, model ini berpotensi mendukung pengembangan sistem manajemen dokumen cerdas, aplikasi edukasi, dan pelestarian warisan budaya Islam.

8. DEKLARASI

8.1. Tentang Penulis

Indra Riyanto (IR) https://orcid.org/0009-0002-2738-7068
Muhamad Rido (MR) https://orcid.org/0009-0008-1057-6421
Marta Rodriguez (MZ) https://orcid.org/0009-0000-1367-0511

8.2. Kontribusi Penulis

Konseptualisasi: IR dan MR; Metodologi: IR dan MR; Perangkat Lunak: IR dan MR; Validasi: IR dan MR; Analisis Formal: IR dan MR; Investigasi: IR dan MR; Sumber daya: IR dan MR; Kurasi Data: IR dan MR; Penulisan Draf Awal: IR dan MR; Peninjauan dan Penyuntingan Tulisan: IR dan MR; Visualisasi: IR dan MR; Semua penulis, IR, MR dan MZ, telah membaca dan menyetujui naskah yang telah diterbitkan.

8.3. Pernyataan Ketersediaan Data

Data yang disajikan dalam studi ini tersedia atas permintaan dari penulis terkait.

8.4. Pendanaan

Penulis tidak menerima dukungan finansial untuk penelitian, kepenulisan, dan/atau penerbitan artikel ini.

8.5. Deklarasi Konflik Kepentingan

Penulis menyatakan bahwa mereka tidak memiliki konflik kepentingan, konflik kepentingan finansial yang diketahui, atau hubungan pribadi yang dapat memengaruhi pekerjaan yang dilaporkan dalam makalah ini.

DAFTAR PUSTAKA

[1] S. Alghyaline, "Arabic optical character recognition: A review," *Computer Modeling in Engineering & Sciences*, vol. 135, no. 3, pp. 1825–1861, 2023. [Online]. Available: https://doi.org/10.32604/cmes.2022.024555

[2] M. Cleti, "Attention mechanisms in deep learning: A comprehensive survey from foundations to modern applications," Tech. Rep., 2025. [Online]. Available: https://doi.org/10.13140/RG.2.2.18159.29605

[3] A. A. Zainuddin, R. M. Nor, D. Handayani, M. I. M. Tamrin, K. Subramaniam, and S. F. N. Sadikan, "Smart attendance in classroom (cobot): Iot and facial recognition for educational and entrepreneurial impact," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 6, no. 3, pp. 608–622, 2024.

[4] M.-H. Guo *et al.*, "Attention mechanisms in computer vision: A survey," *Computational Visual Media*, vol. 8, no. 3, pp. 331–368, Sep. 2022. [Online]. Available: https://doi.org/10.1007/s41095-022-0271-y

[5] K. Han, Y. Wang, H. Chen, X. Chen, J. Guo, Z. Liu, Y. Tang, A. Xiao, C. Xu, Y. Xu, Z. Yang, Y. Zhang, and D. Tao, "A survey on vision transformer," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 1, pp. 87–110, Jan. 2023. [Online]. Available: https://doi.org/10.1109/TPAMI.2022.3152247

[6] M. S. Kasem, M. Mahmoud, and H.-S. Kang, "Advancements and challenges in arabic optical character recognition: A comprehensive survey," 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2312.11812

[7] Y. Kristiana, V. A. H. Goeltom, and T. C. Nathalia, "Ai-driven pos, adaptable workload, and service culture influence on frontline employee performance," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 7, no. 1, pp. 217–227, 2025.

[8] H. Lamtougui, H. El Moubtahij, H. Fouadi, and K. Satori, "An efficient hybrid model for arabic text recognition," *Computers, Materials & Continua*, vol. 74, no. 2, pp. 2871–2888, 2023. [Online]. Available: https://doi.org/10.32604/cmc.2023.032550

[9] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: Analysis, applications, and prospects," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 12, pp. 6999–7019, Dec. 2022. [Online]. Available: https://doi.org/10.1109/TNNLS.2021.3084827

[10] Z. Lv, K. Chen, T. Zhang, J. Zhao, and W. Wang, "Multi-feature generation network-based imputation method for industrial data with high missing rate," *Expert Systems with Applications*, vol. 227, p. 120229, Oct. 2023. [Online]. Available: https://doi.org/10.1016/j.eswa.2023.120229

[11] E. Susetyono, D. S. Priyarsono, A. Sukmawati, and P. Nurhayati, "Improving risk management maturity in ultra micro soe holding companies," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 8, no. 1, pp. 310–324, 2026.

[12] S. Münster *et al.*, "Artificial intelligence for digital heritage innovation: Setting up a R&D agenda for europe," *Heritage*, vol. 7, no. 2, pp. 794–816, Feb. 2024. [Online]. Available: https://doi.org/10.3390/heritage7020038

[13] V. Poulopoulos and M. Wallace, "Digital technologies and the role of data in cultural heritage: The past, the present, and the future," *Big Data and Cognitive Computing*, vol. 6, no. 3, p. 73, Jul. 2022. [Online]. Available: https://doi.org/10.3390/bdcc6030073

[14] S. O. Salih and K. Jacksi, "KSTRV1: A scene text recognition dataset for central kurdish in (arabic-based) script," *Data in Brief*, vol. 60, p. 111648, Jun. 2025. [Online]. Available: https://doi.org/10.1016/j.dib.2025.111648

[15] I. R. Maulana, U. Rahardja, N. Azizah, M. Rakhmansyah, and M. A. Komara, "Leveraging ipfs to build secure and decentralized websites in the web 3.0 era," *IAIC Transactions on Sustainable Digital Innovation (ITSDI)*, vol. 7, no. 1, pp. 1–12, 2025.

[16] T. B. A. Gader and A. K. Echi, "Attention-based deep learning model for arabic handwritten text recognition," *Machine Graphics & Vision*, vol. 31, no. 1/4, pp. 49–73, 2022.

[17] P. A. Suraya, T. Ramadhan, N. Lutfiani, A. Khoirunisa, and U. Rahardja, "Blockchain, information and speculation calculations in indonesia: Recent work," in *2022 10th International Conference on Cyber and IT Service Management (CITSM)*. IEEE, 2022, pp. 1–8.

[18] E. T. Attar, "Deep convolutional neural network for isolated arabic handwritten character recognition: design, evaluation, and comparative study," *Scientific Reports*, vol. 15, no. 1, p. 42467, 2025.

[19] S. L. Sitorus, R. T. H. Safariningsih, A. H. A. N. Karsa, M. A. Komara, and R. Evans, "Optimizing smart contracts and blockchain for sustainable digital fashionpreneurship," *Blockchain Frontier Technology*, vol. 5, no. 1, pp. 37–48, 2025.

[20] K. M. Nahar, I. Alsmadi, R. E. Al Mamlook, A. Nasayreh, H. Gharaibeh, A. S. Almuflih, and F. Alasim, "Recognition of arabic air-written letters: machine learning, convolutional neural networks, and optical character recognition (ocr) techniques," *Sensors*, vol. 23, no. 23, p. 9475, 2023.

[21] A. A. H. Al-Radaideh, M. S. bin Mohd Rahim, W. Ghaban, M. Bsoul, S. Kamal, and N. Abbas, "Arabic words extraction and character recognition from picturesque image macros with enhanced vgg-16 based model functionality using neural networks." *KSII Transactions on Internet & Information Systems*, vol. 17, no. 7, p. 1807, 2023.

[22] M. Elkhayati, Y. Elkettani, and M. Mourchid, "Segmentation of handwritten arabic graphemes using a directed convolutional neural network and mathematical morphology operations," *Pattern Recognition*, vol. 122, p. 108288, 2022.

[23] L. Philip, T. Pradiani *et al.*, "Influence brand experience, viral marketing and brand image to brand loyalty to service users streaming spotify in indonesia," *ADI Journal on Recent Innovation (AJRI)*, vol. 5, no. 2, pp. 127–135, 2024.

[24] T. B. Aicha and A. K. Echi, "Attention-based cnn-convlstm for handwritten arabic word extraction," *ELCVIA: electronic letters on computer vision and image analysis*, vol. 21, no. 1, pp. 121–129, 2022.

[25] K. Miloud, M. L. Abdelmounaim, B. Mohammed, and B. R. Ilyas, "Advancing ancient arabic manuscript restoration with optimized deep learning and image enhancement techniques," *Traitement du Signal*, vol. 41, no. 4, p. 2203, 2024.

[26] S. Rahayu, N. Septiani, R. Z. Ikhsan, Y. M. Kareem, U. Rahardja *et al.*, "Revolutionizing renewable energy systems throughadvanced machine learning integration approaches," *Journal of Computer Science and Technology Application*, vol. 2, no. 2, pp. 23–34, 2025.

[27] S. S. Qasim and S. H. Oleiwi, "Advancing arabic handwritten digit recognition with ai-enhanced neural network architectures," *Babylonian Journal of Artificial Intelligence*, vol. 2024, pp. 146–157, 2024.

[28] S. A. Alsuhibany and T. Alwehaid, "Security and usability evaluation of handwritten arabic captcha systems based on saudi calligraphic fonts," *Scientific Reports*, 2026.

[29] C. Salhi, A. El Allati, and M. Ben Ahmed, "Exploring classical and quantum-inspired approaches for arabic handwritten text recognition: A comprehensive and detailed review," in *International Conference on Digital Age & Technological Advances for Sustainable Development*. Springer, 2025, pp. 64–76.

[30] F. Haque, A. S. M. Miah, M. Al-Hasan, S. T. Mou, M. A. Rahim, and J. Shin, "Multichannel attention networks with ensembled transfer learning to recognize bangla handwritten character," in *Data Driven Applications for Industry 4.0 and Beyond*. Chapman and Hall/CRC, 2025, pp. 172–186.

[31] K. Mirdad, O. P. M. Daeli, N. Septiani, A. Ekawati, U. Rusilowati *et al.*, "Optimizing student engagement and performance usingai-enabled educational tools," *Journal of Computer Science and Technology Application*, vol. 1, no. 1, pp. 53–60, 2024.

[32] S. Saber and M. G. Mahdi, "Urdu handwriting recognition with deep learning: Current methods and

future prospects," *International Journal of Computers and Informatics (Zagazig University)*, vol. 3, pp. 32–43, 2024.

[33] M. Salaheldin Kasem, M. Mahmoud, and H.-S. Kang, "Advancements and challenges in arabic optical character recognition: A comprehensive survey," *ACM Computing Surveys*, vol. 58, no. 4, pp. 1–37, 2025.

[34] M. Ayadi, N. Masmoudi, L. Almuqren, R. O. Aljohani, and H. S. Alshahrani, "Empowering accessibility in handwritten arabic text recognition for visually impaired individuals through optimized generative adversarial network (gan) model," *Journal of disability research*, vol. 4, no. 1, p. 20240110, 2025.

[35] B. A. Tuama and F. Mohamed, "A systematic literature review of deep learning methods for handwritten text recognition in historical arabic manuscripts," *Engineering, Technology & Applied Science Research*, vol. 15, no. 4, pp. 25 772–25 782, 2025.

[36] A. A. AGUNG, "Implementasi kegiatan ekstrakurikuler kaligrafi dalam mengembangkan kecerdasan visual spasial peserta didik di ma diniyyah putri lampung," Ph.D. dissertation, UIN RADEN INTAN LAMPUNG, 2026.

[37] A. A. Chandio, M. Asikuzzaman, M. R. Pickering, and M. Leghari, "Cursive text recognition in natural scene images using deep convolutional recurrent neural network," *IEEE Access*, vol. 10, pp. 10 062–10 078, 2022.

[38] Z. A. Umar and M. Yunus, "Optimalisasi teknologi ocr dalam digitalisasi manuskrip hadis: Studi akurasi dan tantangan linguistik arab klasik," *Hamidah: Jurnal Ilmu Hadis*, vol. 1, no. 1, pp. 43–54, 2025.

[39] A. Waly, B. Tarek, A. Feteha, R. Yehia, G. Amr, W. Gomaa, and A. Fares, "Invizo: Arabic handwritten document optical character recognition solution," *arXiv preprint arXiv:2502.05277*, 2025.

[40] J. Wang, X. Liao, Y. Che, P. Wang, Y. Zhang, F. Wang, F. Marias, and H. Rong, "Lstm-cnn-cbam-based prediction of nox emissions with zonal control strategies for flexible peak-load nox reduction in cfb coal-fired units," *Journal of Cleaner Production*, vol. 545, p. 147705, 2026.

[41] Z. Qin, L. Xu, C. Chen, Y. Qi, Y. Li, and Y. Li, "Cnn-cbam prediction model for moisture content determination of raw materials for wood-based panels based on near-infrared spectroscopy," *Analytical Methods*, vol. 18, no. 3, pp. 688–699, 2026.

[42] S. N. A. Norhairi, A. S. A. Nasir, H. B. Affandy, and F. Abdullah, "Hybrid cnn-cbam architecture for facial expression recognition across multiple benchmark datasets," in *2026 2nd International Conference on Electrical, Robotics and Intelligent Systems (ELECRIS)*. IEEE, 2026, pp. 450–455.

[43] A. S. Bist, B. Rawat, U. Rahardja, M. Hardini, and M. Kamal, "Unveiling ethical challenges and malpractice in academic research and integrity," in *2024 3rd International Conference on Creative Communication and Innovative Technology (ICCIT)*. IEEE, 2024, pp. 1–6.

[44] X. Chen, "Design of a fruit sorting system based on deep learning models optimized by cbam attention mechanism and its simulation verification in coppeliasim," in *2025 5th International Conference on Electronic Information Engineering and Computer Science (EIECS)*. IEEE, 2025, pp. 967–972.

[45] X. Xu, J. Sun, and Z. Li, "Research on writer identification based on attention mechanism," in *Sixteenth International Conference on Graphics and Image Processing (ICGIP 2024)*, vol. 13539. SPIE, 2025, pp. 322–329.

[46] M. Nurkanti, U. Rahardja, M. Lubis, and A. Shukri, "Deep learning-based biology learning with ethnopedagogy and local wisdom to support sustainable development goals," *Jurnal Pendidikan IPA Indonesia*, vol. 15, no. 1, pp. 103–111, 2026.

[47] R. Alabdan, M. A. AlAqil, E. M. Alanazi, and A. S. Zamani, "Enhance real-time activity recognition of disabled individuals using hybridisation of convolutional neural network with attention mechanism," *Signal, Image and Video Processing*, vol. 19, no. 14, p. 1268, 2025.

[48] OECD, *Enhanced Monitoring Report on the Implementation of the Standard on Transparency and Exchange of Information on Request 2025*. Paris: OECD Publishing, 2025. [Online]. Available: https://doi.org/10.1787/4db1a02b-en

[49] M. K. Sastry and S. Aparna, "Developing adaptive yolov5-based telugu handwritten character segmentation and classification framework using enhanced chef-based optimization algorithm and deep learning networks," *Computers and Electrical Engineering*, vol. 126, p. 110533, 2025.

[50] K. Khoirudin, S. Y. J. Prasetyo, S. Wijono, E. Maria, and U. Rahardja, "Transfer learning in data-scarce agricultural yield forecasting: A bibliometric and systematic literature review," in *2nd International Conference on Social Environment Diversity (ICOSEND 2025)*. Atlantis Press, 2026, pp. 467–477.