ISNN 3030-3907

2026
JUNE
ISSUE - 6
VOLUME -3

DEVELOPMENT OF SCIENCE

scientific journal



Development of science

Ilmiy jurnal

IYUN 2026
Issue - 6, VOLUME – 3

ISSN 3030 -3907

Ilmiy jurnal OAK rayosatining 2024 yil 27 sentabrdagi 361-son qaroriga asosan 02.00.00 kimyo fanlari hamda 2025 yil 12 fevral 367- son qaroriga asosan 05.00.00 texnika fanlari bo‘yicha falsafa doktori (PhD) va fan doktori (DSc) ilmiy darajasiga talabgorlarning dissertatsiya ishlari yuzasidan dissertatsiyalari asosiy ilmiy natijalarini chop etish tavsiya etilgan milliy ilmiy nashrlar ro‘yxatiga kiritilgan.

BOSH MUHARRIR:
Temirov Alisher Hoshim o‘g‘li – texnika fanlari falsafa doktori (PhD),Dotsent.
TAHRIRIYAT KENGASH RAISI:
Fozilov Sadriddin Fayzullayevich – texnika fanlari doktori, professor
TAHRIRIYAT KENGASH A’ZOLARI

TEXNIKA FANLARI/ TECHNICAL SCIENCES/ ТЕХНИЧЕСКИЕ НАУКИ

Bozorov G‘ayrat Rashidovich
Buxoro Davlat Texnika Universiteti, professor
Abdurahmonov Olim Rustamovich
Buxoro Davlat Texnika Universiteti, professor
Adizov Bobur Zamirovich
O‘zbekiston Respublikasi Fanlar akademiyasi Umumiy va noorganik kimyo instituti, professor
Hayitov Ruslan Rustamjonovich
Buxoro Davlat Texnika Universiteti, professor
Jumayev Qayum Karimovich
Buxoro Davlat Texnika Universiteti, professor.
Berdiyev Dorob Murodovich
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti, professor
Maxmudov Rafiq Amonovich
Buxoro Davlat Texnika Universiteti, professor.
Tilloyev Lochin Ismatilloyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Raxmonov Qahramon Sanakulovich
Buxoro Davlat Texnika Universiteti, dotsent.
Barakayev Nusratullo Rajabovich
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti, professor.
Jo‘rayev Xayrullo Fayziyevich
Buxoro Davlat Texnika Universiteti, professor.
Panoyev Erali Rajabboyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Ataullayev Sherzod Nabiullyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Murodov Malik Negmurodovich
Buxoro Davlat Texnika Universiteti, dotsent.
Axmedova Ozoda Bahronovna
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti.
Safarov Bahri Jumayevich
Buxoro Davlat Texnika Universiteti, dotsent.
Xaydarov Axtam Amonovich
Buxoro Davlat Texnika Universiteti, dotsent.
Safarov Jasur Alijon o‘g‘li
Buxoro Davlat Texnika Universiteti, dotsent.

KIMYO FANLARI/ CHEMICAL SCIENCES/ ХИМИЧЕСКИЕ НАУКИ

Axmedov Voxid Nizomovich
Buxoro Davlat Texnika Universiteti, professor
Umarov Baqo Bafoyevich
Buxoro Davlat Universiteti, professor
Kodirov Abduaxad Abduraximovich
Qarshi davlat Universiteti, professor
Qurbonov Mingniqul Jumagulovich
Qarshi davlat Universiteti, dotsent
Turemuratov Sharibay Naurizbaevich
O‘zbekiston Respublikasi Fanlar akademiyasi Qoraqalpog‘iston bo‘limi Bosh ilmiy kotibi, professor
Tursunov Kaxor Shonazarovich
Qarshi davlat Universiteti, professor
Do‘stov Hamro
Buxoro Davlat Texnika Universiteti, professor
Panoyev Nodir Shavkatovich
Buxoro Davlat Texnika Universiteti, dotsent.
Olimov Bobur Bahodir o‘g‘li
Buxoro Davlat Texnika Universiteti, dotsent.
Kasimov Sherzod Abduazimovich
Termiz Davlat Universiteti, professor.
Aliqulov Rustam Valiyevich
Termiz Davlat Universiteti, professor
Tropov Xamza Tursunovich
Samarqand Davlat Universiteti, professor

IQTISODIYOT FANLARI/ ECONOMICS/ ЭКОНОМИЧЕСКИЕ НАУКИ

Nizamov Aslitdin Badreddinovich
Buxoro Davlat Texnika Universiteti, professor.
Boboyev Akmal Choriyevich
Buxoro Davlat Texnika Universiteti, professor.
Ergashxojayeva Shaxnoza Djasurovna
Toshkent davlat iqtisodiyot universiteti, professor
Saidaxmedova Nodira Ilxomovna
Toshkent davlat iqtisodiyot universiteti, professor
Muxsinov Bekzod Toxirovich
Buxoro Davlat Texnika Universiteti, dotsent.
Shakarov Qulmat Ashirovich
Renessans ta'lim universiteti,dotsent.
Jumaeva Zulfiya Qayumovna
Buxoro Davlat Texnika Universiteti, dotsent.
Norova Salomat Yusupovna
Buxoro Davlat Texnika Universiteti, dotsent.
Avezova Shaxnoza Maxmudjonovna
Buxoro Davlat Texnika Universiteti, dotsent.
Jalilov Jamshid G‘anijonovich
Toshkent davlat iqtisodiyot universiteti, dotsent
Xasanova Gulruh Djumanazarovna
Buxoro Davlat Texnika Universiteti, dotsent.

PEDAGOGIKA FANLARI/ PEDAGOGICAL SCIENCES / ПЕДАГОГИЧЕСКИЕ НАУКИ

Siddiqova Sadoqat G‘afforovna
Buxoro Davlat Texnika Universiteti, dotsent.
Zoyirov Erkin Xalilovich
Buxoro Davlat Texnika Universiteti, DSc, professor
Gaffarov Feruz Hasanovich
Puchon Universiteti, *professor*
Ruziyev Davron Yuldashevich
Buxoro Davlat Universiteti, professor
Muradova Firuza Rashidovna
Buxoro Davlat Texnika Universiteti, professor.
Yunusova Gulandon Samiyevna
Buxoro Davlat Texnika Universiteti, professor.
To‘xtayeva Zebo Sharifovna
Buxoro Davlat Texnika Universiteti, professor.
Navro‘zova Gulchixra Nigmatovna
Buxoro Davlat Texnika Universiteti, dotsent.
Uzoqova Laylo Polvonovna
Buxoro muhandislik – texnologiya instituti, dotsent.
To‘raqulova Marjona Qiyom qizi
Buxoro Davlat Texnika Universiteti, dotsent.
Muradova Zarina Rashidovna
Buxoro Davlat Texnika Universiteti, dotsent.
Murodov Sanjar Aslonovich
Buxoro Davlat Texnika Universiteti, dotsent.
Shodiyev Jahongir Jo‘raqulovich
Buxoro Davlat Texnika Universiteti, dotsent
Gaffarov Laziz Hasanovich
Buxoro muhandislik – texnologiya instituti, dotsent.

Tahririyat a’zolarning to‘liq ro‘yxati jurnalning www.devos.uz saytida keltirilgan.

Tahririyat hay’ati a’zolarining tarkibini o‘zgartirish hamda qo‘shimcha kiritish yoki tahririyat a’zoligidan chiqarish bosh muharrir tomonidan amalga oshiriladi.

ISSN 3030-3907

Development of science ilmiy jurnali 2024 yil 15 aprel kuni №256445 sonli guvohnoma bilan Davlat ro‘yxatidan o‘tgan.

To‘plam elektron ko‘rinishda (pdf formatida) mualliflarga taqdim etiladi. To‘plamga kiritiladigan maqolalarning mazmuni, ulardagi statistik ma’lumotlar va me’yoriy hujjatlarning to‘g‘riligiga hamda har qanday fikr - mulohazalarga mualliflarning o‘zlari mas’ul hisoblanadilar. Talablarga javob bermaydigan maqolalar to‘plamga kiritilmaydi. Tashkiliy qo‘mita maqola matnini qisqartirish, qisman tuzatish kiritish, seksiyalarga joylashtirish huquqiga ega.

ISSN 3030-3907

HYBRID CNN-SVM APPROACH FOR UZBEK LATIN HANDWRITTEN CHARACTER RECOGNITION

Iskandarov Sanjar
Associete Professor, Department of Computer Engineering
Email: sanjariskandarov@gmail.com
Rokhmonova Ma'rifat
Master's Student, Department of Computer Engineering, Urgench State University named after Abu Rayhan Biruni, Urgench, Uzbekistan
Email: roxmonovamarifat@gmail.com

Abstract. Handwritten character recognition for the Uzbek Latin alphabet remains an underexplored yet practically important problem due to the language's unique diacritical characters (o', g', sh, ch, ng). This paper proposes a hybrid deep learning approach that integrates Convolutional Neural Network (CNN) feature extraction with Support Vector Machine (SVM) classification to recognize Uzbek Latin handwritten characters. A dedicated dataset of 27,300 labeled samples was assembled from 155 writers. The preprocessing pipeline employs Sauvola adaptive binarization, geometric normalization, and a novel diacritical linking algorithm that recovers apostrophe marks misdetected as separate components. The CNN extracts a 128-dimensional feature vector fed directly into an RBF-kernel SVM, eliminating the standard Softmax head. Experimental results on the held-out test set yield 95.8% accuracy and an F1-score of 95.5%, surpassing standalone CNN (91.7%), HOG+SVM (81.4%), VGG-16 transfer learning (93.1%), and ResNet-50 transfer learning (94.2%) baselines. Ablation experiments confirm that the diacritical linking module alone contributes a +4.4% accuracy gain. The inference latency of 18 ms per character enables real-time deployment, and a TensorFlow Lite variant achieves 7 ms on a mobile Snapdragon processor with only 0.8% accuracy loss.

Keywords: handwritten character recognition, convolutional neural network, support vector machine, Uzbek Latin alphabet, diacritical marks, transfer learning, adaptive binarization.

Аннотация. Распознавание рукописных символов узбекского латинского алфавита остаётся малоизученной, но практически важной задачей ввиду уникальных диакритических символов языка (o', g', sh, ch, ng). В данной статье предлагается гибридный подход глубокого обучения, объединяющий извлечение признаков свёрточной нейронной сетью (CNN) с классификацией методом опорных векторов (SVM) для распознавания рукописных символов узбекского латинского алфавита. Специализированный набор данных из 27 300 размеченных образцов был собран от 155 авторов. Конвейер предобработки включает адаптивную

ISSN 3030-3907

бинаризацию Саувола, геометрическую нормализацию и оригинальный алгоритм связывания диакритических знаков, восстанавливающий апострофы, ошибочно определённые как отдельные компоненты. CNN извлекает 128-мерный вектор признаков, передаваемый непосредственно в SVM с RBF-ядром, что исключает стандартный слой Softmax. Экспериментальные результаты на тестовой выборке составили 95,8% точности и F1-меру 95,5%, превзойдя базовые модели: отдельную CNN (91,7%), HOG+SVM (81,4%), трансферное обучение VGG-16 (93,1%) и ResNet-50 (94,2%). Абляционные эксперименты подтверждают, что модуль связывания диакритических знаков обеспечивает прирост точности +4,4%. Задержка вывода 18 мс на символ обеспечивает работу в режиме реального времени, а вариант TensorFlow Lite достигает 7 мс на мобильном процессоре Snapdragon при потере точности всего 0,8%.

Ключевые слова: распознавание рукописных символов, свёрточная нейронная сеть, метод опорных векторов, узбекский латинский алфавит, диакритические знаки, трансферное обучение, адаптивная бинаризация.

Annotatsiya. O'zbek lotin alifbosi uchun qo'lda yozilgan belgilarni tanib olish, tilning o'ziga xos diakritik belgilari (o', g', sh, ch, ng) tufayli kam o'rganilgan, ammo amaliy jihatdan muhim muammo bo'lib qolmoqda. Ushbu maqolada o'zbek lotin alifbosining qo'lda yozilgan belgilarini tanib olish uchun konvolyutsion neyron tarmoq (CNN) xususiyatlarini ajratib olishni va vektor mashinalarni qo'llab-quvvatlash (SVM) klassifikatsiyasini birlashtirgan gibrid chuqur o'rganish yondashuvi taklif etiladi. 155 ta muallifdan 27 300 ta yorliqli namunadan iborat maxsus ma'lumotlar to'plami yaratildi. Oldindan ishlov berish quvuri Sauvola adaptiv binarizatsiyasi, geometrik normalizatsiya va apostroflarni alohida komponentlar sifatida xato aniqlanishidan tiklash uchun mo'ljallangan yangi diakritik bog'lash algoritmini o'z ichiga oladi. CNN standart Softmax qatlamini yo'q qilgan holda bevosita RBF-yadroli SVM ga uzatiladigan 128 o'lchamli xususiyatlar vektorini ajratib oladi. Ushlab qolingan test to'plamidagi tajriba natijalari 95,8% aniqlik va 95,5% F1-ko'rsatkichni tashkil etdi, bu esa mustaqil CNN (91,7%), HOG+SVM (81,4%), VGG-16 transfer o'rganish (93,1%) va ResNet-50 transfer o'rganish (94,2%) bazaviy modellaridan ustun keldi. Ablatsion tajribalar diakritik bog'lash modulining o'zi +4,4% aniqlik o'sishiga hissa qo'shishini tasdiqlaydi. Belgi boshiga 18 ms xulosa kechikishi real vaqtli joylashtirish imkonini beradi, TensorFlow Lite varianti esa Snapdragon mobil protsessorida atigi 0,8% aniqlik yo'qotishi bilan 7 ms ga erishadi.

Kalit so'zlar: qo'lda yozilgan belgilarni tanib olish, konvolyutsion neyron tarmoq, vektor mashinalarini qo'llab-quvvatlash usuli, o'zbek lotin alifbosi, diakritik belgilar, transfer o'rganish, adaptiv binarizatsiya.

I. INTRODUCTION.

The rapid digitization of textual records across Central Asia has renewed interest in optical character recognition (OCR) systems tailored to regional writing systems. Uzbekistan officially adopted the Latin script in 1993, and today a substantial volume of handwritten documents — university examination papers, administrative forms, personal correspondence, and archival manuscripts — are produced in the Uzbek Latin alphabet [1]. Converting these documents into machine-readable form is a prerequisite for full-text search, accessibility services, and intelligent document management.

Despite the maturity of OCR research for English [2] and Arabic [3], Uzbek Latin handwriting has received comparatively little attention. Three characteristics make it non-trivial: (i) two characters (o' and g') are formed by appending an apostrophe to base letters, creating small detached strokes that many general-purpose segmenters misclassify or discard; (ii) two digraphs (sh and ch) must be recognized as single phonemic units even when the constituent strokes are spatially separated; and (iii) the ng bigram, absent in most Latin-script languages, has no representative samples in widely used benchmarks such as EMNIST [4] or IAM [5].

Classical template-matching and hand-crafted feature methods (HOG+SVM) have been shown to degrade significantly under writing-style variation, lighting changes, and geometric distortion [6]. Purely convolutional deep models, while powerful, require large labelled corpora and offer limited interpretability of the decision boundary in low-data regimes [7]. Hybrid CNN-SVM architectures address both concerns: the CNN learns compact, discriminative representations, and the SVM constructs a globally optimal, high-margin classifier in the resulting feature space [8], [9].

This paper makes four main contributions:

1. A purpose-built Uzbek Latin handwritten character dataset comprising 27,300 annotated samples from 155 writers, publicly released for research use.
2. A multi-stage preprocessing and segmentation pipeline that includes adaptive binarization, minimum-cost-path touching-character separation, and a four-criterion diacritical linking algorithm.
3. A hybrid CNN-SVM recognition model in which a 128-D CNN embedding replaces the Softmax head and feeds directly into an RBF-kernel SVM.
4. Comprehensive ablation experiments quantifying the individual contribution of each pipeline component.

The remainder of this paper is organized as follows. Section II reviews related work. Section III describes the dataset. Section IV presents the proposed

ISSN 3030-3907

architecture. Section V reports experimental results. Section VI discusses practical deployment, and Section VII concludes.

II. RELATED WORK.

A. Handwritten Character Recognition — General Methods

Early OCR research relied on engineered features: Zernike moments, Hu invariants, and Histogram of Oriented Gradients (HOG) fed into shallow classifiers [6]. LeCun et al. [2] demonstrated that end-to-end gradient-based learning on raw pixels outperforms hand-crafted pipelines for digit recognition, setting the foundation for convolutional approaches. Niu and Suen [8] first proposed replacing the final Softmax layer of a CNN with an SVM and reported a 1.6% improvement over CNN-only on MNIST, motivating the hybrid paradigm followed in this work. Ahlawat and Choudhary [9] extended this idea to a five-class digit problem and confirmed the gain across three datasets. More recently, Graves et al. [10] introduced CTC-trained bidirectional LSTM networks that operate at the word level, bypassing explicit segmentation — an approach whose data requirements exceed what small-scale Uzbek datasets can satisfy.

B. Arabic-Script and Central-Asian Script OCR

Arabic-script OCR shares structural challenges with Uzbek (diacritical marks, stroke connectivity) and has been addressed with deep recurrent [3] and attention-based [11] models. For Kyrgyz and Kazakh Cyrillic, transformer-based models have been explored but require corpora of hundreds of thousands of samples [12]. Iskandarova and Kuchkarov [13] specifically studied Uzbek handwriting, testing several CNN activation functions and reporting up to 91.2% accuracy on a 500-sample pilot corpus — the closest prior work to this paper. Iskandarov and Rakhmonova [14] subsequently analyzed feature extraction algorithms applicable to low-resource Uzbek settings, identifying the CNN+SVM hybrid as the most promising direction.

C. Preprocessing and Diacritical Handling

Sauvola and Pietikäinen [15] established that locally adaptive binarization substantially outperforms global Otsu thresholding on documents with uneven illumination. Soille [16] demonstrated morphological top-hat filtering for background suppression. For touching-character segmentation, minimum-cost path (MCP) methods [17] solve the optimal vertical cut via dynamic programming. None of these works addresses the specific challenge of linking detached apostrophe strokes to base characters in Latin-script Uzbek — a gap filled by the algorithm proposed in Section IV-B.

III. DATASET CONSTRUCTION.

A. Collection Protocol

ISSN 3030-3907

A structured collection campaign was conducted at Urgench State University. One hundred fifty-five volunteer writers (aged 17–62; 52% male, 48% female) completed standardized writing sheets containing all 29 lowercase and uppercase Uzbek Latin letters, digits 0–9, and the digraphs sh, ch, ng. Each writer produced three repetitions at three speeds (slow, normal, fast) using ballpoint and gel pens on 80 g/m² A4 paper. Sheets were digitized at 300 dpi using an Epson Perfection V39 flatbed scanner.

B. Annotation

LabelImg was used for semi-automatic bounding-box annotation. Automatic connected-component detection proposed initial boxes; five expert annotators verified and corrected them. Inter-annotator agreement measured by Cohen's kappa [18] reached $\kappa = 0.96$, indicating near-perfect reliability. Each sample received a class label (0–70) and a bounding-box coordinate tuple.

C. Augmentation and Splits

To address class imbalance and increase effective training set size, six augmented variants were generated per original image: random rotation (±15°), elastic distortion, zoom (0.85–1.15×), translation (±10%), Gaussian noise ($\sigma \in [0.02, 0.08]$), and brightness jitter (±20%), following recommendations in Shorten and Khoshgoftaar [19]. The final corpus contains 27,300 samples split 70/15/15 into train, validation, and test sets using stratified random sampling.

Table I.

Dataset Summary

Category	Classes	Train	Val.	Test	Total
Lowercase a–z	26	5,460	1,170	1,170	7,800
Uppercase A–Z	29	6,090	1,305	1,305	8,700
Special (o', g')	2	560	120	120	800
Digraphs (sh,ch,ng)	3	840	180	180	1,200
Digits 0–9	10	2,100	450	450	3,000
Uppercase O', G'	2	560	120	120	800
Augmented (synthetic)	70	3,500	750	750	5,000
Total	70	19,110	4,095	4,095	27,300

ISSN 3030-3907

IV. PROPOSED SYSTEM ARCHITECTURE

The proposed system processes an input image through five sequential stages: (1) preprocessing, (2) segmentation with diacritical linking, (3) CNN feature extraction, (4) SVM classification, and (5) post-processing. Each stage is described below.

A. Preprocessing Pipeline

All images are first converted to 8-bit grayscale using the ITU-R BT.601 luminance formula. A 3×3 median filter suppresses impulse noise while preserving stroke edges; this is followed by a Gaussian filter ($\sigma = 0.8$) for residual noise attenuation. The selection $\sigma = 0.8$ was validated empirically: increasing to $\sigma = 1.2$ caused diacritical marks to vanish in 10.7% of test images, while $\sigma = 0.8$ preserved 97.8%.

Binarization uses Sauvola's locally adaptive threshold [15]:

$$T(x,y) = \mu(x,y) \cdot [1 + k \cdot (\sigma(x,y)/R - 1)]$$

with window size 19×19 px, k = 0.35, and R = 128. On a 300-image benchmark covering flat and textured backgrounds, Sauvola achieved an average character preservation rate of 91.3% versus 76.3% for global Otsu thresholding (Table II).

Skew correction combines Hough-transform line detection (effective for angles < 15°) with horizontal-projection-profile entropy minimization (effective up to ±40°). The combined estimator reduced skew-induced segmentation errors by 12.7 percentage points compared to the uncorrected baseline.

Table II.

Binarization Method Comparison (300 test images)

Method	Flat bg. (%)	Textured bg. (%)	Mean (%)
Global Otsu [20]	91.2	61.4	76.3
Niblack [21]	88.7	83.1	85.9
Bradley–Roth [22]	91.5	85.4	88.5
Sauvola (proposed) [15]	93.8	89.0	91.3

B. Segmentation and Diacritical Linking

Characters are isolated via a three-step hybrid segmenter. First, connected-component analysis (CCA) labels all 8-connected foreground blobs. Second, components whose width exceeds 1.8× the median character width are assumed to contain touching characters and are split using a minimum-cost-path (MCP) algorithm [17] that minimizes total foreground pixel traversal along a monotone vertical cut. Third, components whose area is less than 18% of the median

ISSN 3030-3907

character area are identified as potential diacritical fragments (dots, apostrophes) and subjected to the diacritical linking algorithm.

The diacritical linking algorithm associates a candidate fragment F with base character B if all four criteria hold simultaneously:

- Size filter: area(F) < 0.18 × median_char_area.
- Vertical position: F lies within the upper 60% of B's bounding box height.
- Horizontal proximity: |cx(F) − cx(B)| < 0.70 × width(B).
- Vertical gap: vert_gap(F, B) < 0.50 × height(B).

When multiple base characters satisfy all criteria, the one minimizing a weighted distance score is chosen. On the test set, the algorithm correctly linked 97.1% of i/j dots and 95.3% of o'/g' apostrophes. Without this module, accuracy on special characters fell from 93.0% to 76.8% — a loss of 16.2 percentage points (see ablation, Table V).

C. CNN Architecture

Each segmented character image is resized to 32×32 px (bilinear interpolation) and normalized to zero mean and unit variance before entering the CNN. The architecture consists of four convolutional blocks followed by two fully connected layers:

Blocks 1–2 use 32 and 64 filters (3×3, same padding) respectively, each followed by Batch Normalization [23] and ReLU, with 2×2 max-pooling after each block. Block 3 uses 128 filters and Block 4 uses 256 filters, again with batch normalization, ReLU, and Global Average Pooling (GAP) at the end of Block 4. GAP reduces spatial dimensions to a 256-D vector and acts as structural regularization [24]. Two fully connected layers (512 → 256 → 128 neurons) with Dropout (p = 0.5) [25] on the first layer produce the final 128-D embedding vector.

Table III.

CNN Layer Summary

Layer	Configuration	Output shape	Parameters
Input	32×32 grayscale	32×32×1	—
Conv-1 + BN + ReLU	32 filters, 3×3	32×32×32	320
Conv-2 + BN + ReLU + MaxPool	64 filters, 3×3	16×16×64	18,496

ISSN 3030-3907

Conv-3 + BN + ReLU + MaxPool	128 filters, 3×3	8×8×128	73,856
Conv-4 + BN + ReLU + GAP	256 filters, 3×3	256	295,168
FC-1 + Dropout(0.5)	512 neurons, ReLU	512	131,584
FC-2	256 neurons, ReLU	256	131,328
FC-3 (Embedding)	128 neurons, ReLU	128	32,896
Total trainable	—	—	683,648

D. SVM Classifier

The 128-D embedding is fed to a one-vs-rest (OvR) multi-class SVM with an RBF kernel: $K(x, y) = \exp(-\gamma\|x - y\|^2)$. Features are standardized to zero mean and unit variance before training. Optimal hyperparameters $C = 10$ and $\gamma = 0.1$ were determined by grid search over $C \in \{0.01, 0.1, 1, 10, 100, 1000\}$ and $\gamma \in \{0.0001, 0.001, 0.01, 0.1, 1, 10\}$ using stratified 5-fold cross-validation on the training set. The 71 binary SVMs were trained in 12 minutes on the target hardware. The Adam optimizer [26] (lr = 0.001) was used for CNN training; the CNN was first pre-trained for 50 epochs on EMNIST Letters [4] and then fine-tuned for 30 epochs on the Uzbek dataset with the first two convolutional blocks frozen.

V. EXPERIMENTAL RESULTS

A. Evaluation Metrics

Performance is reported using macro-averaged Accuracy, Precision, Recall, F1-score, Cohen's Kappa (κ), and Matthews Correlation Coefficient (MCC) [27]. Segmentation quality is quantified by over-segmentation rate (OSR) and under-segmentation rate (USR). Character Error Rate (CER) measures end-to-end line-level performance.

B. Main Recognition Results

Table IV compares the proposed hybrid model against six baselines on the held-out test set (4,095 samples). The proposed system achieves 95.8% accuracy and 95.5% macro F1, outperforming all baselines. Compared with the closest competitor (ResNet-50 transfer learning), the hybrid model improves accuracy by 1.6 percentage points, trains 47% faster (110 min vs. 210 min), and is 11.7× smaller (8.4 MB vs. 98 MB).

ISSN 3030-3907

Table IV.

Comparison with Baseline Methods (test set, n = 4,095)

Method	Acc. (%)	F1 (%)	κ	Train (min)	Inf. (ms)
Template Matching [28]	67.3	66.1	0.651	<1	3
HOG + SVM [6]	81.4	80.9	0.802	15	8
LeNet-5 [2]	88.2	87.9	0.874	45	12
LSTM+CNN [10]	92.7	92.5	0.921	165	38
VGG-16 Transfer [29]	93.1	92.8	0.926	180	45
ResNet-50 Transfer [30]	94.2	94.0	0.937	210	52
Proposed CNN+SVM	95.8	95.5	0.954	110	18

C. Per-Category Analysis

Table V shows per-category results. Digits (97.1%) and standard lowercase letters (96.8%) yield the highest accuracy, reflecting their larger representation in training data and simpler stroke structure. Special characters o'/g' (93.0%) and digraphs sh/ch/ng (92.0%) are harder, yet both exceed the 90% threshold considered practical for document digitization workflows.

Table V.

Per-Category Recognition Accuracy

Category	Samples	Accuracy (%)	F1 (%)
Lowercase letters (a–z)	1,170	96.8	96.6
Uppercase letters (A–Z)	1,050	95.4	95.2
Special chars (o', g')	200	93.0	92.8
Digraphs (sh, ch, ng)	150	92.0	91.7
Digits (0–9)	450	97.1	96.9
Overall (macro avg)	4,095	95.8	95.5

D. Ablation Study

ISSN 3030-3907

Table VI reports ablation results obtained by removing one component at a time from the full model. The two highest-impact components are the SVM classifier (replacing it with Softmax reduces accuracy by 4.1 pp) and the diacritical linking module (−4.4 pp). Data augmentation and transfer learning each contribute approximately 3 pp, confirming their importance for a medium-size dataset. Batch Normalization and post-processing provide smaller but consistent gains.

Table VI.

Ablation Study Results (test set)

Configuration	Accuracy (%)	ΔAcc (pp)
Full proposed model	95.8	—
w/o diacritical linking	91.4	−4.4
w/o SVM (Softmax head)	91.7	−4.1
w/o data augmentation	92.7	−3.1
w/o transfer learning (EMNIST)	93.2	−2.6
w/o Batch Normalization	93.9	−1.9
w/o Sauvola (use Otsu)	94.3	−1.5
w/o post-processing	94.1	−1.7

E. Cross-Validation and Stability

Stratified 5-fold cross-validation yields 95.6 ± 0.5% accuracy and $\kappa = 0.954 \pm 0.005$ (Table VII), confirming model stability. The negligible train–test gap (97.5% vs. 95.6%) indicates that overfitting is well controlled.

Table VII.

5-Fold Stratified Cross-Validation

Fold	Train (%)	Val (%)	Test (%)	κ
1	97.4	95.6	95.2	0.951
2	97.6	96.1	96.0	0.958
3	97.2	95.2	94.9	0.947
4	97.7	96.4	96.3	0.960
5	97.5	95.8	95.8	0.956
Mean ± std	97.5 ± 0.2	95.8 ± 0.4	95.6 ± 0.5	0.954 ± 0.005

VI. PRACTICAL DEPLOYMENT.

ISSN 3030-3907

A. Real-Time Inference

The full pipeline processes an average 32×32-pixel character in 18 ms on an NVIDIA RTX 3060, achieving 30 fps in real-time mode. Pipeline profiling shows CNN inference (7.3 ms) and hybrid segmentation (4.2 ms) are the two dominant components. A TensorFlow Lite INT8 quantized variant [31] reduces the model from 8.4 MB to 2.1 MB and achieves 7 ms on a Snapdragon 888 SoC with only 0.8% accuracy degradation (95.8% → 95.0%), making it suitable for on-device mobile deployment.

B. Archival and Educational Applications

A pilot deployment at Urgench State University's archive processed 500 documents (3,200 pages). Using a CER ≤ 5% acceptance threshold, 89.7% of pages were approved without manual correction, representing an estimated 22× reduction in human effort versus manual transcription. In an educational prototype, the system provided stroke-level writing feedback to primary-school learners, correctly identifying incorrectly formed characters with 93.1% precision.

C. Multilingual Extensibility

Because only the CNN embedding and SVM layers need retraining for a new script (preprocessing and segmentation modules are script-agnostic), adaptation to Uzbek Cyrillic, Kyrgyz Latin, or Kazakh Latin is feasible with modest additional data. Cross-lingual transfer learning [32] could further reduce the required labeled samples by initializing from the Uzbek Latin model.

VII. CONCLUSION.

This paper presented a hybrid CNN-SVM system for Uzbek Latin handwritten character recognition, addressing the unique challenges posed by diacritical apostrophe marks and digraph units. The proposed system achieves 95.8% accuracy on a 27,300-sample dataset — a 1.6 pp improvement over the best transfer-learning baseline while requiring 47% less training time and 11.7× less storage. Ablation experiments demonstrate that the diacritical linking module contributes the largest single accuracy gain (+4.4 pp), highlighting the importance of script-specific preprocessing for low-resource languages.

Future work will focus on: (i) scaling the dataset to 100,000+ samples via crowdsourcing; (ii) extending the system to line-level and page-level recognition using attention-based sequence models; (iii) integrating a large language model for context-aware post-correction; and (iv) adapting the architecture to Uzbek Cyrillic and other Central Asian Latin orthographies.

ISSN 3030-3907

REFERENCES

1. UNESCO, "A Guide to Best Practices for the Digitization of Handwritten Cultural Heritage," UNESCO Digital Library, Paris, France, Tech. Rep., 2021.
2. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proc. IEEE, vol. 86, no. 11, pp. 2278–2324, Nov. 1998, doi: 10.1109/5.726791.
3. A. Graves and J. Schmidhuber, "Offline Arabic handwriting recognition using multidimensional recurrent neural networks," in Guide to OCR for Arabic Scripts, Springer, 2012, pp. 297–313.
4. G. Cohen, S. Afshar, J. Tapson, and A. van Schaik, "EMNIST: Extending MNIST to handwritten letters," in Proc. IJCNN, Anchorage, AK, May 2017, pp. 2921–2926, doi: 10.1109/IJCNN.2017.7966217.
5. U.-V. Marti and H. Bunke, "The IAM-database: An English sentence database for offline handwriting recognition," Int. J. Doc. Anal. Recognit., vol. 5, no. 1, pp. 39–46, 2002, doi: 10.1007/s100320200071.
6. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Proc. IEEE CVPR, San Diego, CA, Jun. 2005, pp. 886–893, doi: 10.1109/CVPR.2005.177.
7. I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA: MIT Press, 2016.
8. X.-X. Niu and C. Y. Suen, "A novel hybrid CNN–SVM classifier for recognizing handwritten digits," Pattern Recognit., vol. 45, no. 4, pp. 1318–1325, Apr. 2012, doi: 10.1016/j.patcog.2011.09.021.
9. S. Ahlawat and A. Choudhary, "Hybrid CNN-SVM classifier for handwritten digit recognition," Procedia Comput. Sci., vol. 167, pp. 2554–2560, 2020, doi: 10.1016/j.procs.2020.03.309.
10. A. Graves, M. Liwicki, S. Fernández, R. Bertolami, H. Bunke, and J. Schmidhuber, "A novel connectionist system for unconstrained handwriting recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 5, pp. 855–868, May 2009, doi: 10.1109/TPAMI.2008.137.
11. S. Peng, X. Liu, and C. Liu, "Attention-based Arabic handwriting recognition with multi-scale features," Pattern Recognit. Lett., vol. 138, pp. 610–617, 2020, doi: 10.1016/j.patrec.2020.08.022.
12. I. Lozinskiy, A. Suleymanov, and B. Nurmukhanov, "Low-resource language models for Central Asian languages: Challenges and methods," in Proc. ACL Workshop Low-Resource Lang., 2023, pp. 112–119.
13. S. Iskandarova and Z. Kuchkarov, "Recognition of Uzbek handwritten texts by selection of activation functions for convolutional neural network," Science and Innovation, vol. 2, no. A2, pp. 16–21, 2023.

ISSN 3030-3907

14. S. Iskandarov and M. Rakhmonova, "Theoretical foundations and efficient algorithms for feature detection and extraction," Raqamli Transformatsiya va Sun'iy Intellekt, vol. 3, no. 6, pp. 122–126, Dec. 2025, ISSN: 3030-3346.
15. J. Sauvola and M. Pietikäinen, "Adaptive document image binarization," Pattern Recognit., vol. 33, no. 2, pp. 225–236, Feb. 2000, doi: 10.1016/S0031-3203(99)00055-2.
16. P. Soille, Morphological Image Analysis: Principles and Applications, 2nd ed. Berlin: Springer, 2003.
17. J. Liang, D. Doermann, and H. Li, "Page segmentation using a Voronoi+MST approach," in Proc. SPIE Document Recognit. Retrieval, vol. 5010, 2003, pp. 96–104.
18. J. Cohen, "A coefficient of agreement for nominal scales," Educ. Psychol. Meas., vol. 20, no. 1, pp. 37–46, 1960, doi: 10.1177/001316446002000104.
19. C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," J. Big Data, vol. 6, no. 1, pp. 1–48, Jul. 2019, doi: 10.1186/s40537-019-0197-0.
20. N. Otsu, "A threshold selection method from gray-level histograms," IEEE Trans. Syst., Man, Cybern., vol. 9, no. 1, pp. 62–66, Jan. 1979, doi: 10.1109/TSMC.1979.4310076.
21. W. Niblack, An Introduction to Digital Image Processing. Englewood Cliffs, NJ: Prentice Hall, 1986.
22. D. Bradley and G. Roth, "Adaptive thresholding using the integral image," J. Graph. Tools, vol. 12, no. 2, pp. 13–21, 2007, doi: 10.1080/2151237X.2007.10129236.
23. S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in Proc. ICML, Lille, France, 2015, vol. 37, pp. 448–456.
24. M. Lin, Q. Chen, and S. Yan, "Network in network," in Proc. ICLR, Banff, Canada, 2014, arXiv:1312.4400.
25. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," J. Mach. Learn. Res., vol. 15, no. 1, pp. 1929–1958, 2014.
26. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in Proc. ICLR, San Diego, CA, 2015, arXiv:1412.6980.
27. M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," Inf. Process. Manag., vol. 45, no. 4, pp. 427–437, 2009, doi: 10.1016/j.ipm.2009.03.002.

28. D. Bhatt, C. Patel, H. Ghetiya, D. Pandya, and C. Bhatt, "Survey on various image classification techniques," J. Comput. Sci., vol. 17, no. 3, pp. 235–253, 2021.
29. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in Proc. ICLR, San Diego, CA, 2015, arXiv:1409.1556.
30. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE CVPR, Las Vegas, NV, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
31. A. G. Howard et al., "MobileNets: Efficient convolutional neural networks for mobile vision applications," arXiv:1704.04861, Apr. 2017.
32. S. Ruder, "Neural transfer learning for natural language processing," Ph.D. thesis, Dept. Comput. Sci., Nat. Univ. Ireland, Galway, 2019.

ISSN 3030-3907