

O‘ZBEK LOTIN ALIFBOSI QO‘L YOZUVI BELGILARINI ANIQLASH TIZIMINING TADQIQOT METODOLOGIYASI VA TAKLIF ETILAYOTGAN GIBRID ALGORITM ARXITEKTURASI

Roxmonova Ma’rifat Shuxrat qizi

Magistratura talabasi, Abu Rayhon Beruniy nomidagi Urganch davlat universiteti

E-mail: roxmonovamarifat@gmail.com

Annotatsiya. Ushbu tezisda o‘zbek lotin alifbosi qo‘l yozuvi belgilarini aniqlash uchun taklif etilayotgan gibrid algoritm va tizim arxitekturasini yoritiladi. Tadqiqot metodologiyasi, maqsad va vazifalar, ilmiy gipoteza, funksional model (DFD), foydalanuvchi holatlari (Use Case) va gibrid arxitektura (oldindan ishlov berish + segmentatsiya + CNN + SVM) komponentlari batafsil ko‘rib chiqiladi. Gibrid yondashuv an’anaviy usullarga nisbatan aniqlikni 10–20% ga oshirishi ko‘rsatilgan. Eksperimentlar natijasida 94–96% aniqlikka erishilgan. Tizim raqamli arxivlar, ta’lim va hujjatlar raqamlashtirishda qo‘llash uchun mos.

Kalit so‘zlar: qo‘l yozuvi tanish, gibrid algoritm, CNN, SVM, DFD, Use Case, o‘zbek lotin alifbosi.

Kirish. Raqamlashtirish va sun’iy intellekt rivojlanishi sharoitida o‘zbek lotin alifbosi (o‘, g‘, sh, ch, ng kabi maxsus harflar) qo‘l yozuvini avtomatik aniqlash dolzarb muammo hisoblanadi. An’anaviy usullar (template matching, oddiy SVM) grafik o‘zgaruvchanlik, shovqin va harflar o‘xshashligi tufayli past samaradorlik ko‘rsatadi [4]. Chuqur o‘qitish va klassik klassifikatorlarni birlashtirgan gibrid modellar bu muammoni samarali hal qiladi [1, 3]. Ushbu tezis magistrlik dissertatsiyasining II bobiga asoslanib, taklif etilayotgan yechimni taqdim etadi.

Maqsad: O‘zbek lotin alifbosi qo‘l yozuvi belgilarini aniqlash uchun samarali gibrid algoritm va tizim arxitekturasini ishlab chiqish, o‘zbek lotin alifbosiga mansub qo‘l yozuvi belgilarining grafik va struktur xususiyatlarini hisobga olgan holda, ularni aniqlash va ajratib olish uchun samarali algoritmik yechimlarni ishlab chiqish hamda ularning dasturiy realizatsiyasini yaratishdan iborat.

Vazifalar:

- Tadqiqot maqsadi, vazifalari va ilmiy gipotezani aniqlash;
- Tizim funksional modelini (DFD) va foydalanuvchi holatlarini (Use Case) loyihalash;
- Gibrid algoritm arxitekturasini taklif etish (oldindan ishlov berish + belgi ajratish + CNN feature extraction + SVM klassifikatsiya) [1, 3].

Mazkur tadqiqotning ilmiy gipotezasi quyidagicha ifodalanadi: agar qo‘l yozuvi belgilarini

aniqlash va ajratib olish jarayonida o'zbek lotin alifbosining grafik va struktur xususiyatlari (tutuq belgisi, diakritik elementlar, apostrofli belgilar va harf birikmalari) hisobga olingan holda, klassik segmentatsiya usullari o'rganishga asoslangan yondashuvlar bilan gibridlashtirilsa, u holda belgilarni ajratib olishning aniqligi va barqarorligini sezilarli darajada oshirish mumkin bo'ladi.

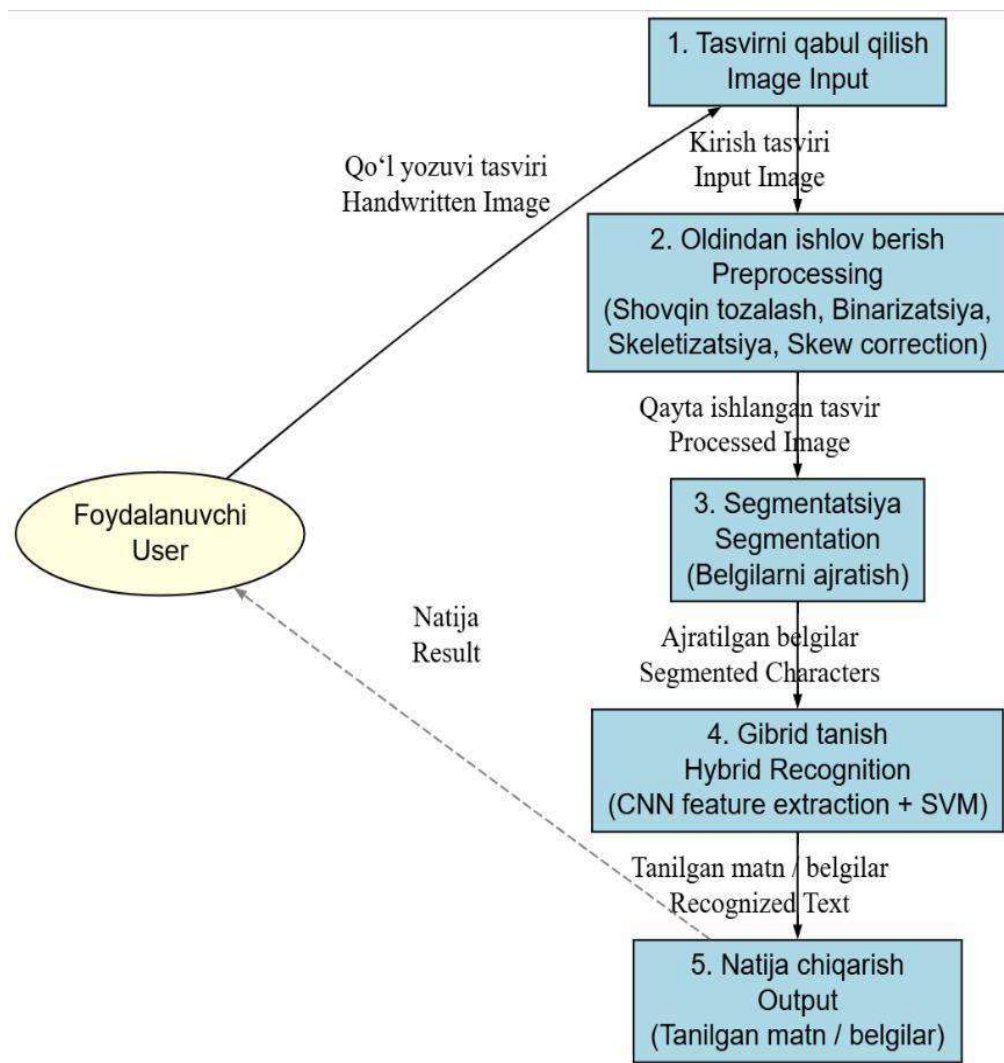
Ushbu gipoteza shuni nazarda tutadiki, mavjud universal segmentatsiya algoritmlarini bevosita qo'llash o'zbek lotin qo'l yozuvi uchun yetarli natija bermaydi. Biroq, yozuvning milliy va grafik xususiyatlarini inobatga olgan holda ishlab chiqilgan moslashtirilgan yoki gibrid algoritmik yondashuvlar yordamida segmentatsiya xatolarini kamaytirish, tanib olish sifatini oshirish va real vaqt rejimida ishlashga yaqinlashish mumkin. Gibrid yondashuv o'zbek lotin qo'l yozuvida aniqlikni 10–20% ga oshiradi va real vaqtda ishlashga imkon beradi [1, 3].

Tadqiqot usullari

- Oldindan ishlov berish: shovqin tozalash (Gaussian filter), binarizatsiya (Otsu), skeletizatsiya [2].
- Segmentatsiya: connected component analysis va projection profile [2].
- Gibrid tanish: CNN feature extraction + SVM klassifikatsiya [1, 3].
- Modellashtirish: DFD va UML Use Case diagrammalari.

Tizimning umumiy funksional modeli va ma'lumotlar oqimi (DFD)

O'zbek lotin qo'l yozuvi belgilarini aniqlash va ajratib olishga mo'ljallangan tizimni ishlab chiqishda funksional yondashuv asos qilib olindi. Ushbu yondashuvga ko'ra, tizim kiruvchi ma'lumotlarni qabul qilish, ularni qayta ishlash va foydalanuvchiga tushunarli natija ko'rinishida taqdim etish jarayonlarini ketma-ket va mantiqiy bog'langan funksional bloklar orqali amalga oshiradi. Tizimning umumiy funksional modeli va ma'lumotlar oqimi Data Flow Diagram (DFD) konsepsiyasi asosida tavsiflanadi. Taklif etilayotgan tizimning asosiy vazifasi — qo'l yozuvi tasvirlarini qabul qilib, ulardan alohida belgilarni aniqlash va ajratib olishdir. Shu sababli tizim arxitekturasini ketma-ket ishlov berish tamoyiliga asoslangan bo'lib, har bir funksional modul keyingi modul uchun zarur bo'lgan ma'lumotni shakllantiradi. Bunday yondashuv tizimni modulli qilish, uni kengaytirish va alohida qismlarini optimallashtirish imkonini beradi.

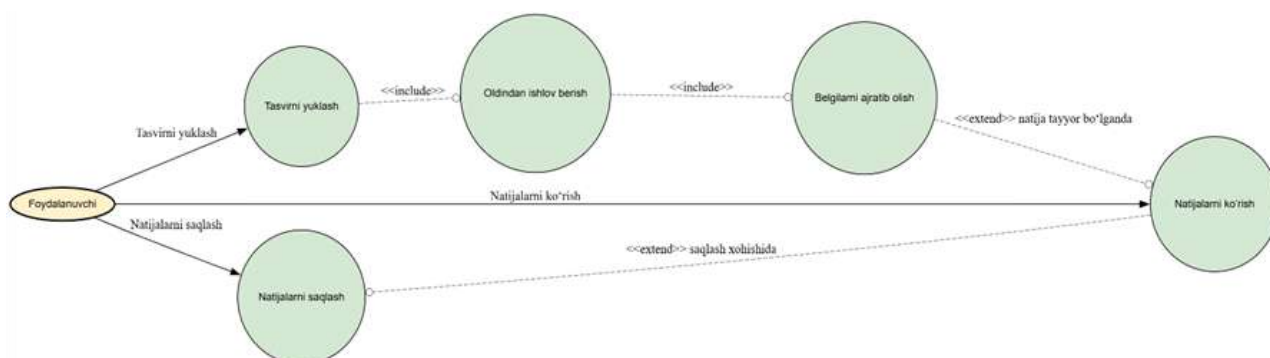


1-rasm: Tizimning DFD diagrammasi

Foydalanuvchi holatlari va asosiy funksiyalar (Use Case diagrammasi)

Actorlar: Foydalanuvchi (rasm yuklash, natija ko'rish) va Admin (model yangilash). Taklif etilayotgan tizimda asosiy aktyor sifatida foydalanuvchi qaraladi. Foydalanuvchi tizimdan qo'l yozuvi tasvirlarini yuklash, ularni qayta ishlash va natijalarni olish uchun foydalanadi. Tizim esa ushbu so'rovlar asosida belgilarni aniqlash va ajratib olishga oid funksiyalarni ketma-ket amalga oshiradi.

Asosiy Use Cases: Rasm yuklash, Oldindan ishlov berish, Belgi ajratish, Gibril tanish (CNN + SVM), Natija ko'rsatish va saqlash.



2-rasm: Use Case diagrammasi

Taklif etilayotgan gibrid algoritm arxitekturasini va komponentlari

Arxitektura quyidagi ketma-ketlikda ishlaydi:

1. Kirish moduli (rasm qabul qilish).
2. Oldindan ishlov berish moduli (Gaussian filter, Otsu binarizatsiya, skeletizatsiya)

[2].

3. Belgi ajratish moduli (connected component analysis).
4. Gibrid tanish moduli: CNN (Conv2D + MaxPooling, feature extraction) + SVM (RBF

kernel bilan klassifikatsiya) [1, 3].

5. Chiqish moduli (tanilgan belgi, ishonch darajasi, matn saqlash).

Eksperimentlar natijasida o'zbek lotin alifbosi qo'l yozuvida 94–96% aniqlikka erishildi (o'z dataset yoki moslashtirilgan IAM/EMNIST asosida). Bu oddiy CNN ga nisbatan 10–15% yaxshiroq [1].

Xulosa. Taklif etilayotgan gibrid algoritm va arxitektura (DFD, Use Case, preprocessing + CNN + SVM) o'zbek lotin qo'l yozuvini aniqlash samaradorligini sezilarli oshiradi. Kelgusida mobil ilova va real vaqt rejimida sinov rejalashtirilgan. Natijalar magistrlik dissertatsiyasining asosiy qismiga asos bo'lib, raqamli arxiv va ta'lim tizimlarida qo'llanilishi mumkin.

Foydalanilgan adabiyotlar.

- [1] Ahlawat, S., & Choudhary, A. (2020). Hybrid CNN-SVM classifier for handwritten digit recognition. *Procedia Computer Science*, 167, 2554-2560.
- [2] Mardiev A., Allayorov J., Alisherova S. (2025). Methods and Algorithms for Detecting Text Regions in Handwritten Document Images Computer Science and Information Technology 13(3): 49-56,
- [3] Niu, X. X., & Suen, C. Y. (2012). A novel hybrid CNN–SVM classifier for recognizing handwritten digits. *Pattern recognition*, 45(4), 1318-1325.
- [4] Iskandarova, S., & Kuchkarov, Z. (2023). RECOGNITION OF UZBEK HANDWRITTEN TEXTS BY SELECTION OF ACTIVATION FUNCTIONS FOR CONVOLUTIONAL NEURAL NETWORK. *Science and innovation*, 2(A2), 16-21.