

UDC 004.352.243

## QO'LYOZMA MATN TASVIRINI GIBRID MODEL YORDAMIDA TANIB OLISH

*Iskandarova S.N.<sup>1</sup>*

<sup>1</sup> Muhammad al Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti,  
Toshkent, O'zbekiston  
sayyora5@mail.ru

**Annotatsiya.** *O'zbek matni qo'lyozma harflarni tanib olishda segmentatsiya masalasi va uni tanib olish muhim hisoblanadi. Qo'lyozma matnning egriligi, bir-birini takrorlashi va tegishi bilan bog'liq muammolar, kursiv bog'lanish, noto'g'ri pozitsiya diakritik belgilar, ko'tarilish va tushish borligi segmentatsiya masalasini qayta ishlashga olib keladi. Qo'lyozmalarda CNN+LSTM+CTC neyron tarmoq modeli asosida dasturiy ta'minotdan olingan natija asosida qator segmentatsiyasi va so'zlarga ajratish samaradorligi ko'rsatib berilgan.*

**Kalit so'zlar:** *segmentatsiya, neyron tarmoq, filtratsiya, topologiya, qo'l yozma, python.*

### I. KIRISH

Matn qatori va so'zlarni ajratish - bu jarayon matnli hujjat rasmidagi asosiy ob'ektlarni ajratib olish. Bu qo'lda yozilgan matn uchun off-line tanib olishda muhim qadamdir. Har qanday grafikada qo'l yozma matn tanib olishda segmentatsiya masalasi tanish samaradorligini oshirishga xizmat qiladi.

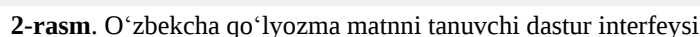
Qo'lda yozilgan matn satrini segmentatsiya qilish uchun yuqoridan pastga, pastdan yuqoriga usullari mavjud. Yuqoridan pastga tushadigan usullar butunlikni qayta ishlaydi. Bu usullarda harflarni ajratib olishda ayrim xatoliklarga uchrashimiz mumkin. Qo'lyozma tasvirda harflarning turli hajmdaligi, qo'shib yozilishi va hattoki so'zlarning turli o'rinda yozilishi boshqa xatlarga qaraganda bir muncha segmentatsiya masalalarida muammolarni keltirib chiqaradi. Kompyuterda terilgan lotin xarflarini tanishda bu muammolar hal etilgan ammo, qo'l yozma matnlarda

qisman ko'plab izlanuvchilar tamonidan ko'rib chiqilgan. Qo'lyozma matn tasvirlarni o'rganish ish samaradorligini oshirishga xizmat qiladi. Ularni avtomatlashtirilgan tarzda kompyuterda terilgan matnga o'tkazish imkoniyatini beradi. Qo'l yozma manbalarni elektron shaklda har bitta harfni qayta ishlash, turli krill va lotin yozuvlariga o'tkazish, Matn mazmunini izohli lug'atlari yordamida xosil qilishni avtomatlashtirish, o'qilishi qiyin manbalarni oson o'qib olish imkoniyatini beradi[1].

### II. ASOSIY QISM

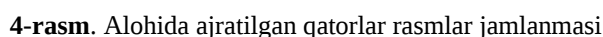
So'zlarning bir-biriga qo'shib ketilishi, harflarning birikib yozilishi lotin yozuvi grafikasida ayrim muammolarni keltirib chiqaradi.

Shu sababli segmentatsiya masalasini to'g'ri yechimini topishda har bitta qatorni alohida ajratib olamiz. Dasturda har bitta qatorni ajratishda ularning o'lchami hisobga olingan.



2. Fanning boshqa fanlar bilan uzviy bog'liqlig.

Ushbu segmentatsiya yolgʻon qatorlarni alohida qator harflari segmentatsiya deb nomlanadi. Matnni aralashtirmasdan ajratib tanishda biz yuqori natija erishishda



To'g'ri chiziqli segmentlangan ma'lumotlarni olish jarayonida balandlik chegarasi qutilari ma'lum chegara chegarasidan past bo'lgan barcha rasmlarni e'tiborsiz qoldiriladi. Tajriba orqali 64 qiymatidan foydalanish kerakli natijani berishini aniqlandi. Bundan tashqari, olingan tasvirlarning tartibini saqlab qolish juda muhimdir. Buning uchun tasvirning boshlang'ich piksel qiymatidan kelib chiqqan holda konturlardan tasvirlarni o'qishi ta-

minlanadi. Shunday qilib, tasvirlar yuqori chap burchakdan pastki o'ng burchakgacha o'qiladi[2,3].

CNN qatlamlarini oziqlantirish uchun Keras TimeDistributed o'ramidan foydalanildi. Bu nima qiladi, u kirishning turli freymlarini oladi va uni kadrma-kadrga qayta ishlaydi. Keyin har bir keyingi kadr uchun 4 qadam bilan kiritilgan tasvirdan 64 x 64 piksellik kadrlar olinadi. Keras TimeDistributed o'rami kirish ramkalarini CNN qatlamlariga o'tkazish uchun ishlatiladi. TimeDistributed o'rami biz kirishdan olingan tasvirlar ramkalarining vaqtinchalik ketma-ketligini saqlashda yordam beradi [4-6].

Har bir yo'nalishda 256 birlikdan iborat uchta ikki yo'nalishli LSTMdan keyin stacked 13 konvolyutsiya qatlamidan foydalanildi. Maksimal birlashtirish ba'zi konvolyutsiya qatlamlaridan keyin qo'llaniladi, jami 5 ta maksimal birlashtiruvchi qatlam qo'llaniladi. Konvolyutsiya qatlamlarida chiziqli bo'lmaganlikni joriy qilish uchun ReLU faollashtirish funksiyasidan foydalaniladi. Konvolyutsiya qatlamlarining og'irliklari normal holatga keltiriladi, shuningdek, tarmoqning tezroq konvergensiyasiga yordam beradigan og'irliklarni normallashtirish uchun barcha konvolyutsiya qatlamlaridan keyin ommaviy normalizatsiya qatlamlari qo'llaniladi. Treningni tark etish darajasi LSTM qatlamlarining har birida 0,25 ga o'rnatiladi. CNN chiqishi va LSTM-larning kiritilishi o'rtasida zich qatlam mavjud bo'lib, bu ma'lumotlarni o'qitish uchun ishlatiladigan parametrlar sonini kamaytirishda juda samarali.

Tanib olish uchun qo'llaniladigan chiqish sinflarining umumiy sonini 90 ga o'rnatildi (A-Z, o'g', 0-9, ingliz klaviaturasidagi barcha standart maxsus belgilar, shuningdek, noma'lum uchun

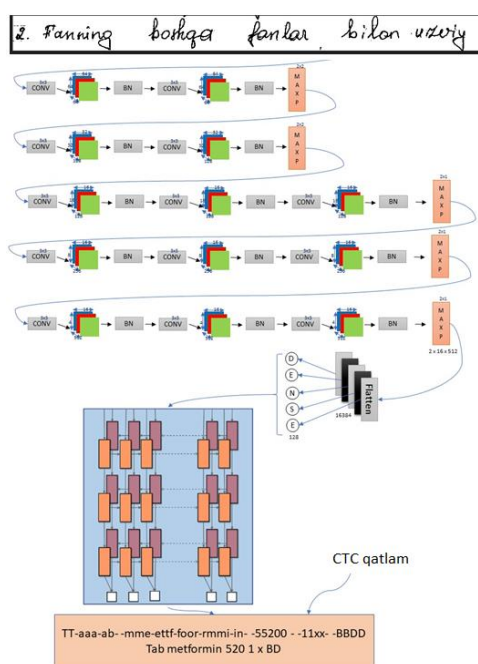
bitta qo'shimcha sinf). Shunday qilib, neyron tarmog'ining oxirgi qatlami LSTM ramkasining har bir chiqishi uchun 90 ta vaznga ega edi. Softmax qatlami oxirgi qatlamda ramkada eng yuqori ehtimollikka ega bo'lgan chiqish sinfini olish uchun ishlatiladi.

Yo'qotishni minimallashtirish uchun foydalaniladigan maqsad funksiyasi Connectionist Temporal Categorical (CTC) yo'qotish funksiyasidir. Boshqa yo'qotish funksiyasi bitta maqsad funksiyasini optimallashtirgan bo'lsa-da, CTC yo'qolishi bashorat qilingan ketma-ketlikning uzunligini va bashorat qilingan ketma-ketlik sinflarini optimallashtirish uchun maxsus ishlab chiqilgan, chunki kirish tasviri tabiatan o'zgaradi. Konvolyutsiya filtrlari va LSTM og'irliklari orqaga tarqalish protsedurasi doirasida birgalikda o'rganiladi. Odam optimallashtiruvchisi boshlang'ich o'rganish tezligi 0,001 bo'lgan trening uchun ishlatiladi.

Tarmoqning ishlashini kuzatish uchun ishlatiladigan ko'rsatkichlar Levenshtein masofasi (tahrirlash masofasi sifatida ham tanilgan) bo'lib, u kuzatilgan ketma-ketlik va bashorat qilingan ketma-ketlik o'rtasidagi farqni o'lchash orqali qator ko'rsatkichlarini o'lchash uchun keng tarqalgan ko'rsatkichdir.

Kiritilgan barcha so'zlarning lug'atini yaratildi. CTC yo'qotish funksiyasi aslida butun so'zni to'g'ri bermaydi, chunki TimeDistributed o'ramidan foydalangan holda oziqlanadigan freymlar o'rtasida o'xshashlik mavjud. Bitta alifbo bir nechta freymlarda bo'lishi mumkinligi sababli, chiqish so'zlarning takrorlanishini o'z ichiga oladi. Shunday qilib, biz butun korpusimizdagi chiqish sifatida berilgan so'zga eng yaqin mos keladigan eng yaqin so'z nima ekanligini aniqlash uchun

metrik sifatida masofani tahrirlashdan foydalanishimiz kerak.



5-rasm. CNN+LSTM+CTC model arxitekturası

Qo'lyozma matn modelini hosil qilishda neyron tarmoq bilan o'qitish amalga oshiriladi. O'qitish jarayonida berilgan satr tasvirlarini segmentatsiya qilish masalasi CNN+LSTM+CTC orqali amalga oshiriladi. Har bir qator topologiyasi va so'zlari o'rganilib chiqilib modelda saqlanadi.

Rasm shakldagi matnni tanib olishda segmentatsiya masalasida gibrid neyron tarmoqlarning qo'llanilishi tanish samaradorligini ko'tarishga xizmat qiladi.

### III. XULOSA VA TAVSIYALAR

Xozirgi kunga qadar tasvirni tanishda turli xil filtratsiya, segmentatsiya va tanib olish algoritmlari alohida-alohida olib borilgan va bular ko'plab yozuv grafiqlarida yuqori tanib olish ko'rsatkichlari bilan ta'minlagan. Dasturlash imkoniyatlarining va neyron tarmoqlar modellarining yangi ko'rinishlari yutuqlaridan foydalanib

tanib olish samaradorligini amalga oshirish mumkin. Matnda harflarni ajrtib olishda so'zlar bazasi saqlanib qolishi uchun CNN(convolutional neural network) orqali asosiy tasvir fragmentlarini LSTM+CTClar yordamida xotirada saqlab, tanishda harflar ketma-ketligida mos so'zdan foydalanish imkoniyatini beradi. Ya'nikim so'z tasvirdagi topologiyasiga yaqin bo'lgan so'zni neyron orqali tanib olinadi. Ushbu metoddan foydalanish qo'lyozma manbalarni tanishda 30-40 foizgacha bo'lgan matnni o'qitish orqali butun kitobni tanib olish topologiyasini shakllantirishga olib keladi.

### ADABIYOTLAR

- [1] S. Espana-Boquera, M. J. Castro-Bleda, J. Gorbe-Moya, and F. ZamoraMartinez, "Improving offline handwritten text recognition with hybrid hmm/ann models," IEEE transactions on pattern analysis and machine intelligence, vol. 33, no. 4, pp. 767–779, 2010.
- [2] T. M. Breuel, "High performance text recognition using a hybrid convolutional-lstm implementation," in 2017 14th IAPR international conference on document analysis and recognition (ICDAR), vol. 1. IEEE, 2017, pp. 11–16.
- [3] H. Scheidl, S. Fiel, and R. Sablatnig, "Word beam search: A connectionist temporal classification decoding algorithm," in 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR). IEEE, 2018, pp. 253–258.
- [4] R. R. Ingle, Y. Fujii, T. Deselaers, J. Baccash, and A. C. Popat, "A scalable handwritten text recognition system," in 2019 International Conference on

- Document Analysis and Recognition (ICDAR). IEEE, 2019, pp. 17–24.
- [5] D. Edmundson and G. Schaefer, “An overview and evaluation of jpeg compressed domain retrieval techniques,” in Proceedings ELMAR-2012. IEEE, 2012, pp. 75–78.
- [6] B. Rajesh, P. Jain, M. Javed, and D. Doermann, “Hh-compwordnet: Holistic handwritten word recognition in the compressed domain,” in 2021 Data Compression Conference (DCC). IEEE, 2021, pp. 362–362.

Поступила в редакцию 22.07.2022

**Citation:** Iskandarova S.N.. Qo‘lyozma matn tasvirini gibrid model yordamida tanib olish. // Raqamli texnologiyalarning nazariy va amaliy masalalari xalqaro jurnali. – 2022. – № 1(1). – B. 48-52.

## USING A HYBRID MODEL FOR IMAGE RECOGNITION OF HANDWRITTEN TEXT

Iskandarova S.N.<sup>1</sup>

<sup>1</sup> Tashkent University of information technologies named after Muhammad al-Khwarizmi, Uzbekistan, Uzbekistan  
sayyora5@mail.ru

**Abstract.** *In the recognition of Uzbek handwritten letters, an important issue is segmentation and its recognition. The curvature of the handwriting, cursive and word-to-word problems, incorrect placement of diacritics, and the presence of ups and downs lead to the processing of the segmentation problem. In the manuscripts, the performance of linear and word segmentation is shown based on results obtained from software based on the CNN+LSTM+CTC neural network model.*

**Keywords:** *segmentation, neural network, filtering, topology, handwriting, python.*

## ИСПОЛЬЗОВАНИЕМ ГИБРИДНОЙ МОДЕЛИ ДЛЯ РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ РУКОПИСНОГО ТЕКСТА

Искандарова С.Н.<sup>1</sup>

<sup>1</sup> Ташкентский университет информационных технологий имени Мухаммада ал-Хорезми, Ташкент, Узбекистан  
sayyora5@mail.ru

**Аннотация.** *В распознавании узбекских рукописных букв важным вопросом является сегментация и ее распознавание. Кривизна рукописного текста, проблемы, связанные со скорописью и касанием слов друг-другу, неправильный расположении диакритических знаков и наличием подъемов и спадов приводят к обработке проблемы сегментации. В рукописях эффективность линейной сегментации и сегментации на слова показана на основе результатов, полученных от программного обеспечения на основе модели нейронной сети CNN+LSTM+CTC.*

**Ключевые слова:** *сегментация, нейронная сеть, фильтрация, топология, почерк, python.*