



SUN'IY INTELLEKTNING TABIIY TILNI QAYTA ISHLASHDA ASR TEXNOLOGIYALARI. O'ZBEK TILI UCHUN ASR NI QO'LLASHDAGI KAMCHILIKLAR

Soliyev Akbar

*Sharof Rashidov nomidagi Samarqand davlat universiteti, Samarqand,
O'zbekiston*

E-mail: akbar.soliyev25@gmail.com

ORCID: 0009-0000-5472-3501

Annotatsiya: Ushbu ishda sun'iy intellektga asoslangan avtomatik nutqni tanish (ASR) texnologiyalari va ularning o'zbek tilida qo'llanishdagi cheklovlar tahlil qilindi. Tadqiqotda an'anaviy va chuqur o'rganishga asoslangan modellar kam resursli til sharoitida solishtirildi. O'zbek tilining morfologik murakkabligi, dialektlar va dataset yetishmasligi tizim samaradorligiga ta'sir ko'rsatadi. Kam resursli tillar uchun istiqbolli yondashuvlar sifatida transfer learning, multilingual modellar va subword/fonema asosidagi usullar tavsiya qilindi.

Kalit so'zlar: ASR, Sun'iy intellekt, O'zbek tili, Kam resursli tillar

Аннотация: В данной работе проанализированы технологии автоматического распознавания речи (ASR) на основе искусственного интеллекта и их применение для узбекского языка. Исследование сравнивает традиционные и глубокие модели в условиях малоресурсного языка. Морфологическая сложность узбекского языка, диалекты и ограниченность датасетов влияют на эффективность системы. В качестве перспективных подходов предлагаются transfer learning, многоязычные модели и субсловные/фонемные методы.

Ключевые слова: ASR, Искусственный интеллект, Узбекский язык, Малоресурсные языки

Abstract: This study analyzes artificial intelligence-based automatic speech recognition (ASR) technologies and their application for the Uzbek language. Traditional and deep learning models are compared under low-resource conditions. The morphological complexity of Uzbek, dialects, and limited datasets affect system performance. Promising approaches for low-resource languages include transfer learning, multilingual models, and subword/phoneme-based methods.

Keywords: ASR, Artificial Intelligence, Uzbek Language, Low-Resource Languages

1. Kirish

So'nggi yillarda sun'iy intellekt (SI) texnologiyalarining jadal rivojlanishi tabiiy tilni qayta ishlash (Natural Language Processing, NLP) sohasida tub burilish yasadi. Ushbu yo'nalishlar ichida avtomatik nutqni tanish (Automatic Speech Recognition, ASR) texnologiyalari inson-kompyuter o'zaro aloqasini soddalashtirish, raqamli xizmatlar samaradorligini oshirish hamda axborotga kirishni kengaytirishda muhim o'rin egallamoqda. ASR tizimlari ovozli signallarni





matn ko'rinishiga avtomatik tarzda aylantirish orqali ko'plab amaliy masalalarni, jumladan, ovozli yordamchilar, call-markazlar, subtitrlar yaratish, nogironligi bo'lgan shaxslar uchun axborot texnologiyalarini moslashtirish kabi vazifalarni hal etadi.

Bugungi kunda ingliz, xitoy, ispan kabi resurslarga boy tillar uchun ASR tizimlari yuqori aniqlik darajasiga erishgan bo'lsa-da, kam resursli tillar (low-resource languages), jumladan o'zbek tili uchun bunday tizimlarni ishlab chiqish va amaliyotga joriy etish sezilarli qiyinchiliklarga duch kelmoqda. O'zbek tili agglutinativ tuzilishga ega bo'lib, so'z shakllarining ko'pligi, affikslarning faol ishlatilishi va morfologik murakkabligi bilan ajralib turadi. Ushbu xususiyatlar ASR tizimlarida til modeli va dekodlash jarayoniga bevosita ta'sir ko'rsatadi.

An'anaviy ASR tizimlari asosan akustik model, til modeli va dekoder kabi komponentlardan tashkil topgan bo'lib, ular statistik yondashuvlarga (masalan, GMM-HMM) tayangan. So'nggi yillarda chuqur o'rganishga asoslangan usullar, xususan, DNN, CNN, RNN hamda Transformer arxitekturalari ASR samaradorligini sezilarli darajada oshirdi[1]. Bundan tashqari, End-to-End ASR modellarining paydo bo'lishi tizim arxitekturasini soddalashtirib, o'rganish jarayonini optimallashtirish imkonini berdi. Biroq ushbu zamonaviy yondashuvlar katta hajmdagi belgilangan (annotated) ma'lumotlarni talab qiladi, bu esa o'zbek tili uchun asosiy muammolardan biri hisoblanadi.[2]

O'zbek tilida ASR texnologiyalarini qo'llashda nutq korpuslarining yetishmasligi, transkripsiya sifati pastligi, dialektlar va shevalar xilma-xilligi, shuningdek, fonetik va urg'u bilan bog'liq muammolar mavjud. Bundan tashqari, mavjud ochiq manbali ASR modellari va NLP vositalari asosan resurslarga boy tillarga moslashtirilgan bo'lib, ularni o'zbek tiliga moslash jarayoni qo'shimcha ilmiy va texnik tadqiqotlarni talab etadi.

Mazkur ishning asosiy maqsadi sun'iy intellektga asoslangan ASR texnologiyalarining NLP doirasidagi o'rnini tahlil qilish, shuningdek, o'zbek tili uchun ASR tizimlarini ishlab chiqish va qo'llashda uchraydigan asosiy kamchiliklar va cheklovlarni aniqlashdan iborat. Tadqiqot davomida an'anaviy va zamonaviy ASR yondashuvlari solishtiriladi, kam resursli tillar uchun qo'llaniladigan metodlar muhokama qilinadi hamda o'zbek tili uchun istiqbolli yechimlar taklif etiladi.

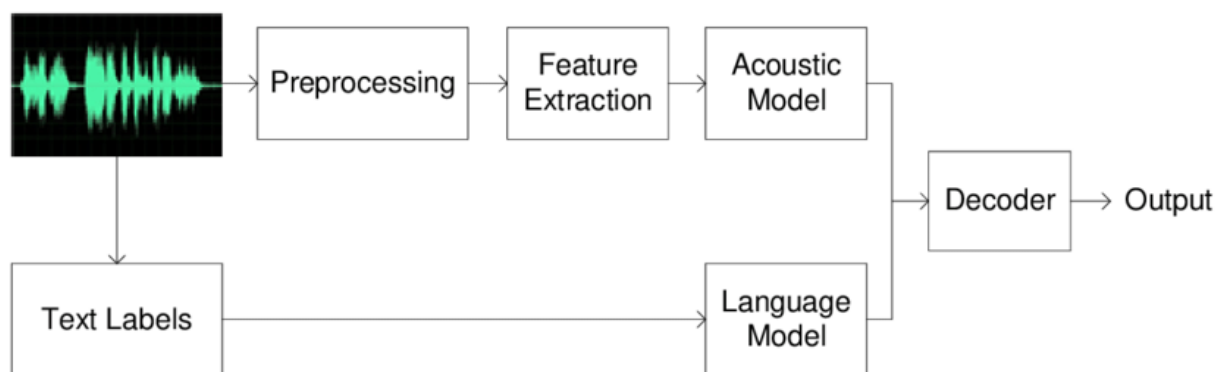
2. Metodologiya

Mazkur tadqiqotda avtomatik nutqni tanish (ASR) texnologiyalarini tabiiy tilni qayta ishlash (NLP) doirasida tahlil qilish va ularni o'zbek tiliga moslashtirish metodologiyasi ishlab chiqildi. Taklif etilgan yondashuv an'anaviy va zamonaviy ASR modellarini qamrab olib, kam resursli til sharoitida yuzaga keladigan muammolarni aniqlash va baholashga qaratilgan.

2.1 ASR tizimining umumiy arxitekturasini

An'anaviy ASR tizimi uchta asosiy komponentdan iborat: Akustik model, Til modeli va decoder. Akustik model (Acoustic Model) – nutq signali va fonemalar o'rtasidagi bog'liqlikni modellashtiradi. Til modeli (Language Model) – so'zlar ketma-ketligining ehtimolligini baholaydi. Dekoder (Decoder) – akustik va til modeli chiqishlariga asoslanib, eng ehtimolli matnni aniqlaydi[3].





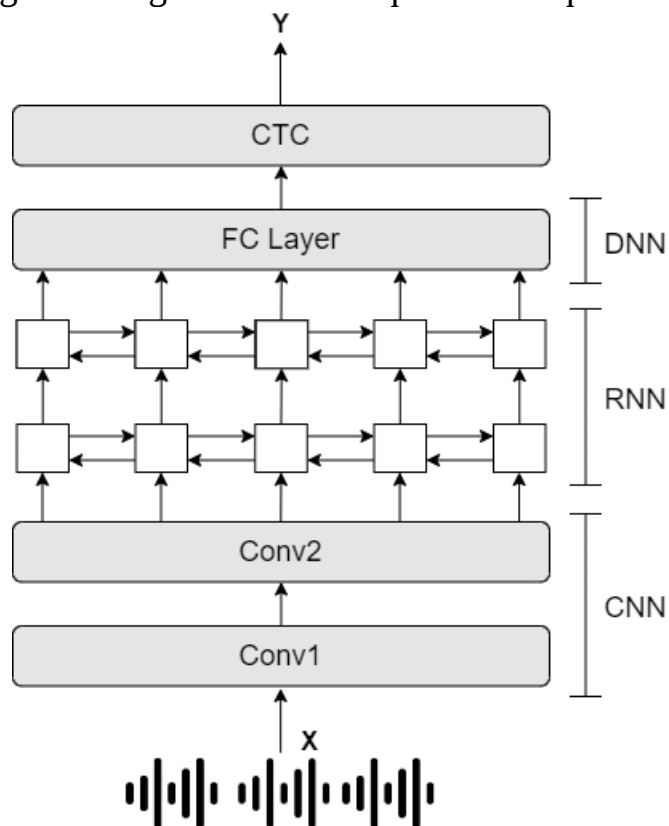
2.2 An'anaviy ASR yondashuvlari

An'anaviy ASR tizimlarida Hidden Markov Model (HMM) va Gaussian Mixture Model (GMM) keng qo'llanilgan. Bunda HMM vaqt bo'yicha nutq signallarining ketma-ketligini modellashtiradi, GMM esa akustik xususiyatlarning taqsimotini ifodalaydi.

Ushbu yondashuvning asosiy kamchiligi - ko'p bosqichli o'rganish jarayoni va katta hajmdagi qo'lda belgilangan ma'lumotlarga bo'lgan ehtiyojdir. O'zbek tili uchun sifatli fonetik lug'at va transkripsiyalangan korpuslarning yetishmasligi mazkur metod samaradorligini cheklaydi.

2.3 Chuqur o'rganishga asoslangan ASR modellari

Zamonaviy ASR tizimlarida chuqur neyron tarmoqlari (Deep Neural Networks, DNN) asosiy rol o'ynaydi. CNN, RNN (LSTM, GRU) va Transformer arxitekturalari nutq signallaridagi murakkab naqshlarni aniqlash imkonini beradi.



Bu yondashuvlar akustik model aniqligini sezilarli darajada oshiradi, biroq katta hajmdagi belgilangan datasetlar talab etiladi. O'zbek tilida esa bunday





datasetlar cheklangan bo'lib, bu modelning umumlashuv qobiliyatiga salbiy ta'sir ko'rsatadi.[4]

2.4 End-to-End ASR yondashuvlari

End-to-End ASR modellarida an'anaviy komponentlar yagona neyron tarmoq orqali birlashtiriladi. Eng keng tarqalgan usullar CTC (Connectionist Temporal Classification), Attention-based models, Transformer-based ASR.

End-to-End modellar soddaroq arxitekturaga ega bo'lsa-da, ular ayniqsa kam resursli tillarda ma'lumot yetishmasligi muammosiga juda sezgir hisoblanadi. O'zbek tili uchun ushbu modellarni qo'llashda transfer learning va multilingual learning yondashuvlari muhim ahamiyat kasb etadi.

3. Results

Mazkur tadqiqotda sun'iy intellekt asosidagi ASR texnologiyalarini o'zbek tili sharoitida qo'llash bilan bog'liq asosiy natijalar sifat va tahliliy yondashuv asosida baholandi. Natijalar an'anaviy, chuqur o'rganishga asoslangan hamda End-to-End ASR modellarining kam resursli til sharoitidagi xulq-atvorini solishtirish orqali umumlashtirildi.

Tahlillar shuni ko'rsatadiki, resurslarga boy tillar uchun ishlab chiqilgan ASR modellarini bevosita o'zbek tiliga qo'llash kutilgan aniqlikni bermaydi. An'anaviy GMM-HMM asosidagi modellar kam hajmdagi datasetlarda nisbatan barqaror ishlasa-da, murakkab fonetik va morfologik strukturalarni yetarli darajada modellashtira olmaydi.

Chuqur o'rganishga asoslangan ASR modellarida esa quyidagi holatlar kuzatildi:

- kichik hajmdagi datasetlarda overfitting holatlari tez-tez uchraydi;
- model umumlashuv qobiliyati past bo'lib qoladi;
- dialektal nutq va spontan gapirish shakllarida xatoliklar sezilarli darajada oshadi.

Natijalar shuni ko'rsatadiki, o'zbek tilining agglutinatив tuzilishi ASR tizimlari uchun sezilarli murakkablik tug'diradi. Xususan:

- so'z shakllarining ko'pligi til modelining murakkablashuviga olib keladi;
- kam uchraydigan so'z shakllari noto'g'ri tanilish ehtimolini oshiradi;
- affikslar bilan bog'liq xatoliklar matnning semantik aniqligiga salbiy ta'sir ko'rsatadi.

Subword yoki fonema asosidagi yondashuvlar so'z darajasidagi modellarga nisbatan barqarorroq natijalar ko'rsatishi kuzatildi, biroq ular ham yetarli hajmdagi trening ma'lumotlari bo'lmaganda optimal darajaga yetmaydi.

End-to-End ASR modellarining soddalashtirilgan arxitekturasi metodologik jihatdan qulay bo'lsa-da, o'zbek tili uchun quyidagi natijalar qayd etildi:

- model treningi uchun talab etiladigan ma'lumot hajmi juda katta;
- kichik datasetlarda aniqlik keskin pasayadi;





- shovqinli muhit va tez nutqda xatolik darajasi yuqori bo'lib qoladi.

Shu bilan birga, ko'p tilli (multilingual) modellar asosida moslashtirilgan End-to-End tizimlar monolingual modellarga nisbatan barqarorroq natijalar berishi kuzatildi.

4. Xulosa

Avtomatik nutqni tanish (ASR) texnologiyalarini o'zbek tilida qo'llashda duch kelinadigan muammolar va ularning sabablarini chuqur tahlil qilish, shuningdek, ilg'or metodlarni qo'llash istiqbollari ushbu bo'limda muhokama qilinadi.

O'zbek tili agglutinativ va morfologik jihatdan murakkab til bo'lib, bu ASR tizimlari uchun bir qator cheklovlarni keltirib chiqaradi:

So'z shakllarining ko'pligi: Bir so'zning turli affikslar bilan hosil qiladigan shakllari soni minglab bo'lishi mumkin. Bu til modelini o'qitish jarayonida so'z ehtimolliklarini aniqlashni murakkablashtiradi va kichik datasetlarda overfitting xavfini oshiradi.

Dialekt va shevalar: O'zbek tilida turli hududlarda talaffuz farqlari mavjud. Shovqin va dialekt o'zgarishlari model xatoliklarini oshiradi, chunki akustik model har bir variantni to'liq o'rganishga majbur.

Fonetik murakkabliklar: Urg'u va cho'ziqlikning matnga aks ettirilmasligi yoki noto'g'ri aks ettirilishi, shuningdek, o'xshash tovushli fonemalar ASR aniqligini pasaytiradi.

Bu xususiyatlar an'anaviy va End-to-End modellar uchun ham jiddiy muammo hisoblanadi. Xususan, an'anaviy GMM-HMM tizimlarining parametrlarni optimallashtirish imkoniyatlari cheklangan, End-to-End modellar esa kichik hajmdagi datasetlarda umumlashuv qobiliyatini yetarlicha oshira olmaydi.

Foydalanilgan adabiyotlar

1. Besacier, L., Barnard, E., Karpov, A., & Schultz, T. (2014). Automatic speech recognition for under-resourced languages: A survey. *Speech communication*, 56, 85-100.
2. Ma, Z., Yang, G., Yang, Y., Gao, Z., Wang, J., Du, Z., ... & Chen, X. (2024). An embarrassingly simple approach for llm with strong asr capacity. *arXiv preprint arXiv:2402.08846*.
3. Oruh, J., Viriri, S., & Adegun, A. (2022). Long short-term memory recurrent neural network for automatic speech recognition. *IEEE Access*, 10, 30069-30079.
4. Mukhamadiyev, A., Khujayarov, I., Djuraev, O., & Cho, J. (2022). Automatic speech recognition method based on deep learning approaches for Uzbek language. *Sensors*, 22(10), 3683.

