

**МУХИДДИНОВ МУҲРИДИН НУРИДИН ЎҒЛИ**

**ТАСВИРЛАРДАГИ ОБЪЕКТЛАРНИ АЖРАТИШ УСУЛЛАРИ ВА**  
**АЛГОРИТМЛАРИНИ ИШЛАБ ЧИҚИШ**

05.01.01 – Мухандислик геометрияси ва компьютер графикаси.  
Аудио ва видеотехнологиялари

**ТЕХНИКА ФАНЛАРИ БЎЙИЧА ФАЛСАФА ДОКТОРИ (PhD)**  
**ДИССЕРТАЦИЯСИ АВТОРЕФЕРАТИ**



**ТОШКЕНТ АХБОРОТ ТЕХНОЛОГИЯЛАРИ УНИВЕРСИТЕТИ**  
**ҲУЗУРИДАГИ ИЛМӢЙ ДАРАЖАЛАР БЕРУВЧИ**  
**DSc.13/30.12.2019.T.07.02 РАҚАМЛИ ИЛМӢЙ КЕНГАШ**  

---

**ТОШКЕНТ АХБОРОТ ТЕХНОЛОГИЯЛАРИ УНИВЕРСИТЕТИ**

**МУХИДДИНОВ МУҲРИДДИН НУРИДДИН ЎҒЛИ**

**ТАСВИРЛАРДАГИ ОБЪЕКТЛАРНИ АЖРАТИШ УСУЛЛАРИ ВА**  
**АЛГОРИТМЛАРИНИ ИШЛАБ ЧИҚИШ**

05.01.01 – Мухандислик геометрияси ва компьютер графикаси.  
Аудио ва видеотехнологиялари

**ТЕХНИКА ФАНЛАРИ БЎЙИЧА ФАЛСАФА ДОКТОРИ (PhD)**  
**ДИССЕРТАЦИЯСИ АВТОРЕФЕРАТИ**

Техника фанлари бўйича фалсафа доктори (PhD) диссертацияси мавзуси Ўзбекистон Республикаси Вазирлар Маҳкамаси ҳузуридаги Олий аттестация комиссиясида В2019.4.PhD/T1387 рақам билан рўйхатга олинган.

Диссертация Тошкент ахборот технологиялари университетида бажарилган.

Диссертация автореферати уч тилда (Ўзбек, инглиз, рус (резюме)) Илмий кенгаш веб саҳифасида ([www.tuit.uz](http://www.tuit.uz)) ва «Ziyounet» ахборот-таълим порталида ([www.ziyounet.uz](http://www.ziyounet.uz)) жойлаштирилган.

**Илмий раҳбар:** Хамдамов Уткир Раҳматиллаевич  
техника фанлари доктори, доцент

**Расмий оппонентлар:** Кучкарова Дилором Файзуллаевна  
техника фанлари доктори, профессор

Рахимов Бахтияр Саидович  
техника фанлари номзоди, доцент

**Етакчи ташкилот:** Фаргона политехника институти

Диссертация ҳимояси Тошкент ахборот технологиялари университети ҳузуридаги DSc.13/30.12.2019.T.07.02 рақамли Илмий кенгашнинг 2020 йил «26» сентябр да соат 12<sup>00</sup> даги мажлисида бўлиб ўтади. (Манзил: 100084, Тошкент шаҳри, Амир Темур кўчаси, 108-уй. Тел.: (99871) 238-64-43; факс: (99871) 238-65-52; e-mail: [tuit@tuit.uz](mailto:tuit@tuit.uz)).

Диссертация билан Тошкент ахборот технологиялари университетининг Ахборот-ресурс марказида танишиш мумкин (154 рақам билан рўйхатга олинган). Манзил: 100084, Тошкент шаҳри, Амир Темур кўчаси, 108-уй. Тел.: (99871) 238-64-43).

Диссертация автореферати 2020 йил «    » \_\_\_\_\_ кунни тарқатилди.  
(2020 йил «    » \_\_\_\_\_ даги \_\_\_\_\_ рақамли реестр баённомаси).



**И.Х. Сиддиқов**  
Илмий даражалар берувчи  
илмий кенгаш раиси,  
т.ф.д., профессор

**Х.Э. Хужаматов**  
Илмий даражалар берувчи  
илмий кенгаш илмий котиби,  
техника фанлари бўйича фалсафа доктори (PhD)

**Ф.М. Нуралиев**  
Илмий даражалар берувчи  
илмий кенгаш ҳузуридаги илмий семинар раиси,  
т.ф.д., доцент

## **КИРИШ (фалсафа доктори (PhD) диссертацияси аннотацияси)**

Диссертация мавзусининг долзарблиги ва зарурати. Жаҳонда тасвирлардан муҳим объектларни ажратиб олиш, контурини аниқлаш ва матнларни таниб олиш тадқиқотларига катта қизиқиш ва алоҳида эътибор берилмоқда. Тасвирлардаги визуал маълумотларни аниқ ва самарали олиш учун рақамли ишлов беришда филтрлаш, сиқиш, аниқлаш, ажратиш ва таниш масалаларини ечишда янги компьютер кўриши алгоритмларини яратиш ҳамда ижтимоий соҳаларда қўллаш учун дастурий воситаларни ишлаб чиқиш масалалари ҳам муҳим аҳамият касб этмоқда. Инсонларнинг кўриш тизими тасвирлардаги муҳим объектларни визуал равишда тез ва аниқ таниш учун ақл бовар қилмайдиган қобилиятга эга бўлиб, бу равшанлик деб аталади. Ушбу визуал равшанлик механизми тасвирдаги баъзи объектларни атрофидаги нарсалардан ажратиб туради. Бу эса кўриш қобилияти заифларга мураккаб табиий тасвирларни идрок этишда муҳим ҳисобланади. Тасвирдаги муҳим объектларни ажратиб олиш ҳамда матнларни таниш алгоритмлари ва усулларини ишлаб чиқиш муаммолари бўйича АҚШ, Германия, Буюк Британия, Хитой, Япония, Ҳиндистон, Жанубий Корея ва Россия Федерацияси каби дунёнинг ривожланган мамлакатларида илғор тадқиқотлар олиб берилмоқда.

Жаҳонда ақли қурилмалар сонининг ортиши ва уларнинг универсал бўлиб бориши, ижтимоий қўллаб-қувватлаш, объектларни таниш, ёрдамчи навигация, геокодинг каби тасвир ҳамда видео билан ишлайдиган дастурий воситаларга эҳтиёж ортиб боришини кўрсатмоқда ҳамда уларга қаратилган қатор илмий тадқиқот ишлари олиб борилмоқда. Ушбу соҳада, кўзи ожиз шахслар визуал маълумотларни идрок этиши учун табиий тасвирлардаги муҳим объектларни ажратиб олиш, контурларини белгилаш ва тактил графикасида ифодалаш усуллари, алгоритмлари ҳамда дастурий воситаларини ишлаб чиқишга алоҳида эътибор қаратилмоқда. Шу билан бирга табиий тасвирлардан ўзбек тилидаги матнларни белгилаш, ажратиб олиш, таниш ва товушли талаффуз қилиш алгоритмларини ҳамда дастурий воситасини ишлаб чиқиш энг долзарб масалалардан бири ҳисобланади.

Республикамизда ҳозирги кунда ногиронларнинг манфаатларини ҳимоя қилиш, уларни реабилитация қилиш ва ижтимоий мослашишга ёрдам беришга алоҳида эътибор қаратилмоқда. Тасвирларга ишлов бериш усулларини қўллаган, илмий ва инновацион технологиялар билан жиҳозланган ихтисослаштирилган мактаб-интернатлари, марказлар ва кутубхоналар ташкил этиш бўйича кенг қамровли чора-тадбирлар амалга оширилмоқда. 2017-2021 йилларда Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегиясида, жумладан «...имконияти чекланган шахсларни давлат томонидан қўллаб-қувватлашни кучайтириш, ижтимоий хизмат кўрсатишни яхшилаш, ...кўзи ожизлар учун таълим олиш

учун кенг имкониятлар яратиш...»<sup>1</sup> вазифалари белгиланган. Мазкур вазифаларни амалга ошириш, жумладан кўзи ожизлар учун визуал маълумотларни идрок этишлари учун тасвирдаги муҳим объектларни ажратиб олиш алгоритмлари асосида тактил графикаси яратиш, тасвирлардаги матнларни таниш алгоритмлари орқали уларни талаффуз қилувчи нутқ синтезатори дастурий воситаларини ишлаб чиқиш ҳамда амалга ошириш муҳим масалалардан ҳисобланади.

Ўзбекистон Республикаси Президентининг 2017 йил 7 февралдаги ПФ-4947-сон «Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегияси тўғрисида»ги ва 2017 йил 1 декабрдаги ПФ-5270-сон «Ногиронлиги бўлган шахсларни давлат томонидан қўллаб-қувватлаш тизимини тубдан такомиллаштириш чора-тадбирлари тўғрисида»ги Фармонлари, Ўзбекистон Республикаси Вазирлар Маҳкамасининг 2018 йил 18 ноябрдаги 739-сон «Ногиронлиги бўлган шахсларга ахборот-кутубхона хизмати кўрсатиш тизимини такомиллаштириш тўғрисида»ги Қарори ҳамда мазкур фаолиятга тегишли бошқа меъёрий-ҳуқуқий ҳужжатларда белгиланган вазифаларни амалга оширишда ушбу диссертация тадқиқоти муайян даражада хизмат қилади.

Тадқиқотнинг республика фан ва технологиялари ривожланишининг устувор йўналишларига мослиги. Мазкур тадқиқот республика фан ва технологиялари ривожланишининг IV. «Ахборотлаштириш ва ахборот-коммуникация технологияларини ривожлантириш» устувор йўналиши доирасида бажарилган.

Муаммоннинг ўрганилганлик даражаси. Тасвирлардаги муҳим объектларни аниқлаш ва ажратиб олиш, матнни аниқлаш ва таниб олиш, шунингдек, самарали ҳисоблаш алгоритмлари ва дастурий воситаларини ишлаб чиқишни ўрганиш ҳамда амалга оширишнинг мақбул параметрларини аниқлаш учун ҳисоблаш тажрибалари натижаларини таҳлил қилиш бўйича кенг қамровли тадқиқотлар олиб борилмоқда. Бундай усуллар ва алгоритмларни яратишга йуналтирилган масалалар Springer, Association for Computing Machinery, IEEE Xplore, Cyberleninka, Web of Science, Elsevier нашриётларининг маълумотлар базасига киритилган, асосан, IEEE Xplore ва Scopus рўйхатидаги илмий журналларда охириги йилларда чоп этилган мақолалардан ўрганилди.

Визуал ахборотни олиш учун тасвирлардаги объектларни ажратиш ва матнни таниш муаммолари жаҳон илмий адабиётида анча кенг ёритилган. Ушбу тадқиқот йўналишига М.М.Ченг, Л.Итти, А.Ахилиш, Р.Смит, Ш.Ли, Ю.И.Журавлев, Л.Неуманн, М.Б.Медведев, С.Гоферман, А.С.Кузнецов, Х.Зоу, Д.Брадлей, Б.Эпштейн, С.Яо, С.Жинсоо, Д.Е.Супрун, Н.Такаги, М.Жадерберг, Р.Марголин, П.Ф.Фелзенсвалб, Т.Дутоит ва бошқа таниқли хорижий олимларнинг ишлари бағишланган. Шунингдек, Республикамизда тимсолларни таниб олиш ва тасвирларга ишлов беришнинг назарий

<sup>1</sup>Ўзбекистон Республикаси Президентининг 2017 йил 7 февралдаги ПФ-4947-сон «Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегияси тўғрисида»ги Фармони

асосларини ривожлантиришга М.М.Камилов, Ш.Х.Фозилов, Н.М.Мирзаев, У.Р.Ҳамдамов, С.С.Раджабов, Т.С.Жумаев, А.А.Маҳкамов, О.Н.Джураев ҳамда бошқа олим ва тадқиқотчилар ушбу йўналишдаги илмий-тадқиқотларга ўз ҳиссаларини қўшишган.

Соҳага оид олиб борилган таҳлил натижалари шуни кўрсатдики, тасвирлардан муҳим объектларни ажратиб олиш ва матнни таниш йўналишларининг ривожланишига қарамай, кўзи ожизлар учун тактил графикасини яратишда тасвирлардаги муҳим объектларни аниқлаш ва ажратиб олиш, шунингдек матнни нутққа айлантириш механизмларини ишлаб чиқишда тасвирлардаги матнларни аниқлаш ва таниб олиш усуллари ҳамда алгоритмлари долзарб бўлиб, ўзбек тили учун етарли даражада ўрганилмаган.

Диссертация тадқиқотининг диссертация бажарилган олий таълим муассасасининг илмий-тадқиқот ишлари режалари билан боғлиқлиги. Диссертация тадқиқоти Мухаммад ал-Хоразмий номидаги Тошкент ахборот технологиялари университетининг илмий-тадқиқот ишлари режасининг №БВ-Ф4-011 - «Сигналларни ва тасвирларни қайта ишлаш масалаларида параллел ҳисоблаш назарияси, методлари ва воситаларни ривожлантириш» (2017-2019) ва №БВ-Атех-2018-251 – «Кўзи ожиз шахслар учун компьютер техникасидан фойдаланиш, матнларни ўқиш ва ёзиш имконини берувчи ўзбек тилига асосланган гапирувчи дастурий таъминот ва овоз синтезаторини ишлаб чиқиш» (2018-2019) мавзуларидаги илмий лойиҳалари доирасида бажарилган.

Тадқиқотнинг мақсади тасвирлардаги муҳим объектларни ажратиб олиш, контурларини аниқлаш ва матнларни таниш, товушли талаффуз қилиш усуллари, алгоритмлари ҳамда дастурий воситасини ишлаб чиқишдан иборат.

#### **Тадқиқотнинг вазифалари:**

тасвир ранг қийматларининг глобал контрастини ошириш ва локал мослашувчан уч қийматли чегаралаш асосида тасвирдаги муҳим объектларни ажратиб олиш усуллари таҳлил қилиш;

тўлиқ конволюцион нейрон тармоқ модели асосида табиий тасвирлардан ўзбек тилидаги матнларни белгилаш, ажратиб олиш ва таниш жараёнининг алгоритмик усулини ишлаб чиқиш;

тасвирлардаги график объектларни аниқлаб, уларни ички ва ташқи чегараларини белгилаган ҳолда тактил графикасида ифодалаш алгоритми ва дастурий воситасини ишлаб чиқиш;

кўзи ожиз шахслар идрок этиши учун табиий тасвирлардан ўзбек тилидаги матнларини товушли талаффуз қилиш алгоритминини яратиш ва дастурий воситасини ишлаб чиқиш.

Тадқиқотнинг объекти сифатида табиий тасвирлар, видео маълумотлар, ўзбек тили овоз синтезатори учун сўзлар ва аудио маълумотлар ҳамда тасвирларга ишлов бериш воситалари олинган.

Тадқиқотнинг предметини тасвирдаги объектларни аниқлаш, ажратиш ва матнларни таниб олиш муаммоларини самарали ҳал қилишда тасвирларга ишлов бериш алгоритмлари, нейрон тармоқ моделлари ҳамда дастурлаш кутубхоналари ташкил этади.

Тадқиқотнинг усуллари. Тадқиқот жараёнида тасвирларга ишлов беришнинг замонавий усуллари, объектларни аниқлаш, матнларни аниқлаш, чуқур ўрганиш, эҳтимоллик назарияси, математик таҳлил, алгоритмлаш, тузилмалар дастурлаш технологияси ва ҳисоблаш экспериментларни ўтказиш усуллари қўлланилган.

Тадқиқотнинг илмий янгилиги куйидагилардан иборат:

тасвир ранг қийматларининг глобал контрастини ошириш ва локал мослашувчан уч қийматли чегаралаш асосида тасвирдаги муҳим объектларни ажратиб олиш усули такомиллаштирилган;

тўлиқ конволюцион нейрон тармоқ модели асосида табиий тасвирлардан ўзбек тилидаги матнларини белгилаш, ажратиб олиш ва таниш жараёнининг алгоритмик усули ишлаб чиқилган;

тасвирлардаги график объектларни аниқлаб, уларни ички ва ташқи чегараларини белгилаган ҳолда тактил графикасида ифодалаш алгоритми ва дастурий воситаси ишлаб чиқилган;

кўзи ожиз шахслар идрок этиши учун тасвирлардаги ўзбек тили матнларини товушли талаффуз қилиш алгоритми ва дастурий воситаси яратилган.

Тадқиқотнинг амалий натижаси куйидагилардан иборат:

кўзи ожизлар учун тактил графикаларини яратиш жараёнида тасвирлардаги муҳим объектларни аниқлаш ва ажратиб олиш бўйича дастурий воситаси яратилган;

кўзи ожизлар учун ўзбек тилидаги матнларни талаффуз қилувчи овоз синтезаторини яратиш жараёнида тасвирлардаги матнни аниқлаш ва таниб олиш дастурий воситаси яратилган;

таклиф этилган тасвирлардаги объектларни аниқлаш ва ажратиб олиш дастурий воситаси 77-сонли ихтисослаштирилган кўзи ожизлар мактаб-интернатида ҳамда Халқ таълими соҳасида ахборот-коммуникация технологияларини ривожлантириш марказида муваффақиятли татбиқ этилган, шунингдек кўзи ожизларга визуал маълумот бериш учун тактил графикаларни яратиш жараёнида қўлланилган;

таклиф этилган тасвирлардаги матнларни аниқлаш ва таниш дастурий воситаси Ўзбекистон Кўзи ожизлар жамияти марказий бошқарувида ҳамда Имконияти чекланган ёшлар ва болалар марказида муваффақиятли татбиқ этилган, шунингдек кўриш қобилиятига эга бўлмаганлар учун ўзбек тилидаги матнларини талаффуз қилувчи овоз синтезаторини яратиш жараёнида қўлланилган.

Тадқиқот натижаларининг ишончлилиги. Тадқиқот натижаларининг ишончлилиги таклиф этилаётган усуллар ва алгоритмларни бошқа замонавий усуллар билан миқдорий ва сифат жиҳатидан таққослаш орқали аниқ ва



самарали натижаларга эришилганлиги ҳамда мос гувоҳномалар ва жорий қилинганлик тўғрисидаги далолатномалар билан асосланган. Бундан ташқари, таклиф қилинаётган муҳим объектни аниқлаш, муҳим объектни ажратиб олиш, матнларни аниқлаш ва матнларни таниб олиш усуллари тажрибавий ҳисоблаш ва солиштириш орқали баҳоланган.

**Тадқиқот натижаларининг илмий ва амалий аҳамияти.** Тадқиқот натижаларининг илмий аҳамияти шундан иборатки, тасвир ранг қийматларининг гистограммасини тенглаштириш ёрдамида глобал контрастни оширишга асосланган объектни аниқлаш усули, интеграл тасвирлардан фойдаланган ҳолда локал мослашувчан уч қийматли чегаралаш асосида объектни ажратиб олиш усули, тўлиқ конволюцион нейрон тармоғига асосланган матнни аниқлаш усули ва Tesseract OCR моделини ўзбек тилининг лотин ва кирилл алифболари билан ўқитиш асосида ўзбек тилидаги матнларни таниш усуллари билан изоҳланади.

Тадқиқот натижаларининг амалий аҳамияти шундан иборатки, кўзи ожизлар учун тактил графикаларини яратишда тасвирлардаги объектларни ажратиб олиш, уларни контурларини аниқлаш дастурий воситаси, шунингдек, ўзбек тилида талаффуз қилувчи овоз синтезаторини яратишда тасвирлардан матнни аниқлаш ва таниб олиш дастурий воситаси ишлаб чиқилганлиги билан изоҳланади.

**Тадқиқот натижаларининг жорий қилиниши.** Кўзи ожизлар учун тасвирлардаги объектларни ажратиб олиш, матнларни таниш ва товушли талаффуз қилиш усуллари, алгоритмлари ҳамда дастурий воситаларини яратиш бўйича олинган натижалар асосида:

тасвир ранг қийматларининг глобал контрастини ошириш ва локал мослашувчан уч қийматли чегаралаш асосида тасвирдаги муҳим объектларни ажратиб олишнинг такомиллаштирилган усули, алгоритми ва дастурий воситаси Ўзбекистон Республикаси Халқ таълими вазирлиги 77-сонли ихтисослаштирилган кўзи ожизлар мактаб-интернатида ишлаб чиқариш ва ўқув жараёнларига жорий қилинган (Ахборот технологиялари ва коммуникацияларини ривожлантириш вазирлигининг 2019 йил 3 декабрдаги 33-8/8552-сон маълумотномаси). Натижада кўзи ожиз ўқувчилар томонидан уларга тақдим этилган турли тасвир объектларидан иборат тактил графикаларининг 74 фоизини аниқ таниш имконига эришилган;

тўлиқ конволюцион нейрон тармоқ модели асосида табиий тасвирлардан ўзбек тилидаги матнларни белгилаш, ажратиб олиш ва таниш жараёнининг алгоритмик усули, алгоритми ва дастурий воситаси Ўзбекистон кўзи ожизлар жамияти марказий бошқарувида ахборот-хизмат кўрсатиш жараёнларига жорий қилинган (Ахборот технологиялари ва коммуникацияларини ривожлантириш вазирлигининг 2019 йил 3 декабрдаги 33-8/8552-сон маълумотномаси). Натижада электрон файллардаги ўзбек тили матнларини компьютер тизимида талаффуз қилиш орқали кўзи ожизларнинг брайл ёзувидаги маълумотларга нисбатан қисқа вақт ичида катта ҳажмдаги

ахборотларни олишига эришилган ҳамда тасвирлардаги ёзувларни ўқиш ва англаш имконияти яратилган;

тасвирлардаги муҳим объектларни ажратиб олиш, матнларни таниш ва товушли талаффуз қилиш усуллари, алгоритмлари ва дастурий воситаси Ўзбекистон ёшлар иттифоқи қошидаги Имконияти чекланган ёшлар ва болалар маркази ҳамда Халқ таълими соҳасида ахборот-коммуникация технологияларини ривожлантириш марказининг ишлаб чиқариш, ўқув ва ижтимоий хизмат кўрсатиш жараёнларига жорий қилинган (Ахборот технологиялари ва коммуникацияларини ривожлантириш вазирлигининг 2019 йил 3 декабрдаги 33-8/8552-сон маълумотномаси). Натижада кўзи ожиз шахсларнинг касб-хунарга йўналтирилган профессионал кўникмаларини шакллантириш бўйича ўқув курслари ва тўғарақларида тактил графикасидаги ўқув, тарқатма, кўргазмавий материаллардан фойдаланиш орқали уларнинг касбий қизиқиши ва қобилиятларини ривожлантириш, тасаввур оламини бойитиш ҳамда жамиятдаги имкониятларини кенгайтиришга эришилган.

**Тадқиқот натижаларининг апробацияси.** Мазкур тадқиқот натижалари 5 та халқаро ва 2 та республика илмий-амалий анжуманларида маъруза қилинган ва муҳокамадан ўтказилган.

**Тадқиқот натижаларининг эълон қилинганлиги.** Диссертация мавзуси бўйича жами 23 та илмий иш чоп этилган, шулардан, Ўзбекистон Республикаси Олий аттестация комиссиясининг диссертациялар асосий илмий натижаларини чоп этишга тавсия этилган илмий нашрларда 8 та мақола, шундан 4 таси республика ва 4 таси хорижий журналларда, 3 та илмий мақолалар бошқа журналларда нашр қилинган, шунингдек ЭҶМ учун яратилган дастурнинг расмий рўйхатдан ўтказилганлиги тўғрисида 5 та гувоҳнома олинган.

**Диссертациянинг тузилиши ва ҳажми.** Диссертация таркиби кириш, тўртта боб, умумий хулоса, фойдаланилган адабиётлар рўйхати, шартли белгилар ва атамалар рўйхати ҳамда иловалардан иборат. Диссертациянинг ҳажми 117 бетни ташкил этган.

## **ДИССЕРТАЦИЯНИНГ АСОСИЙ МАЗМУНИ**

**Кириш** қисмида диссертация мавзусининг долзарблиги ва зарурияти асосланган, тадқиқотнинг мақсади ва вазифалари шакллантирилган, объект ва предмети тавсифланган, Ўзбекистон Республикаси фан ва технологиялар тараққиётининг устувор йўналишларига мослиги кўрсатилган, тадқиқотнинг илмий янгилиги ва амалий натижалари баён қилинган, олинган натижаларнинг ишончилиги асослаб берилган, тадқиқот натижаларини амалда жорий қилиш ҳолати, нашр этилган ишлар ва диссертация тузилиши бўйича маълумотлар келтирилган.

**Диссертациянинг «Тасвирлардаги объектларни ажратиш ҳамда таниб олиш усуллари ва алгоритмларини таҳлил қилиш»** деб номланган биринчи бобида, мавжуд усуллар изоҳланган ва таҳлил қилинган, шунингдек муаммонинг долзарблиги аниқланган. Тасвирлардаги объектларни

ажратишнинг ва матнларни танишнинг самарали усуллари бўйича хорижий давлатлар ва Республикамизда олиб борилаётган илмий ва амалий тадқиқот ишлари атрофлича таҳлил қилинган. Тадқиқотлар натижасида компьютер кўришининг муаммоли жиҳатларидан бири бўлган объектни ажратиш олиш ва таниш усуллари яратиш ҳамда тасвирлардаги муҳим объектларни ва матнларни самарали аниқлашда уларнинг ранг қийматларининг контрасти муҳим омиллардан бири эканлигига хулоса қилинган. Шунингдек, муҳим объектларни мураккаб табиий тасвирлардан ажратиш олиш усулида кўп қийматли чегаралашни ҳисоблаш орқали бир қийматли чегаралашга нисбатан барқарор натижаларга эришиш мумкинлиги аниқланган.

Бундан ташқари, табиий тасвирлардаги матнларни аниқлаш ва таниш усуллари компьютер кўришининг ечим топмаган ва мураккаб муаммоси ҳисобланади, чунки матнлар одатда табиий тасвирнинг кичик қисмини эгаллайди ва улар нотекис муҳитга эга бўлиб, самарали натижалар олиш учун кўш нурига аксанганини, турли кичик тўсиқларни, доғларни ҳамда тасвир шовқинларини инобатга олиш зарур. Тасвирлардаги матнларни таниш усулларида дастлаб матнларни жойлашган жойини аниқлаб олиш натижанинг ишончли бўлишида муҳимдир. Бундай, матнларнинг тасвирдаги жойлашган ўрнини аниқлаш муаммоларига ечим топишда турли нейрон тармоқ моделларини қўллаш энг самарали усуллардан бири эканлиги таҳлил қилинган.

Шунингдек, диссертациянинг ушбу бобида кўзи ожизларнинг визуал ахборотларни идрок этишларида тасвирлардан ажратиш олинган объект контурларининг жуда содда ва тушунишга осон бўлиши тактил графикани ишлаб чиқишда муҳим вазифалардан бири эканлиги аниқланган. Визуал ахборотлар одатда тасвирлар, шакллар, графикалар, диаграммалар, ранглар ва бошқаларни ўз ичига олади. Шунинг назарда тутиш лозимки, кўзи ожизлар визуал ахборотларни асосан қуйидаги иккита усул орқали тушунишади: (1) ушлаб кўриш, яъни графикалар, тасвирлар, диаграммалар ва рақамлар тактил графикаси ёрдамида тақдим этилиши керак; (2) эшитиш, яъни кўриш қобилияти чекланганларга атроф-муҳит ҳақида овозли маълумот бериш зарур.

Манбаларнинг таҳлили шуни кўрсатадики, кўзи ожизлар учун тасвирлардаги муҳим объектларнинг контурларини тактил графикасида содда ва осон тушунарли қилиб яратиш кераклиги ҳамда ўзбек тили синтезатори учун тасвирлардаги матнларни таниш усули ва алгоритмларини яратиш лозимлиги аниқланди.

Диссертация ишида қўйилган мақсадга эришиш учун ўтказилган тадқиқотлар диссертация доирасида ишлаб чиқилган тадқиқот методологияси асосида олиб борилди. Тадқиқот методологияси объектларни аниқлаш, ажратиш ва уларнинг контурларини белгилаш ҳамда матнларни аниқлаш ва таниш учун тасвирларга рақамли ишлов бериш усуллари таҳлил қилишни, мавжуд муаммоларга ечим топиш учун самарали усуллар ва алгоритмларни яратишда компьютер кўриш воситалари ва кутубхоналарини

тадқиқ қилиш ва танлашни, объектларни ажратиб олиш ва матнларни таниш алгоритмлари асосида тактил графикаларини ҳамда матнларни ўзбек тилида талаффуз қилиш дастурий воситаларини ишлаб чиқишни ўз ичига олади.

Диссертациянинг «Кўзи ожизлар учун тасвирлардаги муҳим объектларни ажратиб олиш усуллари ва алгоритмларини яратиш» деб номланган иккинчи бобида глобал контрастни ошириш асосида муҳим объектларни аниқлаш усули ҳамда локал мослашувчан уч қийматли чегаралаш асосида муҳим объектларни ажратиб олиш усули ишлаб чиқилган. 1-расмда кўрсатилганидек, тасвирлардаги муҳим объектларни аниқлаш ва ажратиб олиш бўйича ишлаб чиқилган усуллар қуйидагилардан иборат: 1) гистограммани тенглаштириш асосида глобал контрастни ошириш, 2) объектнинг муҳимлилик харитасини яратиш, 3) локал мослашувчан уч қийматли чегаралашга асосланган тасвирдаги муҳим худудларни ажратиб олиш, ва 4) муҳим объектнинг ташқи ва ички контурларини аниқлаш.

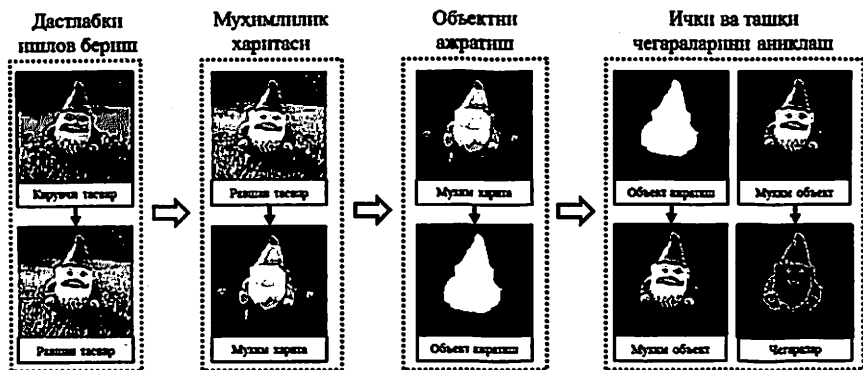
Тасвирнинг ранг қийматлари гистограммасини тенглаштиришга асосланган глобал контрастни ошириш усули атрофидаги объектларга нисбатан муҳим объектнинг контрастни ошириш учун тасвирга дастлабки ишлов бериш босқичида қўлланилади. Тасвирдаги худудларни таққослаш орқали атрофидаги худудларга нисбатан юқори контрастга эга бўлган худуд муҳим объект эканлигини англатади. Тасвир пикселининг равшанлиги унинг тасвирдаги барча бошқа пикселларга нисбатан ранг контрастини солиштириш орқали аниқланади ва  $O(N^2)$  вақтни олади. Буни ҳисоблаш хатто ўртача ўлчамдаги тасвирлар учун ҳам узоқ вақтни олади. Шу сабабли ҳисоблаш вақтини қисқартириш мақсадида тасвирларни сегментация қилиш усули ёрдамида тасвир худудларга ажратилади. Худуднинг равшанлик қиймати тасвирдаги бошқа худудлар билан унинг ранг контрастини таққослаш орқали қуйидагича ҳисобланади:

$$S(r_k) = \sum_{r_k \neq r_l} w(r_l) D(r_k, r_l) \quad (1)$$

бунда  $w(r_l) - r_l$  худуднинг қиймати,  $D(r_k, r_l) -$  иккита худуд ўртасидаги ранглар оралиғининг ўлчови ҳисобланади. Яқинроқ худудларнинг таъсирини кучайтириш ва узоқ худудларнинг таъсирини камайитириш учун худудларнинг жойлашган жойи ҳақидаги маълумот  $S(r_k)$  киритилади. Жойлашиш жойи қийматини ҳисоблашда  $D(r_k, r_l) - r_k$  ва  $r_l$  худудлар орасидаги жойлашиш оралиғини англатади,  $\sigma_s$  жойлашиш оралиғи қиймати кучини назорат қилади.  $w(r_l) - r_l$  худуднинг қиймати бўлиб  $r_l$  худуддаги пикселлар сонидан аниқланади,  $w_s(r_k) = \exp(-9d_k^2)$  жойлашиш қиймати атамаси ҳисобланиб,  $d_k$  тасвирнинг ўртаси билан  $r_k$  худуддаги пикселлар орасидаги ўртача ораликни белгилайди ҳамда  $\sigma_s^2 = 0.4$  га тенг бўлади.

$$S(r_k) = w_s(r_k) \sum_{r_k \neq r_l} e^{-\frac{D_s(r_k, r_l)}{\sigma_s^2}} w(r_l) D_r(r_k, r_l) \quad (2)$$

Бундан ташқари, локал мослашувчан уч қийматли чегаралашга асосланган интеграл тасвирлардан муҳим объектларни ажратиб олиш усули таклиф этилди. 1-расмда кўрсатилгандек тасвирлардаги объектларни аниқлаш ва ажратиб олиш бўйича ишлаб чиқилган усуллар тўртта, яни дастлабки ишлов бериш, муҳимлилик харитаси, объектни ажратиш, ички ва ташқи чегараларини аниқлаш



1- расм. Тасвирлардаги муҳим объектларни аниқлаш ва ажратиб олиш усуллари

Бунда, муҳимлилик харитасидаги кулранг тасвир орқали интеграл тасвир қийматлари ҳисоблаб чиқилади, сўнгра ҳар бир пиксел учун интеграл тасвир ёрдамида маҳаллий ойнани ( $S \times S$ ) аниқлаш ҳамда таққослаш орқали локал мослашувчан уч қийматли чегаралаш қиймати ҳисобланади. Ушбу уч қийматли чегаралаш асосида муҳимлилик харитаси тўртта худудларга ажратилади: аниқ орқа фон, эҳтимолий орқа фон, эҳтимолий объект, аниқ объект. GraphCuts усулидан фойдаланиб, тўртта худудни такрорий ҳисоблашлардан сўнг тўлиқ ҳажмдаги юқори сифатли оқ-қора тасвир яратилади.

Локал мослашувчан уч қийматли чегаралаш усулини ишлаб чиқишдан асосий мақсад, муҳимлилик қийматлари кичик бўлган муҳим объектларни аниқ ажратиб олишдан иборат. Бироқ локал чегаралаш усулларида глобал чегаралаш усулларига қараганда ҳисоблаш учун кўпроқ вақт сарфлаганлиги сабабли, интеграл тасвирларни ҳисоблаш орқали локал чегаралаш усулининг ҳисоблаш тезлигини оширишга эришилди. Интеграл тасвир (йиғилган майдон жадвали деб ҳам номланади) усули пикселлар қийматларининг йиғиндисини ҳисоблашда тез ва самарали усул ҳисобланади ва қуйидагича амалга оширилади:

$$I(x, y) = i(x, y) + I(x - 1, y) + I(x, y - 1) - I(x - 1, y - 1) \quad (3)$$

Бундан ташқари, муҳим объектнинг контурларини аниқлашда Канни контурларни аниқлаш усулини қўллаш самарали натижа олишнинг имконини беради. Ушбу объектларнинг ташқи контурларидан кўзи ожизларга тасвирлардаги визуал маълумотларни етказиш учун тактил графикасини яратишда қўлланилади. Бироқ, олиб борилган тадқиқотлар шуни кўрсатадики баъзи ҳолларда кўзи ожизлар учун объектнинг фақат ташқи контурлари тўғрисидаги маълумотлар етарли ҳисобланмайди. Шу сабабли, объектнинг ички контурларини ҳам аниқлаш тасвирдаги визуал ахборотни тўлиқ идрок этишига хизмат қилади. Тасвирдаги муҳим объектнинг ички контурларини аниқлашда асл кириш тасвир матричасидан оқ-қора тасвир матричаси элементларининг қиймати 1 га тенг бўлган пикселларидан нусха олинад ва қуйидагича амалга оширилади:

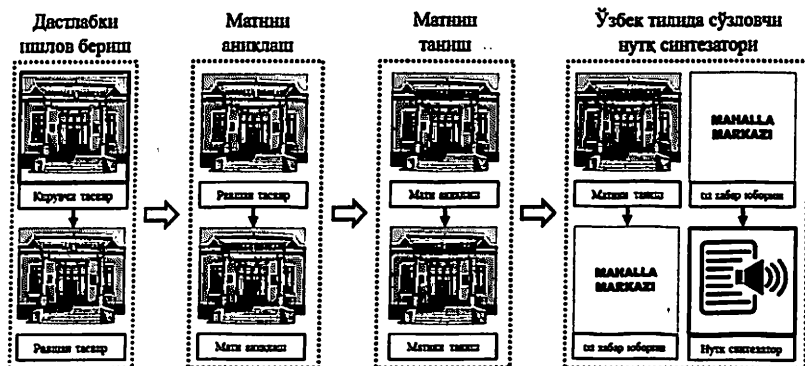
$$S_o = B_m(x, y) * T_i(x, y) \quad (4)$$

бунда  $S_o$  – муҳим объект,  $B_m(x, y)$  – оқ-қора тасвир,  $T_i(x, y)$  – эса асл кириш тасвиридир.

Тадқиқот натижаларига асосланган ҳолда хулоса қилиб айтадиган бўлсак, содда ва тушунарли тактил графикасини яратиш учун тасвирдаги муҳим объектларни ажратиб олиш ва унинг контурларини аниқлаш энг самарали усуллардан бири ҳисобланади.

Диссертациянинг «Нейрон тармоқ асосида матнларни аниқлаш ҳамда таниш усули ва алгоритмини ишлаб чиқиш» деб номланган учинчи бобида тўлиқ конволюцион нейрон тармоқ ва Tesseract OCR моделлари асосида тасвирдаги матнни аниқлаш ва таниш усулларини амалга ошириш келтирилган. 2-расмда кўрсатилгандек, матнларни аниқлаш ва таниш усуллари тўртта асосий босқичда амалга оширилади: 1) гистограммани тенглаштириш асосида глобал контрастни ошириш, 2) тўлиқ конволюцион нейрон тармоқ асосида матнни аниқлаш, 3) Tesseract OCR модели асосида матнни ажратиш ва таниш ҳамда 4) аниқланган сўзларни ўзбек тилида талаффуз қилиш. Тасвирдаги матнларни таниш усули икки қисмдан иборат: матнни аниқлаш ва матнни таниш. Тадқиқот натижаларига асосланган ҳолда шуни таъкидлаш керакки, анъанавий ва нейрон тармоқ моделларига асосланган мавжуд ёндашувлар кўп вақт сарфлайдиган, кўп босқич ва қисмлардан иборат усуллар эканлиги аниқланди. Шунга кўра, бундай ёндашувларнинг самарадорлиги ва аниқлиги ҳали етарли даражада эмас.

Ушбу камчиликларни бартараф этиш учун икки босқичдан иборат бўлган матнни тез ва ишончли аниқлаш усули ишлаб чиқилди. Таклиф этилган усулда қисқа вақтда сўзлар ёки матн сатрларининг жойлашган ўрнини баҳолайдиган тўлиқ конволюцион нейрон тармоқ модели қўлланилган. Тасвирдаги матнларни аниқ жойлашган ўрнини аниқлаш учун конволюцион нейрон тармоқ моделидан кўра тўлиқ конволюцион нейрон тармоқ моделидан фойдаланиш юқори аниқликдаги натижаларни олиш имконини беради.



2- расм. Тасвирлардаги матнни аниқлаш ва таниб олиш усуллари

Бундан ташқари, матнни аниқлаш учун нейрон тармоқ моделларини яратишда турли ҳолатларни ҳисобга олиш талаб этилади. Табиий тасвирлардаги матнли ҳудудларнинг бир-биридан кескин фарқ қилганлиги сабабли, узун матнлар мавжудлигини аниқлаш учун нейрон тармоғининг охириги босқичларидаги хусусиятлари талаб қилиниши, шунингдек қисқа матнлар мавжудлигини аниқлашда бошланғич босқичлардаги хусусиятлари талаб қилиниши аниқланди. Шу сабабли, нейрон тармоқ модели ушбу талабларни қондириш учун турли босқичлардаги хусусиятлардан фойдаланиб амалга оширилиши зарур. Таклиф этилган тўлиқ конволюцион нейрон тармоқ моделида бирин-кетин турли босқичлардаги хусусиятларни бирлаштириш билан бир қаторда қисқа ҳисоблаш вақтини сақлаб қолишга ҳам эришилди. Моделни уч қисмга ажратиш мумкин: хусусият ажратувчи, хусусиятларни бирлаштирувчи ва натижа чиқиш қисми. Хусусиятларни ажратувчи қисми ImageNet базасида олдиндан ўқитилган конволюцион тармоқ ҳисобланади. Хусусият хариталарининг тўртта даражаси  $f_i$  сифатида кўрсатилган, ўлчамлари мос равишда асл кириш тасвирининг  $1/32$ ,  $1/16$ ,  $1/8$  ва  $1/4$  қисмларидан иборат бўлган хусусиятларни ажратувчидан олинади. Хусусиятларни бирлаштириш куйидагича амалга оширилади:

$$g_i = \begin{cases} \text{unpool}(h_i) & \text{агар } i \leq 3 \\ \text{conv}_{3 \times 3}(h_i) & \text{агар } i = 4 \end{cases} \quad (5)$$

$$h_i = \begin{cases} f_i & \text{агар } i = 1 \\ \text{conv}_{3 \times 3}(\text{conv}_{1 \times 1}([g_{i-1}; f_i])) & \text{акс ҳолда} \end{cases} \quad (6)$$

бунда  $g_i$  – бирлашма базаси,  $h_i$  – бирлаштирилган хусусиятлар харитаси ва  $[\ ]$  оператори канал ўқи билан боғланишни аңглатади.

Матн ҳудуди аниқлангандан сўнг кейинги босқичларда, матнни ажратиб уни таниш жараёни амалга оширилади. Бунинг учун ўзбек лотин ва кирилл алифбосидаги ҳарфлар билан ўқитилган Tesseract OCR моделидан

фойдаланилди. Ишлаб чиқилган матнни таниш усули тасвирдан матнларни таниб, ўзбек тилида сўзларни талаффуз қилиш учун нутқ синтезаторига юборади.

Диссертациянинг ушбу бобида ўзбек тили луғатидаги сўзларни ўрганиш ва таҳлил қилиш натижасида 31,5 минг сўздан иборат электрон луғат яратилди ва алфавит тартибда жойлаштирилди. Ўзбек тилида сўзловчи нутқ синтезатори конкатенация усулига асосланган бўлиб сўзларнинг талаффузини ўз ичига олади. Шу сабабли, ўзбек тили электрон луғатидаги 31,5 минг сўзлар ўрганилди ва барча сўзлар 2,5 минг қисмга, яъни бўғинларга ажратилди. Тасвирлардан таниб олинган матнларни тўғри талаффуз қилиш ва ўзбек тилидаги электрон луғат базасини янгилаб бориш учун танилган матнлар электрон луғат базаси билан таққосланади. Агар сўз электрон луғат базасида мавжуд бўлса уни ўзбек тилидаги нутқ синтезаторига юборади, акс ҳолда янги сўзни тасдиқлаш учун экспертга юборилади.

Диссертациянинг «Тасвирлардаги объектларни ажратиш ҳамда таниш усулларини амалга ошириш ва тажрибавий баҳолаш» деб номланган тўртинчи бобида тасвирлардаги объектларни аниқлаш ва ажратиш олиш алгоритмлари, матнларни аниқлаш ва таниш алгоритмларининг дастурий воситалари ишлаб чиқилган. Ишлаб чиқилган алгоритмларнинг самарадорлигини аниқлаш бўйича таққослаш тажрибалари, баҳолаш ва амалий татбиқ этиш натижалари ёритилган. Тасвирдан объектни ажратиш олиш ва матнни таниш алгоритмлари C++ дастурлаш тилида OpenCV кутубхонасидан фойдаланиб амалга оширилган.

Ишлаб чиқилган усуллар энг оммабоп тасвир базаларидан фойдаланган ҳолда бошқа замонавий муқобил усуллар билан таққосланган: муҳим объектларни аниқлаш усули 12 та муқобил усуллар билан MSRA 10K базасидан фойдаланиб таққосланди; муҳим объектларни ажратиш усули 2 та муқобил усуллар билан MSRA 10K базасидан фойдаланиб таққосланди; матнни аниқлаш усули 17 та муқобил усуллар билан ICDAR 2015 ва MSRA-TD500 базаларидан фойдаланиб таққосланди; матнни таниш усули 7 та муқобил усуллар билан ICDAR 2013 базасидан фойдаланиб таққосланди. Шунингдек, таклиф этилган усулларни бошқа усуллар билан миқдорий ва сифат жиҳатдан таққослаш ҳам амалга оширилди.

Таклиф этилган тасвирлардаги муҳим объектларни аниқлаш ва ажратиш олиш усуллари сифат жиҳатидан таққосланганда, ишлаб чиқилган усуллар ёрдамида фонларни самарали бостириб олд объектларни аниқ ажратишга эришилди, шунингдек муҳимлилик харитасидан қийматлари кичик бўлган муҳим объектларни самарали ажратиш олиш ва тасвирдаги бир нечта муҳим объектларни биргаликда ажратиш имконияти яратилди. Матнни аниқлаш ва таниб олиш усулини сифат жиҳатидан таққослашда матн ҳудудлари самарали ва аниқ аниқланди, шунингдек ўзбек тилидаги матнларни тўғри таниб олишга эришилди.

Диссертация доирасида ишлаб чиқилган усулларни миқдорий жиҳатдан таққослашда шу йўналишдаги стандарт баҳолаш ўлчовлари: Аниқлик (P),



Таъсирчанлик (R), F-ўлчов (FM), Қабул қилгичнинг ишлаш хусусиятлари (ROC), эгри чизик остидаги майдон (AUC) ва ўртача мутлақ хатолик (MEA) ўлчовлардан фойдаланилди. Аниқлик ва таъсирчанлик қуйидагича амалга оширилади:

$$\text{аниқлик} = \frac{TDO\alpha GT}{ADO} \quad (7)$$

$$\text{таъсирчанлик} = \frac{TDO\alpha GT}{GT} \quad (8)$$

бунда TDO тўғри аниқланган муҳим объект, ADO барча аниқланган ҳудудлар, ва GT қўлда ишлаб чиқилган рост натижа. F-ўлчов аниқлик ва таъсирчанлик ўртасидаги мувозанат ўлчов бирлигидир ва у қуйидагича амалга оширилади:

$$F\beta = \frac{(1+\beta^2)\text{аниқлик} \times \text{таъсирчанлик}}{\beta^2 \times \text{аниқлик} + \text{таъсирчанлик}} \quad (9)$$

бунда  $\beta^2 = 0.3$  га тенг.

Шунингдек, таклиф этилган маттни аниқлаш усули ICDAR 2015 ва MSRA-TD500 базаларида энг юқори натижаларга эришиш имконини берди. Бу мос равишда 84,7 фоиз ва 84 фоиз F-ўлчов бирлигини ташкил этди. Бундан ташқари, таклиф этилган маттни таниш усули бошқа усулларни ортда қолдириб 94,6 фоиз аниқликда таниш даражасини тақдим этди.

Диссертация доирасида ишлаб чиқилган тасвирдаги муҳим объектни ажратиш олиш ва уни контурларини аниқлаш, матнларни таниб олиш ҳамда талаффуз қилиш усуллари ва алгоритмларидан нафақат тактил графикаси ёки ўзбек тили нутқ синтезаторларида, балки бошқа ахборот технологиялари соҳаларида ҳам фойдаланиш мумкин.

## ХУЛОСА

«Тасвирлардаги объектларни ажратиш усуллари ва алгоритмларини ишлаб чиқиш» мавзусидаги диссертация бўйича олиб борилган тадқиқотлар натижасида қуйидаги хулосалар тақдим этилади:

1. Тасвирнинг ранг қийматлари гистограммасини тенглаштириш ёрдамида унинг глобал контрастини ошириш орқали тасвирнинг муҳим объектларини аниқлаш усули ва алгоритмлари такомиллаштирилди ҳамда 12 та муқобил усуллар билан таққосланди. Тажриба асосида стандарт ўлчов бирликларидан фойдаланиб эришилган натижалар баҳоланди.

2. Локал мослашувчан уч қийматли чегаралашни қўллаб тасвирнинг муҳим объектларини ажратиш олиш усули ва алгоритмлари яратилди ҳамда 2 та муқобил усуллар билан таққосланди. Натижада, тасвирдаги орқа ва олд ҳудудлар маълумотлари жуда ўхшаш бўлган тақдирда ҳам 94-фоиз аниқликда муҳим объектларни ажратиш олиш амалга оширилди.

Ўзбекистон Республикаси  
 Ўқитувчи-адиби  
 ТУШИ

3. Тажриба натижалари шунини кўрсатдики, бошқа муқобил усуллар ёрдамида тасвирдаги фақат битта муҳим объектни ажратиш олиш ҳамда тасвирдаги муҳим объект пикселларининг қийматлари кичик бўлганда муҳим объектларни ажратиш бўлмастиди аниқланди. Бироқ таклиф қилинган усул ёрдамида тасвирдаги бир нечта муҳим объектларни биргаликда аниқлаш ҳамда муҳимлилик хариталарида кичик қийматларга эга бўлган объектларни ҳам орқа ҳудудлар сифатида таснифламасдан ажратиш олишга эришилди.

4. Тўлиқ конволюцион нейрон тармоқ модели негизда тасвирдаги матнларни аниқлаш усули ва алгоритмлари ишлаб чиқилди ҳамда 17 та муқобил усуллар билан таққосланди. Натижада, ICDAR 2015 базасида F-ўлчов бирлигининг 84,7 фоиз, MSRA-TD500 базасида F-ўлчов бирлигининг 84 фоизи миқдорида аниқликка эришилди.

5. Ўзбек тили алифбосига ўқитилган Tesseract OCR модели асосида ўзбек тилидаги матнларни таниш усули ва алгоритмлари ишлаб чиқилди ҳамда 7 та муқобил усуллар билан таққосланди. Натижада таниш олиш кўрсаткичи бўйича 94,6 фоиз аниқликка эришилди.

6. Кўзи ожизлар учун тасвирдаги муҳим объектларнинг ташқи ва ички контурларини аниқлаш ҳамда тактил принтердан фойдаланиб тактил графикасида тасвирлаш алгоритми ишлаб чиқилди. Натижада тасвирлардаги муҳим объектларнинг контурларини юқори аниқликда тактил графикасида тасвирлашга эришилди.

7. Кўзи ожизлар учун тасвирлардан ажратиш олинган матнларни таниш ва ўзбек тили нутқ синтезаторида талаффуз қилиш алгоритми ишлаб чиқилди. Натижада табиий тасвирлардаги ва электрон файллардаги ўзбек тили матнларини компьютер тизимида талаффуз қилишга эришилди.

8. Ишлаб чиқилган тасвирдаги муҳим объектларни ажратиш олиш ҳамда матнларни таниш усуллари, алгоритмлари ва дастурий воситаси “77-сонли ихтисослаштирилган кўзи ожизлар мактаб-интернати”, “Ўзбекистон кўзи ожизлар жамияти марказий бошқаруви”, “Имконияти чекланган ёшлар ва болалар маркази” ҳамда “Халқ таълими соҳасида ахборот-коммуникация технологияларини ривожлантириш маркази” нинг ишлаб чиқариш, ижтимоий хизмат кўрсатиш ва таълим жараёнларига татбиқ этилди. Натижада тасвирлардаги объектларнинг контурларини тактил графикасида тасвирлаш ҳамда матнларни ўзбек тилида талаффуз қилиш орқали кўзи ожиз ёшлар ва ўқувчиларнинг эҳтиёжи, касбий қизиқиши, қобилиятлари ҳамда жамиятдаги имкониятларини кенгайтиришга эришилди.

**SCIENTIFIC COUNCIL AWARDING SCIENTIFIC DEGREES  
DSc.13/30.12.2019.T.07.02 AT TASHKENT UNIVERSITY OF  
INFORMATION TECHNOLOGIES**

---

**TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES**

**MUKHIDDINOV MUKHRIDDIN NURIDDIN UGLI**

**DEVELOPMENT OF METHODS AND ALGORITHMS FOR OBJECTS  
EXTRACTION IN IMAGES**

05.01.01 – Engineering geometry and computer graphics.  
Audio and video technologies

**DISSERTATION ABSTRACT OF THE DOCTOR OF PHILOSOPHY (PhD)  
ON TECHNICAL SCIENCES**

**Tashkent – 2020**

The theme of doctor of philosophy (PhD) on technical sciences was registered at the Supreme Attestation Commission at the Cabinet of Ministers of the Republic of Uzbekistan under number B2019.4.PhD/T1387.

The dissertation has been prepared at Tashkent University of Information Technologies.

The abstract of the dissertation is posted in three languages (Uzbek, English, Russian (resume)) on the website ([www.tuit.uz](http://www.tuit.uz)) and on the website of «Ziyonet» Information and educational portal ([www.ziyonet.uz](http://www.ziyonet.uz).)

**Scientific adviser:** **Khamdamov Utkir Rakhmatillaevich**  
Doctor of Technical Sciences, Docent

**Official opponents:** **Kuchkarova Dilorom Fayzullaevna**  
Doctor of Technical Sciences, Professor

**Rakhimov Bakhtiyar Saidovich**  
Candidate of Technical Sciences, Docent

**Leading organization:** **Ferghana Polytechnic Institute**

The defense of the dissertation will be held « \_\_\_\_ » \_\_\_\_\_ 2020 at \_\_\_\_\_ at the meeting of Scientific council No. DSc.13/30.12.2019.T.07.02 at Tashkent University of Information Technologies (Address: 100084, Tashkent city, Amir Temur street, 108. Ph.: (+99871) 238-64-43, fax: (+99871) 238-65-52, e-mail: [tuit@tuit.uz](mailto:tuit@tuit.uz)).

The dissertation can be reviewed at the Information Resource Centre of Tashkent University of Information Technologies (is registered under No. \_\_\_\_). (Address: 100084, Tashkent city, Amir Temur street, 108. Ph.: (+99871) 238-64-43, fax: (+99871) 238-65-52).

The abstract of dissertation was sent out on « \_\_\_\_ » \_\_\_\_\_ 2020 y.  
(mailing protocol No. \_\_\_\_ on « \_\_\_\_ » \_\_\_\_\_ 2020 y.).



**I.Kh. Siddikov**  
Chairman of the Scientific Council  
awarding scientific degrees,  
Doctor of Technical Sciences, Professor

**H.E. Khujamatov**  
Scientific Secretary of Scientific Council  
awarding scientific degrees,  
PhD on Technical Sciences

**F.M. Nuraliev**  
Chairman of the Academic Seminar at the  
Scientific Council awarding scientific degrees,  
Doctor of Technical Sciences, Docent

## INTRODUCTION (abstract of PhD dissertation)

**The relevance and demanding of the dissertation topic.** A large number of research interest and special attention are raised to developing salient objects extraction, contour detection and text recognition in natural scene images throughout the world. In digital image processing, the issues of filtering, compression, detection, extraction and recognition are also important in the development of new computer vision algorithms and the development of software tools for social sphere to access an accurate and efficient visual information. Human beings have an incredible ability to visually capture relevant targets quickly and accurately, which is called focus of attention or saliency. This visual saliency mechanism makes some objects in a scene stand out from the surroundings. It can assist visually impaired to understand natural scene images. The algorithms and methods of salient objects extraction and text recognition methods in images are considered relevant and widely researched in the developed countries such as USA, Germany, United Kingdom, China, Japan, India, South Korea and Russia.

In the world where number of smart portable devices increasing and becoming universal, there is a growing need for the image and video-based applications such as social support, object recognition, assistive navigation, geocoding, and a number of research studies focused on them. In this area, special attention is paid to the development of methods, algorithms and software tools of salient object extraction, contour detection in natural images and represent them on tactile graphics for visually impaired persons to perceive visual information. At the same time, one of the most important issues is the development of algorithms and software for Uzbek language texts detection, extraction, recognition and pronunciation in natural images.

In the Republic of Uzbekistan, the large scale attention is given for protecting the interests of individuals with disabilities, assisting their rehabilitation and social adaptation by establishing special boarding schools, centers and libraries which are planning to equipped with scientific and innovative technologies that used digital image processing methods. In this matter, essential results are achieved using tactile graphics generation based on salient objects extraction and text recognition from scene images can be used in text to speech technologies. The Strategy for the Further Development of the Republic of Uzbekistan outlines the tasks such as «... increased government support for individuals with disabilities, improving social services, ... providing wide opportunities to education for the visually impaired...»<sup>1</sup>. In order to accomplish these tasks a development of the tactile graphics based on salient objects extraction algorithms and speech synthesizer using text recognition algorithms in images to access visual information for visually impaired people are considered as significant issues.

---

<sup>1</sup> Decree of the President of the Republic of Uzbekistan PD-4947 «On the Strategy for the Further Development of the Republic of Uzbekistan» of 7 February 2017

The dissertation is directly serve the implementation of the tasks set out in the decree of the President of the Republic of Uzbekistan № PD-4947 of 7 February 2017 «On the Strategy for the Further Development of the Republic of Uzbekistan», decree of the President of the Republic of Uzbekistan № PD-5270 of 1 December 2017 «On measures to further improve the system of state support for individuals with disabilities» and decree of Cabinet of Ministers of the Republic of Uzbekistan № 739 of 18 September 2018 «On improving information-library service system for individuals with disabilities» as well as in other regulatory legal acts related to this activity.

**Conformity of the research work with the priority areas of science and technology development of the Republic.** This research was performed in accordance with the priority areas of science and technology IV – «The development of informatization and information communication technologies».

**The level of development of the problem.** Extensive research is being carried out to study development of salient objects detection and extraction, text detection and recognition in natural scene images, as well as computational algorithms and software, and the analysis of the results of computational experiments to identify acceptable parameters of implementation performance. Issues related to the creation of such methods and algorithms have been studied in recent years in articles published in Springer, ACM, IEEE Xplore, Cyberleninka, Web of Science, and Elsevier, mainly in academic journals in the IEEE Xplore and Scopus databases.

This research area is dedicated to the work of scholars such as M.M.Cheng, L.Itti, A.Akhilish, R.Smith, Sh.Li, Y.I.Zhuravlev, L.Neumann, M.B.Medvedev, S.Goferman, A.C.Kuznetsov, X.Zhou, D.Bradley, B.Epshtein, C.Yao, C.Jinsoo, D.E.Suprun, N.Takagi, M.Jaderberg, R.Margolin, P.Felzenszwalb, T.Dutoit and other well-known foreign scholars. In addition, for the development of theoretical foundations of image processing and pattern recognition in the Republic of Uzbekistan M.M. Kamilov, Sh.Kh.Fozilov, N.M.Mirzaev, U.R.Khamdamov, S.S.Radjabov, T.S.Jumaev, A.A.Makhkamov, O.N.Djuraev and other scientists and researchers contributed to the research in this area.

The results of the analysis have shown that, despite the progress in the field of object extraction and text recognition from natural scene images for the visually impaired, detecting and extracting salient objects from natural images when creating tactile graphics as well as text detection and recognition from scene images when developing Uzbek language text-to-speech engine is insufficiently studied.

**Connection of dissertation topic with the plans of research activities of the higher education institution where the dissertation is carried out.** The research was carried out within the framework of the projects on number №BV-F4-011 «Development of theory, methods and tools of parallel computing in signal processing and image processing» (2017-2019), №BV-Atex-2018-251 «Development of Uzbek language Text-To-Speech (TTS) software and speech synthesizer for the visually impaired to use computer technology to read and to

write texts» (2018-2019) of the Tashkent University of Information Technologies named after Muhammad al-Khwarizmi.

**The aim of the research work** is a development of methods, algorithms and software of salient objects extraction, its edge detection and scene text recognition, pronunciation from natural scene images.

**The tasks of the research work:**

analysis of methods of salient objects extraction in images based on global contrast enhancement of the color values with locally adaptive triple thresholding;

development of algorithmic method of Uzbek language texts detection, extraction and recognition in natural images based on fully convolutional neural network model;

development of algorithm and software of graphic objects extraction and inner and outer edge detection in images to represent on tactile graphics;

development of algorithm and software of Uzbek language texts pronunciation in natural images for perception process of visually impaired.

**The object of the research work** are natural scene images, video information, words and audio information for Uzbek language text-to-speech synthesis as well as image processing tools.

**The subject of the research work** – digital image processing algorithms, neural network models, and programming libraries to solve the problem of object detection effectively, extraction and text recognition in images.

**The methods of the research work.** In order to solve this problem modern methods of digital image processing, pattern recognition, character recognition, deep learning, probability theory, mathematical analysis, algorithmization, structured programming technology and performing computational experiments are used.

**The scientific novelty of the research work** as follows:

method of objects extraction in images based on global contrast enhancement of the color values with locally adaptive triple thresholding is improved;

algorithmic method of Uzbek language texts detection, extraction and recognition in natural images based on fully convolutional neural network model is developed;

algorithm and software of graphic objects extraction, inner and outer edge detection in images to represent on tactile graphics are developed;

algorithm and software of Uzbek language texts pronunciation in natural images for perception process of visually impaired are developed.

**The practical results of the research work** are as follows:

the software on detecting and extracting salient objects in natural scene images to generate tactile graphics for the visually impaired is developed;

the software on detecting and recognizing text in natural scene images to generate Uzbek language TTS synthesizer for the visually impaired is developed;

the proposed salient objects detection and extraction method have been implemented in the specialized boarding school for visually impaired № 77, Center

for the development of ICT in public education and used in tactile graphics generation process to give visual information for the visually impaired;

the proposed scene text detection and recognition method have been implemented in the Central Board of the Uzbek Association of the Blind, the Center of youth and children with disabilities and applied in Uzbek language TTS synthesizer creation process for the visually impaired;

**The reliability of the research results.** The reliability of the research results based on the fact that both quantitative and qualitative comparisons of the proposed methods and algorithms with other state-of-the-art approaches performed, achieved accurate and efficient results as well as confirmed by appropriate certificates and documents of implementation. Moreover, proposed salient object detection, salient object extraction, text detection and text recognition methods also evaluated with experimental calculations.

**The scientific and practical significance of the research results.** The scientific significance of the research results are salient objects detection method based on global contrast enhancement using histogram equalization, salient objects extraction method based on locally adaptive triple thresholding using integral images, text detection method based on fully convolutional neural network and Uzbek language text recognition method using trained Tesseract OCR with Uzbek Latin and Cyrillic alphabets.

The practical significance of the research results are software of salient objects detection and extraction from natural scene images can be applied in tactile graphics generation for the visually impaired. Moreover, the software of text detection and recognition from natural scene images can be used in Uzbek language TTS synthesizer for the visually impaired.

**Implementation of the research results.** Based on the results obtained by methods, algorithms and software of salient objects extraction, text recognition and pronunciation for the visually impaired:

improved method, algorithm and developed software of objects extraction in images based on global contrast enhancement of the color values with locally adaptive triple thresholding have been implemented to the production and education processes in the training laboratory of specialized boarding school for visually impaired № 77 under the Ministry of Public Education of the Republic of Uzbekistan (Certificate of the Ministry for Development of Information Technologies and Communications № 33-8/8552 of December 3, 2019). As a result, visually impaired students were able to clearly identify 74% of the tactile graphics that were presented to them;

developed method, algorithm and software of Uzbek language texts detection, extraction and recognition in natural images based on fully convolutional neural network model have been implemented to the information services processes of the Central Board of the Uzbek Association of the Blind (Certificate of the Ministry for Development of Information Technologies and Communications № 33-8/8552 of December 3, 2019). As a result, Uzbek language texts in electronic files



pronounced on computer system, the visually impaired readers were able to obtain a larger amount of information in relatively short period of time than braille text;

developed methods, algorithms and software of salient objects extraction, text recognition and pronunciation in images have been implemented to the production, educational and social services processes of the Center of youth and children with disabilities under Union of the Youth of Uzbekistan and Center for the development of information and communication technologies in public education (Certificate of the Ministry for Development of Information Technologies and Communications № 33-8/8552 of December 3, 2019). As a result, the development of the professional interests and skills, the enrichment of imagination and the expansion of opportunities in society using educational, handout, and demonstration materials in tactile graphics at training courses and circles for developing the professional skills of visually impaired students were achieved.

**The approbation of the research results.** The results of the research were discussed at 5 international and 2 national scientific conferences and scientific seminars.

**The publication of the research results.** The total number of 23 scientific works on the subject of the research, including 8 journal papers recommended in scientific publication of the Higher Attestation Commission of the Republic of Uzbekistan, 4 in republican, 4 in international journals and 3 scientific articles in other journals, 5 in international and 2 national scientific conferences are published and 5 certificates of registration of the software created for the computer are received.

**The structure and volume of the dissertation.** The dissertation consists of an introduction, four chapters, conclusion, references, abbreviations and appendixes. The dissertation consists of 117 pages.

## MAIN CONTENT OF DISSERTATION

**The introduction** provides the actuality and demanding of the dissertation topic, a description of the objectives and main tasks, as well as objects and subjects corresponding to the priority areas of science and technology of the Republic of Uzbekistan, scientific novelties and practical results, the reliability and practical significance of results, information on the implementation of research results in practice, as well as information about the publication of the work and dissertation structure.

In the first chapter of the dissertation, «**Analysis of methods and algorithms for objects extraction and recognition in images**», explained the interpretation and analysis of existing methods as well as the relevance of the problem. It is emphasized that one of the most challenging aspects of the computer vision is automatic object extraction and recognition methods. As a result of analysis of existing methods, it is concluded that color contrast is one of the most important factors for effective detection of salient objects and texts in images. It has also been established that the calculation of multi-level threshold values in the method

of extracting salient objects from complex scene images allows more robust results than a single threshold value.

Furthermore, text detection and recognition in natural scene images is an open and difficult problem of computer vision since text typically settle just a short part of the image, it has a non-uniform environment, it suffers from reflections, occlusions, blur, and noise and perspective results need to be taken into consideration. In recognizing texts in images, firstly it is important to determine the location of texts for reliable result. It has been suggested that the use of different neuron network models is the most appropriate way to solve the text detection and localization problems.

In this chapter of the dissertation, it has been viewed that visual information generation based on tactile graphics is one of the most ambitious research areas since these techniques require important information from an input image. Visual information generally includes pictures, figures, graphs, diagrams, and colors and so on. It is important to remember that visually impaired people access these information mainly through the following two ways: (1) touching, that is, graphs, pictures, diagrams, and figures have to be presented as tactile graphics; (2) hearing, so the visually impaired need to be told the whole environment.

In addition, it has been analyzed that tactile representation of important objects in a natural scene image should be designed as simple and easily understandable for the visually impaired, and a wide range of studies are being conducted to find optimal methods for extracting important objects from natural scene images. Moreover, the stages, methods and process of developing Text to Speech (TTS) synthesizer were analyzed.

The research aimed at achieving the goals and objectives of the dissertation was based on research methodology developed within the dissertation. Research methodology includes: analysis of digital image processing methods such as salient object detection, extraction and edge detection as well as text detection and text recognition; development of methods and algorithms for solving problems; research and selection of relevant computer vision tools and libraries; develop software package for object extraction and text recognition as well as tactile graphics generation and Uzbek language TTS synthesizer.

In the second chapter of the dissertation, «**Creation of methods and algorithms of salient objects extraction in images for visually impaired**», salient object detection methods based on global contrast enhancement and salient object extraction method based on locally adaptive triple thresholding were developed. The proposed method consists of the following steps as shown in Figure 1: 1) global contrast enhancement method using histogram equalization, 2) saliency map generation method 3) salient region extraction method based on locally adaptive triple thresholding, and 4) salient object's boundary and inner edge generation method.

Global contrast enhancement based on histogram equalization of color values of the image is applied into pre-processing step to improve salient region contrast than surrounding regions. A high contrast against the surrounding region is usually

a stronger evidence for saliency of a region than a comparable contrast against a far-away region. Naively evaluating the saliency of a pixel is defined using its color contrast to all other pixels in the image and takes  $O(N^2)$  time, which is computationally too expensive even for medium sized images. To reduce computational time, contrast enhanced image segmented into regions using graph-based images segmentation method and regions' saliency value can be obtained by measuring its color contrast to all other regions in the image:

$$S(r_k) = \sum_{r_k \neq r_i} w(r_i) D(r_k, r_i) \quad (1)$$

where  $w(r_i)$  is the weight of region  $r_i$  and  $D(r_k, r_i)$  is the color distance metric between two regions. Spatial information is incorporated by using a spatial weighting term to increase the effects of closer regions and to decrease the effects of farther regions. In the spatial weighting formulation,  $D_s(r_k, r_i)$  is the spatial distance between regions  $r_k$  and  $r_i$ ,  $\sigma_s$  control the strength of spatial distance weighting,  $w(r_i)$  is the weight of region  $r_i$  defined by the number of pixels in  $r_i$ ,  $w_s(r_k) = \exp(-9d_k^2)$  is a spatial prior weighting term, and  $d_k$  is the average distance between pixels in region  $r_k$  and the center of the image and  $\sigma_s^2 = 0.4$ .

$$S(r_k) = w_s(r_k) \sum_{r_k \neq r_i} e^{-\frac{D_s(r_k, r_i)}{\sigma_s^2}} w(r_i) D(r_k, r_i) \quad (2)$$

In addition, salient object extraction method based on locally adaptive triple thresholding using integral images is proposed. The proposed salient object detection and extraction methods consist of pre-processing, saliency map, saliency cuts, outer and inner edges as shown in Figure 1.

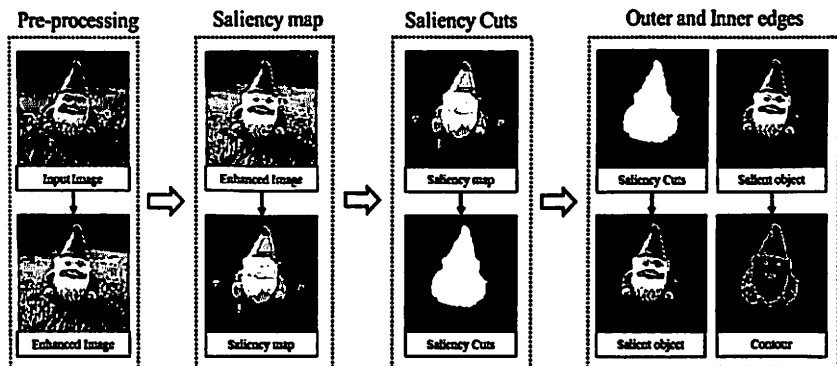


Figure 1. The proposed salient objects detection and extraction methods

It calculates an integral image in the first move through the saliency map grayscale image and then second move figure out the local window ( $S \times S$ ) using an integral image for every pixel in constant time further obtain local adaptive three level threshold values by performing the comparison of further obtaining local

adaptive three level threshold values. These three thresholds marked the saliency map with four different regions, which demonstrates four types of seeds that are assigned an iterative variant of the GrabCuts method, allude to as Saliency cuts and a full resolution, a high quality binary mask is produced.

The purpose of proposing a locally adaptive triple thresholding method accurately extracts salient objects that have small saliency values. However, due to the fact that local thresholding methods are slower than global thresholding methods, it has been able to increase the processing speed by computing integral images. The integral image (also known as a summed-area table) is a method that can be used as a fast and effective way that speed up for computing the sum of values (pixel intensity values) in following formulation:

$$I(x, y) = i(x, y) + I(x - 1, y) + I(x, y - 1) - I(x - 1, y - 1) \quad (3)$$

Moreover, the contour of the salient object can be obtained using the binary mask and applying Canny edge detection method. These outer boundaries can translate into tactile graphics to deliver given natural scene image visual information to the visually impaired. In some cases, however, the visually impaired may not be confident about the object information just by feeling the boundary of a salient object. Therefore, it is very important that internal edges also detected for visually impaired to fully understand the salient object in a natural scene. For obtaining internal edges of salient object with a help of our saliency cuts' results, which is binary mask, create a matrix with the same size and type as an input image to achieve a desired output image. After that, just copy the non-zero elements of binary mask that indicate the original input image matrix elements:

$$S_o = B_m(x, y) * T_l(x, y) \quad (4)$$

where  $S_o$  is the salient object,  $B_m(x, y)$  is the binary mask, and  $T_l(x, y)$  is the input image.

Based on the results of the research, salient objects extraction and its contour detection in images to create a simple and easy-to-understand tactile graphics has been proven to be one of the most effective methods.

In the third chapter of the dissertation, «**Development of text extraction and recognition method and algorithm based on neural network**», text detection and recognition methods based on fully convolutional neural network (FCN) and Tesseract OCR model were developed. The proposed text detection and recognition methods divided into four main stages as shown in Figure 2: 1) pre-processing stage global contrast enhancement using histogram equalization method 2) text detection using FCN, 3) text extraction and recognition, and 4) Uzbek language TTS synthesizer. The end-to-end text recognition approaches consist of two parts: text detection and text recognition. It should be noted that, existing approaches, both conventional and deep neural network-based, principally consist of many steps and parts, which are reasonably sub-optimal and time-consuming.

Accordingly, the efficiency and accuracy of such approaches are still far from sufficient.

To overcome these drawbacks, a quick and reliable scene text detection and localization method that has only two steps proposed. The proposed method uses FCN model that instantly generates word or text-line level prophecies, apart from unnecessary and heavy intermediate stages. Using convolutional neural network (CNN) is not an effective solution to determine the precise location of texts in the image. Therefore, the use of the FCN model will yield higher accuracy results.

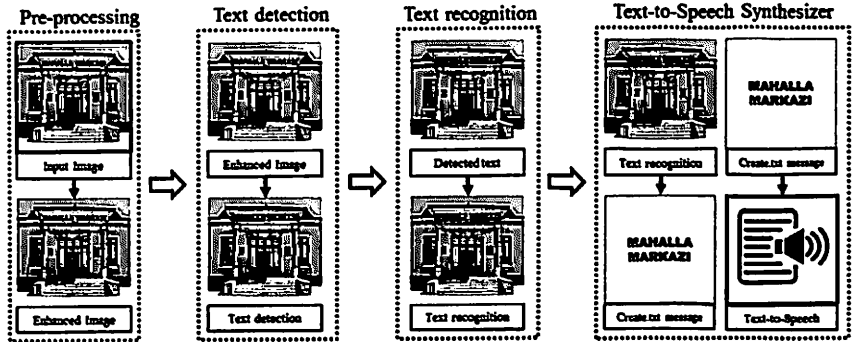


Figure 2. The proposed text detection and recognition methods and Uzbek language TTS synthesizer

Furthermore, different circumstances must be considered when creating neural networks for text detection. Because the areas of text regions differ remarkably, discovering the presence of long sentences would need features from late-stage of a neural network, while predicting correct geometry surrounding a short text regions demand low level knowledge in early stages. For that reason the network must utilize features from various levels to satisfy these demands. The proposed method slowly unites feature maps while preserving the up sampling features merging small. Simultaneously the method concludes with a network that can both use various levels of features and retain a small calculation cost. The model can be decayed into three parts: feature extractor, feature-merging and output layer. The feature extractor might be a convolutional network pre-trained on ImageNet dataset, along with interleaving convolution and pooling layers. Four levels of feature maps, represented as  $f_i$ , are obtained from the feature extractor, whose sizes are  $1/32$ ,  $1/16$ ,  $1/8$  and  $1/4$  of the input image, respectively. Mathematically, feature-merging formulation expressed as:

$$g_i = \begin{cases} \text{unpool}(h_i) & \text{if } i \leq 3 \\ \text{conv}_{3 \times 3}(h_i) & \text{if } i = 4 \end{cases} \quad (5)$$

$$h_i = \begin{cases} f_i & \text{if } i = 1 \\ \text{conv}_{3 \times 3}(\text{conv}_{1 \times 1}([g_{i-1}; f_i])) & \text{otherwise} \end{cases} \quad (6)$$

where  $g_t$  is the merge base, and  $h_t$  is the merged feature map, and the operator  $[:]$  denotes concatenation with the channel axis.

In the next stages, once the text region is detected, the region can be cropped and processed further to recognize the text. To do this, trained Tesseract OCR model with Uzbek Latin and Cyrillic alphabet characters can be used. The proposed method also includes recognized texts send to TTS synthesizer for Uzbek language.

In this chapter of the dissertation, the result of studying and analyzing words in the Uzbek dictionary, an electronic database of 31,5 thousand words was formed and arranged in alphabetical order. The Uzbek language speech synthesizer is based on the concatenation method and contains pronunciation of the words. Therefore, the Uzbek vocabulary with 31,5 thousand words were studied and all words were broken down into 2,5 thousand sections, i.e. syllables. For correct pronouncing of recognized texts and update Uzbek language database, recognized texts are compared with database, if recognized text is exist in Uzbek language database system send it to Uzbek language TTS Synthesizer, else the word send to language specialist to confirm new word.

In the fourth chapter of the dissertation, «**Implementation and experimental results of objects extraction and recognition methods in images**» presents object detection, extraction algorithms, text detection, recognition algorithms and its software as well as the results of comparison experiments and practical implementation. The proposed salient objects extraction and text recognition methods are implemented in C++, programming language using OpenCV library.

The all proposed methods are compared with other state-of-the-art alternate methods using most popular available datasets: salient objects detection method compared with other 12 alternate methods using MSRA 10k dataset; salient object extraction method compared with other 2 alternate methods using MSRA 10k dataset; text detection method compared with other 17 alternate methods using ICDAR 2015 and MSRA-TD500 datasets; text recognition method compared with other 7 alternate methods using ICDAR 2013 dataset. In addition, performed both quantitative and qualitative comparisons of the proposed methods with other methods.

In qualitative comparison of salient objects detection and extraction, it has been determined the proposed methods effectively suppresses background regions and uniformly emphasizes foreground regions as well as extract salient objects with less amount of data about the saliency map without misclassifying them as background regions, and even detect multiple salient objects. Furthermore, text region effectively and accurately detected as well as Uzbek text correctly recognized in qualitative comparison of text detection and recognition.

In this chapter of the dissertation, in quantitative comparison, the standard evaluation metrics are used such as Precision (P), Recall (R), F-Measure (FM), Receiver Operating Characteristics (ROC), Area under ROC Curve (AUC) and Mean Absolute Error (MEA). Precision is the fraction of the detected salient pixel belonging to the salient objects in the ground truth, and recall corresponds to the

percentage of salient pixels correctly assigned. The PR curve is obtained by normalizing the saliency map to [0, 255], generating binary masks with a threshold varying from 0 to 255, and comparing the binary mask against the manually labeled ground truth. Precision and recall rates can be obtained as follows:

$$precision = \frac{TDO \cap GT}{ADO} \quad (7)$$

$$recall = \frac{TDO \cap GT}{GT} \quad (8)$$

where TDO denotes truly detected salient objects, ADO all detected objects, and GT manually labeled ground truth. F-Measure value which balanced measurements between mean of precision and recall rates. A higher F-Measure means a higher performance and it is defined as follows:

$$F_{\beta} = \frac{(1+\beta^2)precision \times recall}{\beta^2 \times precision + recall} \quad (9)$$

where  $\beta^2 = 0.3$ .

The proposed text detection method obtain the most robust results in both ICDAR 2015 and MSRA-TD500 datasets with 84.7% per cent and 84% per cent F-Measure respectively. Moreover, the proposed text recognition method achieves 94.6% per cent recognition rate by leaving behind other methods.

Developed salient objects extraction and its edge detection algorithms for tactile graphics and displays as well as the Uzbek language texts detection, recognition and pronunciation algorithms in images for speech synthesizer can be used in other information technologies sphere.

## CONCLUSION

The following conclusion were obtained during the research of doctoral dissertation on the subject of «Development of methods and algorithms for objects extraction in images»:

1. The proposed salient objects detection method based on global contrast enhancement using histogram equalization has compared with 12 alternate methods and obtained high performance in standard measurement metrics.

2. The proposed salient objects extraction method based on locally adaptive triple thresholding using integral images has compared with 2 alternate methods and obtained robust, full resolution binary mask in precision 0.94 even when background and foreground regions have very similar information.

3. Experimental results shows that, other methods can extract only a single salient objects from a given image and may fail to segment salient objects when pixels in the image have low saliency values. However, the proposed method can reduce these drawbacks and extract salient objects with less amount of data about the saliency map without misclassifying them as background regions, and even detect multiple salient objects.

4. The proposed text detection method and algorithms based on fully convolutional neural network has compared with 17 alternate methods and obtained high accuracy of text detection of 84.7 percent F-measure in ICDAR 2015 dataset as well as 84 percent F-measure in MSRA-TD500 dataset.

5. The proposed text recognition method and algorithms based on trained Tesseract OCR model with Uzbek language alphabets has compared with 7 alternate methods and achieved the highest 94.6 percent recognition rate.

6. An algorithm for visualization of tactile graphics using tactile printer together with salient objects outer and inner edges in images for visually impaired has been developed. As a result, high-resolution tactile graphics of the edges of salient objects in images were achieved.

7. An algorithm of text recognition in natural images and Uzbek language speech synthesizer has been developed for the visually impaired. As a result, Uzbek language texts in natural images and electronic files were able to pronounce using computer systems.

8. The developed methods, algorithms and software of salient objects extraction and text recognition methods in images have been implemented to the production, social service and education processes of the "Specialized boarding school for visually impaired № 77", "Central Board of the Uzbek Association of the Blind", "Center of youth and children with disabilities" as well as "Center for the development of information and communication technologies in public education". As a result, the needs, professional interests, abilities and opportunities of visually impaired youth and students have been enhanced by visualizing the contour of salient objects on tactile graphics and pronouncing Uzbek language text using speech synthesizer.



**НАУЧНЫЙ СОВЕТ DSc.13/30.12.2019.Т.07.02 ПО ПРИСУЖДЕНИЮ  
УЧЕНЫХ СТЕПЕНЕЙ ПРИ ТАШКЕНТСКОМ УНИВЕРСИТЕТЕ  
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ**

---

**ТАШКЕНТСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ  
ТЕХНОЛОГИЙ**

**МУХИДДИНОВ МУХРИДДИН НУРИДДИН УГЛИ**

**РАЗРАБОТКА МЕТОДОВ И АЛГОРИТМОВ ИЗВЛЕЧЕНИЯ  
ОБЪЕКТОВ ИЗ ИЗОБРАЖЕНИЙ**

05.01.01 – Инженерная геометрия и компьютерная графика.  
Аудио и видеотехнологии

**АВТОРЕФЕРАТ ДИССЕРТАЦИИ  
ДОКТОРА ФИЛОСОФИИ (PhD) ПО ТЕХНИЧЕСКИМ НАУКАМ**

Ташкент – 2020



## **ВВЕДЕНИЕ (аннотация диссертации доктора философии (PhD))**

Целью исследования является разработка методов, алгоритмов и программных средств извлечения значимых объектов и выделение контуров, а также распознавание и произношение текста из естественных изображений для понимания визуальной информации людьми с нарушениями зрения.

Объектом исследования являются естественные изображения, видеоинформация, слова и аудиоинформация для синтеза речи на узбекском языке, тактильные дисплеи, а также инструменты обработки изображений.

Научная новизна исследования заключается в следующем:

усовершенствован метод извлечения значимых объектов из изображений на основе операций улучшения глобального контраста и фильтрации локально-адаптивным тройным порогом цветовых значений изображений;

разработан алгоритмический метод процессов определения, извлечения и распознавания текстов на узбекском языке в естественных изображениях на основе модели полной свёрточной нейронной сети;

разработаны алгоритмы и программные средства выделения графических объектов в изображениях, определение внутренних и внешних контуров, а также представление их в тактильной графике;

разработаны алгоритмы и программные средства определения и звукового воспроизведения текстов на узбекском языке из изображений для восприятия незрячими людьми.

**Внедрение результатов исследования.** На основе полученных результатов по разработке методов, алгоритмов и программных средств извлечения значимых объектов, распознавания и произношения текста для слабовидящих:

внедрены методы, алгоритмы и программные средства улучшения глобального контраста изображения на основе выравнивания гистограммы значений цвета пикселей и извлечения значимых объектов из изображений на основе локально-адаптивного тройного порога в производственные и учебные процессы учебной лаборатории специализированной школы интерната для слабовидящих № 77 при Министерстве народного образования Республики Узбекистан (Справка Министерства по развитию информационных технологий и коммуникаций №33-8/8552 от 3 декабря 2019 года). В результате ученики с нарушениями зрения смогли правильно идентифицировать 74% тактильной графики, которая была им представлена во время экспериментов;

внедрены методы, алгоритмы и программные средства определения и распознавания текста в изображениях на основе модели полной свёрточной нейронной сети в процессы информационных услуг Центрального правления общества слепых Узбекистана (Справка Министерства по развитию информационных технологий и коммуникаций №33-8/8552 от 3 декабря 2019 года). В результате голосового произношения электронных текстов на узбекском языке в компьютерной системе были достигнуты более высокие

результаты при получении большего объема информации слабовидящими читателями за относительно короткий срок, чем тексты Брайля;

внедрены методы, алгоритмы и программные средства извлечения значимых объектов, распознавания и произношения текстов на изображениях в производственные и учебные процессы для предоставления социальных услуг Центра молодежи и детей с ограниченными возможностями при Союзе молодежи Узбекистана и Центра развития информационно-коммуникационных технологий в сфере народного образования при Министерстве народного образования (Справка Министерства по развитию информационных технологий и коммуникаций №33-8/8552 от 3 декабря 2019 года). В результате было достигнуто: развитие профессиональных интересов, обогащение воображения и расширение возможностей в обществе с использованием учебных, раздаточных, демонстрационных материалов в тактильной графике на учебных курсах и кружках по развитию профессиональных навыков слабовидящих учащихся.

**Структура и объем диссертации.** Диссертация содержит 117 страниц и состоит из введения, четырех глав, заключения, списка использованной литературы, списка условных обозначений и терминов, приложения.

**ЭЪЛОН ҚИЛИНГАН ИШЛАР РЎЙҲАТИ**  
**LIST OF PUBLISHED WORKS**  
**СПИСОК ОПУБЛИКОВАННЫХ РАБОТ**

**I бўлим (I part; I часть)**

1. Hongchan Yoon, Baek-Hyun Kim, Mukhriddin Mukhiddinov, Jinsoo Cho. Salient Region Extraction based on Global Contrast Enhancement and Saliency Cut for Image Information Recognition of the Visually Impaired. // KSII Transactions on internet and information systems. Vol. 12, No. 5. South Korea, May, 2018. -p. 2287-2312 (№3) Scopus, Cite Score=0.85

2. Khamdamov U.R., Mukhiddinov M.N., Djuraev O.N., Abdusalomov A.B. Image segmentation based on global contrast for salient object extraction. // Descendants of Muhammad al-Khwarizmi scientific-practical and information-analytical journal. Tashkent, -2018, 1(3), -p. 63-66 (05.00.00; №10)

3. Abdusalomov A.B., Mukhiddinov M.N. Image enhancement based on histogram equalization for Indoor environment objects. // Descendants of Muhammad al-Khwarizmi scientific-practical and information-analytical journal. 2(4). Tashkent, -2018. -p. 10-13 (05.00.00; №10)

4. Khamdamov U.R., Mukhiddinov M.N., Mukhamedaminov A.O., Djuraev O.N. A novel method for extracting text from natural scene images and TTS. // "European science review" scientific journal, Premier publishing. 2018. № 11-12 Vol. 1. Vienna, -2018. -p. 30-33 (05.00.00; №3)

5. Khamdamov U., Mukhiddinov M., Djuraev O. An Overview of Deep Learning Based Text Spotting in Natural Scene Images. // Problems of Computational and Applied Mathematics. 2019, № 2(20). Tashkent, -2019. -p. 126-134 (05.00.00; №23)

6. Mukhiddinov M.N. Scene Text Detection and Localization using Fully Convolutional Network. // International Conference of Information science and communications technologies applications, trends and opportunities ICISCT 2019. 4-6 November, 2019. Tashkent, Uzbekistan. (30.09.2019 №269/8 – сон раёсат қарори)

7. Mukhiddinov M.N. Akmuradov B.U., Djuraev O.N. Robust Text Recognition for Uzbek Language in Natural Scene Images. // International Conference of Information science and communications technologies applications, trends and opportunities ICISCT 2019. 4-6 November, 2019. Tashkent, Uzbekistan. (30.09.2019 №269/8 – сон раёсат қарори)

8. Mukhiddinov M.N. Tactile graphics generation from natural scene images for the visually impaired. // Descendants of Muhammad al-Khwarizmi scientific-practical and information-analytical journal. 4(10). Tashkent, -2019. -p. 88-91 (05.00.00; №10)

## II бўлим (II part; II часть)

9. Mukhiddinov Mukhriddin, Jin-Soo Cho. Salient region extraction based on global contrast enhancement and Saliency Cut. // 2nd International Conf. on Electronics, Electrical Engineering, Computer Science: Innovation and Convergence (EEEECS). August 10-13, 2016. Qingdao, China. –p. 91-92

10. Khamdamov U.R., Mukhiddinov M.N., Djuraev O.N., Abdusalomov A.B. Image segmentation based on global contrast for salient object Extraction. // International conference on importance of Information-communication technologies in Innovative development of sectors of economy. Tashkent, – 2018. - p. 573-576

11. Хамдамов У.Р., Мухиддинов М.Н. Алгоритмы сжатия растровых изображений на основе вейвлет функций. // International conference on importance of Information-communication technologies in Innovative development of sectors of economy. Tashkent, – 2018. -p. 654-657

12. Usmanov R.N., Abdusalomov A.B., Kuchkorov T.A., Mukhiddinov M.N. Image enhancement based on histogram equalization for Indoor environment objects. // International conference on importance of Information-communication technologies in Innovative development of sectors of economy. Tashkent, – 2018. - p. 777-780

13. Mukhiddinov M.N., Djuraev O.N. An Improved Text Detection Method for Natural Scene Images. // “Высшая школа”. Научно-практический журнал. №10/ 2018. г.Уфа. Россия. –с. 93-95

14. Khamdamov U., Mukhiddinov M., Zarmasov E. An Overview of and Speech Synthesis Algorithms. // “Высшая школа”. Научно-практический журнал. №18/ 2018. г.Уфа. Россия. –с. 24-25

15. Akmuradov B., Mukhiddinov M., Samatkhonov M. Text to speech conversion using different speech synthesis technologies. // “Высшая школа”. Научно-практический журнал. №18/ 2018. г.Уфа. Россия. –с. 54-56

16. Khamdamov U., Mukhiddinov M. Developing intonation pattern for Uzbek text to speech synthesis system. // Материалы XXIII Международной научно-технической конференции «Современные средства связи». 18-19 октябрь 2018. Минск, Беларусь. –с. 131-133

17. Khamdamov U., Mukhiddinov M., Djuraev O. Text Extraction in Natural Scenes Using Region-Based Method. // Ахборот коммуникация технологиялари ва дастурий таъминот яратишда инновацион ғоялар”, Республика илмий-техник анжуманининг маърузалар тўплами. 2-қисм, 16-17 апрель, 2019. Самарқанд. –б. 52-54

18. Akmuradov B., Mukhiddinov M., Djuraev O. Text Speech Synthesis Using Neural Networks. Ахборот коммуникация технологиялари ва дастурий таъминот яратишда инновацион ғоялар”, Республика илмий-техник анжуманининг маърузалар тўплами. 2-қисм, 16-17 апрель, 2019. Самарқанд. – б. 55-57

19. Хамдамов У.Р., Зармасов Э.М., Мухиддинов М.Н. ва бошқалар. «eYespeakDictionary» ўзбек тилининг электрон изоҳли луғати. // Ўзбекистон Республикаси Интеллектуал мулк агентлиги. Гувоҳнома № DGU 05351, 31.05.2018.

20. Хамдамов У.Р., Зармасов Э.М., Мухиддинов М.Н. ва бошқалар. «eYespeak Synthesizer» - аудио сигналларга ишлов бериш дастури. // Ўзбекистон Республикаси Интеллектуал мулк агентлиги. Гувоҳнома № DGU 06346, 13.05.2019.

21. Хамдамов У.Р., Мухиддинов М.Н., Джураев О.Н. Тасвирлардаги муҳим объектларни аниқлаш ва ажратиб олиш дастурий таъминоти. // Ўзбекистон Республикаси Интеллектуал мулк агентлиги. Гувоҳнома № DGU 06808, 01.08.2019.

22. Хамдамов У.Р., Акмурадов Б.У., Мухиддинов М.Н. График процессорларда Ҳаар вейвлет иёрадаи товуш сигналларига паралел ишлов бериш дастури. // Ўзбекистон Республикаси Интеллектуал мулк агентлиги. Гувоҳнома № DGU 06939, 24.09.2019.

23. Хамдамов У.Р., Мухиддинов М.Н., Джураев О.Н. Табиий тасвирлардаги матнларни аниқлаш ва таниш дастурий таъминоти. // Ўзбекистон Республикаси Интеллектуал мулк агентлиги. Гувоҳнома № DGU 07115, 07.11.2019.

**Автореферат «Муҳаммад ал-Хоразмий авлодлари» илмий журнали таҳририятида таҳрирдан ўтказилди ва ўзбек, инглиз, рус тилларидаги матнларнинг мослиги текширилди.**

**Бичими 60x84<sup>1</sup>/<sub>16</sub>. Ракамли босма усули. Times гарнитураси.  
Шартли босма табоғи: 2,75. Адади 100. Буюртма № 81.**

**Гувоҳнома reestr № 10-3719  
“Тошкент кимё технология институти” босмаҳонасида чоп этилган.  
Босмаҳона манзили: 100011, Тошкент ш., Навоий кўчаси, 32-уй.**