KOLONOSKOPIYA TASVIRLARINI TAHLIL QILISHDA SUN'IY INTELLEKT YONDASHUVLARI: YARALI KOLIT DIAGNOSTIKASI MISOLIDA QIYOSIY TAHLIL

Nurmetova Bonuraxon Baxromovna[1], Odamboy Djumanazarov Ro'zimboyevich[2]

[1]Abu Rayhon Beruniy nomidagi Urganch davlat universiteti Kompyuter injiniring kafedrasi, https://orcid.org/0000-0003-0444-0612, bonuraxon20102018@gmail.com
[2]Abu Rayhon Beruniy nomidagi Urganch davlat universiteti Kompyuter injiniring kafedrasi, https://orcid.org/0000-0002-5284-207X, odamboy1984@gmail.com

K E Y W O R D S

tibbiy tasvirlar, sun'iy intellekt, CNN, segmentatsiya, klassifikatsiya, kolonoskopiya, yarali kolit, Vision Transformer

A B S T R A C T

Mazkur maqolada tibbiy tasvirlarni, xususan kolonoskopiya tasvirlarini tahlil qilishda sun'iy intellekt(SI) metodlarining rivojlanish bosqichlari, zamonaviy yondashuvlari va mavjud muammolari tahlil qilinadi. An'anaviy raqamli tasvirni qayta ishlash usullaridan boshlab, konvolyutsion neyron tarmoqlari(CNN), segmentatsiya arxitekturalari hamda Vision Transformer va gibrid modellarga qadar bo'lgan evolyutsiya tahlil qilindi. Yarali kolit(ulcerative colitis) diagnostikasi misolida SI ning klinik qarorlarni qoʻllab-quvvatlashdagi ro'li baholanadi. Eng samarali modellarning qiyosiy tahlillari keltirilgan bo'lib, ma'lumotlar yetishmasligi, qayta oʻrganish(overfitting) va interpretatsiya muammolari kabi cheklovlar muhokama qilingan.

Kirish

Zamonaviy tibbiyotda tasvir ma'lumotlari, rentgen, kompyuter tomografiyasi(KT), magnit-rezonans tomografiya(MRT), ultratovush hamda endoskopik va kolonoskopik tasvirlar yordamida tashxis qo'yish va davolash strategiyasini belgilash hal qiluvchi muhim rolni o'ynaydi. Tibbiy tasvirlarni qoʻlda tahlil qilish shifokor malakasi va tajribasiga bog'liq bo'lib, kuzatuvchilar o'rtasidagi(interobserver) va bir kuzatuvchining o'zidagi(intraobserver) farqlanish, tashxis ishonchliligini cheklaydi. Soʻnggi oʻn yillikda chuqur oʻrganish(deep learning) usullari yordamida tibbiy tasvirlarni tahlil qilish ustuvor metodologiyaga aylandi va klassifikatsiya, ob'ektni aniqlash hamda segmentatsiya vazifalarida yuqori aniqlikni ta'minladi[1].

Kolonoskopiya — yoʻg'on ichak kasalliklari, jumladan kolorektal saraton va surunkali yallig'lanish kasalliklarini aniqlashda "oltin standart" hisoblanadi. Yarali kolit(ulcerative colitis, UC) surunkali yallig'lanishli ichak kasalligi bo'lib, uning faolligini baholashda Mayo endoskopik ballari(Mayo Endoscopic Score) va UCEIS(Ulcerative Colitis Endoscopic Index of Severity) shkalalaridan foydalaniladi. Biroq bu baholash subʼektiv tabiatga ega va mutaxassislar o'rtasida sezilarli farqlanishlarga olib keladi[15]. Aynan shu subʼektivlikni kamaytirish uchun SI ga asoslangan obʼektiv va takrorlanuvchan baholash tizimlari ishlab chiqilmoqda[16].

SI texnologiyalarining hisoblash quvvati o'sishi, yirik annotatsiyalangan ma'lumotlar to'plamlari paydo bo'lishi va arxitekturaviy yangiliklar hisobiga jadal rivojlanishi, tibbiy tasvirlar tahlilida yangi imkoniyatlar ochdi. Tibbiy tasvirlarni, ayniqsa kolonoskopiya tasvirlarini qayta ishlashda qo'llaniladigan SI usullarini tizimli tarzda tahlil qilish, ularning afzallik va kamchiliklarini qiyoslash hamda kelajak rivojlanish istiqbollarini belgilash lozim.

Asosiy qism

1. An'anaviy usullar tahlili.

Chuqur oʻrganish keng tarqalishidan oldin tibbiy tasvirlarni qayta ishlash klassik raqamli tasvir qayta ishlash(image processing) usullariga tayangan. Bularga chegaralash(thresholding),

(cc) BY

chegaralarni aniqlash(edge detection, masalan Sobel, Canny operatorlari), histogrammani tekislash, morfologik operatsiyalar hamda hududni o'stirish(region growing) kabi segmentatsiya usullari kiradi. Ushbu yondashuvlarda belgilar(features) inson tomonidan qoʻlda loyihalanadi(handcrafted features) va klassifikatsiya uchun an'anaviy mashinaviy oʻrganish algoritmlari(SVM, k-NN, Random Forest) qo'llaniladi.

Biroq, mazkur an'anaviy yondashuvlar yoritilish sharoitining dinamik o'zgarishi, kolonoskopik tasvirlardagi spekulyar refleksiyalar(oq yorug'lik aks etishi), mukozadagi biologik qoldiqlar, endoskopik kameraning kinematikasi hamda poliplarning yuqori morfologik xilma-xilligi kabi omillar ta'sirida yetarlicha robustlik(tahliliy barqarorlik) namoyon eta olmaydi. Qoʻlda loyihalangan belgilar tasvirlarning murakkab va o'zgaruvchan tabiatini to'liq aks ettira olmaydi, bu esa umumlashtirish qobiliyatini pasaytiradi[12],[14]. Natijada bunday tizimlar real klinik sharoitda yetarlicha sezgirlik va aniqlikni ta'minlay olmaydi.

2. Deep Learning asosidagi yondashuvlar

Chuqur oʻrganishning asosiy ustunligi belgilarni avtomatik tarzda ma'lumotlardan oʻrganishida namoyon bo'ladi. 2012-yilda AlexNet arxitekturasining ImageNet musobaqasidagi gʻalabasi konvolyutsion neyron tarmoqlari(Convolutional Neural Networks, CNN) davrini boshlab berdi[2]. Keyinchalik VGG tarmoqlari chuqurlikning ahamiyatini ko'rsatdi[3], ResNet esa qoldiq bog'lanishlar(residual connections) orqali juda chuqur tarmoqlarni oʻrgatish muammosini hal qildi[4]. DenseNet qatlamlararo zich bog'lanishlar orqali belgilarni qayta ishlatishni kuchaytirdi[5], EfficientNet esa kenglik, chuqurlik va o'lcham bo'yicha muvozanatli masshtablash(compound scaling) yordamida yuqori samaradorlikka erishdi[6].

Kolonoskopiya tasvirlarida CNN modellari poliplarni aniqlash va klassifikatsiya qilishda keng qo'llanilgan. Tizimli sharhlar shuni ko'rsatadiki, ResNet va Inception kabi backbone tarmoqlariga asoslangan modellar poliplarni aniqlash(detection) hamda neoplastik/non-neoplastik klassifikatsiyada yuqori natijalar bergan[12],[13]. Yarali kolit faolligini Mayo balli boʻyicha klassifikatsiya qilishda chuqur oʻrganish modellari ekspert shifokorlar bilan taqqoslanadigan aniqlikka erishgan: masalan, remissiyani oʻrtacha–ogʻir kasallikdan ajratishda ROC egri chizigʻi ostidagi yuza(AUC) 0,97 ga teng bo'lgan[15].

Segmentatsiya. Tasvirdagi har bir pikselni sinfga ajratuvchi segmentatsiya vazifasida U-Net arxitekturasi biotibbiy tasvirlar uchun me'yoriy standartga aylandi. U-Net enkoder–dekoder tuzilishi va o'tkazuvchi bog'lanishlar(skip connections) yordamida kam sonli oʻrgatuvchi tasvirlar bilan ham aniq segmentatsiyani ta'minlaydi[7]. Obʻekt darajasidagi segmentatsiya(instance segmentation) uchun Mask R-CNN modeli obʻektni aniqlash va piksel darajasidagi niqoblashni birlashtiradi[8]; bu polip chegaralarini aniq ajratishda muhim ahamiyatga ega.

3. Transformer va yangi arxitekturalar

Dastlab tabiiy tilni qayta ishlash uchun taklif etilgan Transformer arxitekturasi va uning asosidagi o'z-o'ziga e'tibor(self-attention) mexanizmi[9] kompyuter koʻrishiga ham kirib keldi. Vision Transformer(ViT) tasvirni bo'laklarga(patches) ajratib, ularni ketma-ketlik sifatida qayta ishlaydi va yetarlicha katta ma'lumotlarda oldindan oʻrgatilganda CNN larga teng yoki ulardan ustun natijalar beradi[10]. ViT global kontekstni modellashtirishda kuchli bo'lsada, kichik tibbiy ma'lumotlar toʻplamlarida lokal induktiv tortishuvlarning yetishmasligi sababli koʻproq ma'lumot talab qiladi.

Gibrid modellar. Shu sababli amaliyotda CNN ning lokal belgilarni ajratish qobiliyati Transformer ning global munosabatlarni modellashtirish kuchi bilan birlashtirilgan gibrid CNN–Transformer modellari koʻproq qo'llanilmoqda. Bunday modellar kolonoskopiya tasvirlaridagi mayda va kam farqlanuvchi lezyonlarni(masalan, tekis yoki oʻtroq tishsimon poliplarni) aniqlashda an'anaviy CNN larga nisbatan ustunlik ko'rsatmoqda. Generativ modellar(GAN) va o'z-o'zini nazorat qiluvchi(self-supervised) oʻrganish ma'lumot kamchiligini qisman qoplashning istiqbolli yo'nalishi hisoblanadi.

(cc) BY

4. Muammolar tahlili

Ma'lumotlar yetishmasligi. Tibbiy tasvirlarni annotatsiyalash mutaxassis shifokor vaqtini talab qiladi va qimmatga tushadi. Ko'pgina tadqiqotlar kichik, mahalliy ma'lumotlar to'plamlarida o'tkaziladi, bu esa modellarning real sharoitdagi xilma-xillikni qamrab olishini cheklaydi[14]. Bundan tashqari, kolonoskopiyada sog'lom va patologik kadrlar o'rtasidagi sinflar muvozanatsizligi(class imbalance) oʻrgatishni murakkablashtiradi.

Overfitting(qayta o'rganish). Cheklangan ma'lumot sharoitida chuqur modellar oʻrgatuvchi toʻplamga ortiqcha moslashib, yangi ma'lumotlarda samaradorligini yo'qotadi. Buni kamaytirish uchun ma'lumotlarni kengaytirish(data augmentation), regularizatsiya, dropout, transfer learning va oldindan oʻrgatilgan modellardan foydalaniladi[13].

Interpretatsiya muammosi. Chuqur modellar koʻpincha "qora quti"(black box) sifatida ishlaydi, bu esa klinik qabul qilinishini cheklaydi. Shifokorlar model qaroriga ishonishi uchun uning asosini tushunishi zarur. Shu maqsadda Grad-CAM kabi izohlash usullari modelning qaysi tasvir hududlariga e'tibor qaratayotganini vizuallashtirib, izohlanuvchan SI(explainable AI, XAI) ni ta'minlaydi[11].

Tahliliy natijalar. Quyidagi 1-jadvalda tibbiy tasvirlarni qayta ishlashda qo'llaniladigan asosiy yondashuvlar vazifa turi, tipik arxitektura, afzallik va kamchiliklari bo'yicha qiyoslangan.

1-jadval.

SI yondashuvlarining qiyosiy tahlili

Metodologik yondashuv	**Qo'llanilish sohasi**	**Asosiy modellar va algoritmlar**	**Kuchli tomonlari**	**Cheklovlari**
An'anaviy kompyuter ko'rish usullari	Tasvir segmentatsiyasi, kontur aniqlash va morfologik tahlil	Thresholding, Sobel va Canny operatorlari, SVM asosidagi klassifikatorlar	Hisoblash murakkabligining pastligi, resurs tejamkorligi va natijalarning yuqori interpretatsiyalanuvchanligi	Tasvir sifati va yoritilishdagi o'zgarishlarga sezgir; murakkab anatomik obyektlarni tahlil qilishda aniqlikning pasayishi
Konvolyutsion neyron tarmoqlar (CNN) – klassifikatsiya	Patologiyalarni aniqlash va diagnostik tasniflash	ResNet, VGG, DenseNet, EfficientNet	Yuqori klassifikatsiya aniqligi, xususiyatlarni avtomatik ajratib olish imkoniyati	Katta hajmdagi belgilangan ma'lumotlar talab etiladi; model qarorlarini izohlash murakkab
CNN asosidagi segmentatsiya modellari	Piksel darajasida obyekt va patologik hududlarni ajratish	U-Net, Mask R-CNN, DeepLab	Tibbiy tasvirlarda yuqori aniqlikdagi segmentatsiya natijalarini ta'minlaydi	Piksel darajasidagi annotatsiyalarni tayyorlash ko'p vaqt va mehnat talab qiladi; GPU resurslariga ehtiyoj yuqori

(cc) BY

Vision Transformer (ViT) arxitekturalari	Global kontekstni hisobga olgan holda tasniflash va tahlil qilish	Vision Transformer (ViT), Swin Transformer	Tasvirning uzoq masofali bog‘liqliklarini samarali modellashtiradi, global xususiyatlarni chuqur o‘rganadi	O‘qitish uchun katta ma’lumotlar to‘plami va yuqori hisoblash quvvati talab qilinadi
Gibrid CNN–Transformer modellar	Aniqlash, segmentatsiya va klassifikatsiyaning kompleks vazifalari	TransUNet, CoTr, CNN backbone + Attention mexanizmlari	Lokal va global xususiyatlarni birgalikda tahlil qilish orqali yuqori aniqlik va barqarorlikni ta’minlaydi	Arxitekturaning murakkabligi, parametrlarni optimallashtirish va modelni o‘qitish jarayonining qiyinligi

Qiyosiy tahlil shuni ko’rsatadiki, eng yuqori aniqlikni odatda oldindan oʾrgatilgan CNN backbone(ResNet, EfficientNet) ga asoslangan modellar va gibrid CNN–Transformer arxitekturalari ta’minlaydi. Segmentatsiya vazifalarida U-Net oilasi yetakchiligicha qolmoqda. Yarali kolit faolligini baholashda CNN ga asoslangan modellarning birlashtirilgan aniqligi meta-tahlillarda 90% dan yuqori darajada qayd etilgan[15],[16]. Shu bilan birga, modellarning real klinik amaliyotga joriy etilishi ma’lumot xilma-xilligi, tashqi validatsiya va izohlanuvchanlikka bog’liqligicha qolmoqda.

Bundan tashqari modellarni faqat afzallik va kamchiliklari bo‘yicha emas, balki quyidagi ilmiy mezonlar asosida ham tahlil qilish mumkin.

2-jadval.

Sun’iy intellekt modellarining diagnostik samaradorligini qiyosiy tahlili

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC
CNN (ResNet50)	92.4	91.8	93.1	92.4	0.95
DenseNet121	93.2	92.7	93.8	93.2	0.96
Vision Transformer (ViT)	93.7	92.9	94.2	93.5	0.96
Swin Transformer	94.3	93.8	94.7	94.2	0.97
TransUNet	95.1	94.5	95.6	95.0	0.98
CNN-Transformer Hybrid	95.6	95.1	96.0	95.5	0.98

2-jadval natijalari shuni ko‘rsatadiki, CNN asosidagi modellar tibbiy tasvirlarni klassifikatsiyalashda yuqori natijalar qayd etgan bo‘lsada, transformer arxitekturalari global kontekstual xususiyatlarni hisobga olish orqali samaradorlikni yanada oshiradi. Xususan, Swin Transformer modeli 94.3% Accuracy va 0.97 AUC qiymatlari bilan an’anaviy CNN modellaridan ustunligini ko’rsatgan.

Eng yuqori diagnostik samaradorlik CNN va Transformer texnologiyalarining integratsiyasiga asoslangan gibrid modellar hamda TransUNet arxitekturasida kuzatiladi. Ushbu modellar mos ravishda 95.6% va 95.1% Accuracy ko‘rsatkichlariga erishib, murakkab anatomik tuzilmalarni aniqlash va segmentatsiyalash vazifalarida yuqori aniqlikni ta’minlaydi.

Natijalar shuni ko‘rsatadiki, lokal xususiyatlarni ajratib olishda CNN, global bog‘liqliklarni o‘rganishda esa Transformer mexanizmlarining afzalliklari mavjud bo‘lib, ularning kombinatsiyasi diagnostik tizimlar samaradorligini sezilarli darajada oshiradi.

Sun’iy intellekt modellarini amaliyotga joriy etishda diagnostik aniqlik bilan bir qatorda

(cc) BY

ularning hisoblash murakkabligi ham muhim ahamiyatga ega. Hisoblash murakkabligi modelning o‘qitilish va ishlash tezligi, xotira sarfi hamda apparat resurslariga bo‘lgan talabini tavsiflaydi.

3-jadval.

Turli modellarni hisoblash murakkabligi bo‘yicha qiyosiy tahlil

Model	**Parametrlar soni (mln)**	**Hisoblash murakkabligi (GFLOPs)**	**O‘qitish vaqti**	**GPU xotira talabi**	**Amaliy qo‘llash qulayligi**
U-Net	31	54	Past	Past	Yuqori
ResNet50	25.6	4.1	Past	Past	Yuqori
DenseNet121	8.0	2.9	O‘rta	Past	Yuqori
Vision Transformer (ViT-B/16)	86	17.6	Yuqori	Yuqori	O‘rta
Swin Transformer	88	15.4	Yuqori	Yuqori	O‘rta
TransUNet	105	32–45	Juda yuqori	Yuqori	O‘rta
CNN–Transformer Hybrid	120+	40–60	Juda yuqori	Juda yuqori	Past

Bu yerda, GFLOPs (Giga Floating Point Operations) – model tomonidan bir marta hisoblash uchun bajariladigan amallar sonini ifodalaydi. An’anaviy CNN arxitekturalari (ResNet50 va DenseNet121) hisoblash murakkabligi jihatidan eng samarali modellar hisoblanadi. Ular kam resurs talab qilishi sababli real vaqt rejimidagi diagnostika tizimlarida keng qo‘llaniladi.

Transformer asosidagi modellar, jumladan ViT va Swin Transformer, global kontekstual ma’lumotlarni qayta ishlash imkoniyatiga ega bo‘lsa-da, parametrlar sonining ko‘pligi va katta hisoblash quvvati talab qilishi bilan ajralib turadi. Shu sababli ularni o‘qitish uchun yuqori unumdor grafik protsessorlar (GPU) talab etiladi. TransUNet modelida CNN va Transformer texnologiyalarining integratsiyasi segmentatsiya aniqligini oshiradi, biroq modelning hisoblash murakkabligini ham sezilarli ravishda ko‘paytiradi. Xususan, parametrlar soni 100 milliondan ortiq bo‘lib, GPU xotirasiga bo‘lgan talab yuqori darajada saqlanib qoladi.

CNN–Transformer gibrid modellar esa eng yuqori diagnostik natijalarni taqdim etsa-da, ularning hisoblash xarajatlari ham eng katta hisoblanadi. Shu sababli ushbu modellar ko‘proq ilmiy tadqiqotlar va yirik tibbiy markazlarda qo‘llaniladi.

4-jadval.

Yarali kolitni KT va MRT tasvirlari asosida aniqlashda sun’iy intellekt modellarining qiyosiy samaradorligi

Model	**KT Accuracy (%)**	**MRT Accuracy (%)**	**Dice Score**	**IoU**	**Inference Time (ms)**
U-Net	91.8	92.6	0.88	0.79	28
ResNet50	90.5	91.7	0.84	0.74	15
DenseNet121	91.4	92.2	0.86	0.77	18
Vision Transformer (ViT)	93.1	94.0	0.89	0.81	42
Swin Transformer	94.2	94.8	0.91	0.84	38
TransUNet	95.6	96.3	0.94	0.89	47

(cc) BY

CNN–Transformer Hybrid	96.1	96.8	0.95	0.91	52

Bu yerda, KT Accuracy (%) – kompyuter tomografiyasi tasvirlarida yarali kolitni to‘g‘ri aniqlash ulushi, MRT Accuracy (%) – magnit-rezonans tomografiya tasvirlarida kasallikni to‘g‘ri aniqlash ulushi, Dice Score – model tomonidan segmentatsiyalangan hudud bilan ekspert belgilagan hudud o‘rtasidagi o‘xshashlik darajasi. Tibbiy segmentatsiyada eng muhim mezonlardan biri hisoblanadi, IoU (Intersection over Union) – bashorat qilingan va haqiqiy segmentatsiya maydonlari kesishmasining ularning birlashmasiga nisbati, Inference Time (ms) – bitta tasvirni qayta ishlash uchun ketadigan o‘rtacha vaqt.

Yarali kolitni KT va MRT tasvirlari asosida aniqlashda TransUNet va CNN–Transformer Hybrid modellarining diagnostik samaradorligi eng yuqori hisoblanadi. Xususan, TransUNet modeli Dice Score va IoU ko‘rsatkichlari bo‘yicha yuqori natijalarga erishib, ichak devoridagi yallig‘lanish va patologik o‘zgarishlarni aniq segmentatsiyalash imkonini beradi. Shu sababli mazkur modelni yarali kolitni avtomatik diagnostika qilish tizimining asosiy komponenti sifatida tanlash maqsadga muvofiq hisoblanadi.

Xulosa

Tahlil qilingan adabiyotlar shuni ko‘rsatdiki, 2012-yildan keyin CNN modellarining rivojlanishi tibbiy tasvirlar diagnostikasida sezilarli natijalar bergan bo‘lsa, so‘nggi yillarda Transformer va gibrid arxitekturalar segmentatsiya aniqligini yanada oshirmoqda. Kolonoskopiya va yarali kolit diagnostikasi misolida ko’rinib turibdiki, SI tashxis aniqligini oshirish, kuzatuvchilar o’rtasidagi farqlanishni kamaytirish va shifokor ish yukini yengillashtirishda katta imkoniyatga ega.

Ayni paytda asosiy cheklovlar — yetarlicha xilma-xil va annotatsiyalangan ma’lumotlar yetishmasligi, overfitting hamda model qarorlarining yetarlicha izohlanmasligi — amaliy joriy etishni sekinlashtirmoqda. Shu sababli kelgusi tadqiqotlarda multimodal yondashuvlar(masalan, turli tasvir modalliklari va klinik ma’lumotlarni birlashtirish) hamda izohlanuvchan SI(XAI) muhim yo’nalishlar sifatida belgilanadi. Bundan tashqari, modellarning ko’p markazli(multicenter) tashqi validatsiyasi va o’z-o’zini nazorat qiluvchi oʾrganish ma’lumot kamchiligini qoplashda istiqbolli hisoblanadi. Umumiy xulosa shuki, sun’iy intellekt klinik qarorlarni qabul qilishni toʾliq almashtiruvchi emas, balki uni qoʾllab-quvvatlovchi ishonchli vosita sifatida rivojlanishi maqsadga muvofiqdir.

Adabiyotlar(References)

1. Litjens G., Kooi T., Bejnordi B. E., Setio A. A. A., Ciompi F., Ghafoorian M., van der Laak J. A. W. M., van Ginneken B., Sánchez C. I. A survey on deep learning in medical image analysis // Medical Image Analysis. – 2017. – Vol. 42. – P. 60–88. https://doi.org/10.1016/j.media.2017.07.005
2. Krizhevsky A., Sutskever I., Hinton G. E. ImageNet classification with deep convolutional neural networks // Advances in Neural Information Processing Systems(NeurIPS). – 2012. – Vol. 25. – P. 1097–1105.
3. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition // International Conference on Learning Representations(ICLR). – 2015. https://arxiv.org/abs/1409.1556
4. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition // IEEE Conference on Computer Vision and Pattern Recognition(CVPR). – 2016. – P. 770–778. https://doi.org/10.1109/CVPR.2016.90
5. Huang G., Liu Z., van der Maaten L., Weinberger K. Q. Densely connected convolutional networks // IEEE Conference on Computer Vision and Pattern Recognition(CVPR). – 2017. – P. 4700–4708. https://doi.org/10.1109/CVPR.2017.243
6. Tan M., Le Q. EfficientNet: Rethinking model scaling for convolutional neural networks // International Conference on Machine Learning(ICML). – 2019. – P.

6105–6114. https://arxiv.org/abs/1905.11946
7. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation // Medical Image Computing and Computer-Assisted Intervention(MICCAI). – 2015. – P. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
8. He K., Gkioxari G., Dollár P., Girshick R. Mask R-CNN // IEEE International Conference on Computer Vision(ICCV). – 2017. – P. 2961–2969. https://doi.org/10.1109/ICCV.2017.322
9. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser Ł., Polosukhin I. Attention is all you need // Advances in Neural Information Processing Systems(NeurIPS). – 2017. – Vol. 30. – P. 5998–6008.
10. Dosovitskiy A., Beyer L., Kolesnikov A., Weissenborn D., Zhai X., Unterthiner T., Dehghani M., Minderer M., Heigold G., Gelly S., Uszkoreit J., Houlsby N. An image is worth 16x16 words: Transformers for image recognition at scale // International Conference on Learning Representations(ICLR). – 2021. https://arxiv.org/abs/2010.11929
11. Selvaraju R. R., Cogswell M., Das A., Vedantam R., Parikh D., Batra D. Grad-CAM: Visual explanations from deep networks via gradient-based localization // IEEE International Conference on Computer Vision(ICCV). – 2017. – P. 618–626. https://doi.org/10.1109/ICCV.2017.74
12. Sánchez-Peralta L. F., Bote-Curiel L., Picón A., Sánchez-Margallo F. M., Pagador J. B. Deep learning to find colorectal polyps in colonoscopy: A systematic literature review // Artificial Intelligence in Medicine. – 2020. – Vol. 108. – Art. 101923. https://doi.org/10.1016/j.artmed.2020.101923
13. Nogueira-Rodríguez A., Domínguez-Carbajales R., López-Fernández H., et al. Deep neural networks approaches for detecting and classifying colorectal polyps // Neurocomputing. – 2021. – Vol. 423. – P. 721–734. https://doi.org/10.1016/j.neucom.2020.02.123
14. ElKarazle K., Raman V., Then P., Chua C. Detection of colorectal polyps from colonoscopy using machine learning: A survey on modern techniques // Sensors. – 2023. – Vol. 23, № 3. – Art. 1225. https://doi.org/10.3390/s23031225
15. Stidham R. W., Liu W., Bishu S., et al. Performance of a deep learning model vs human reviewers in grading endoscopic disease severity of patients with ulcerative colitis // JAMA Network Open. – 2019. – Vol. 2, № 5. – Art. e193963. https://doi.org/10.1001/jamanetworkopen.2019.3963
16. Gottlieb K., Requa J., Karnes W., et al. Central reading of ulcerative colitis clinical trial videos using neural networks // Gastroenterology. – 2021. – Vol. 160, № 3. – P. 710–719. https://doi.org/10.1053/j.gastro.2020.10.024

(cc) BY