

ՄԵՔԵՆԱՑԱԿԱՆ ԵՎ ԽՈՐ ՈՒՍՈՒՑՄԱՆ ՄԵԹՈՂՆԵՐԻ ԿԻՐԱՌՄԱՍԲ ՕՂԱՏԻԵԶԵՐԱԿԱՆ
ՊԱՏԿԵՐՆԵՐՈՒՄ ՕԲՑԵԿՏՆԵՐԻ ՀԱՅՏՆԱԲԵՐՈՒՄ

Լուսինե Ներսեսի Եղիյան

«Երկրատարածական տեխնոլոգիաների կենտրոն» ՄՊԸ, ք. Երևան, ՀՀ

lusineyeghyan@gmail.com

Օդատիեզերական պատկերները լայնորեն կիրառվում են տարբեր ոլորտներում (շրջակա միջավայրի մշտադիտարկում, քաղաքաշինություն, աղետների կառավարում և այլն), սակայն պատկերից օբյեկտների ավտոմատ հայտնաբերումը մնում է բարդ խնդիր: Հողվածում ուսումնասիրվում են մեքենայական և խոր ուսուցման մեթոդների կիրառումը՝ հենք ընդունելով դրանց կիրառական բաղադրիչն ու արդյունավետությունը: Դիտարկվում են ժամանակակից մոտեցումներ, մասնավորապես, CNN, Mask R-CNN և Segment Anything Model (SAM) մոդելները: Իրականացվում է փորձնական վերլուծություն՝ գնահատելով մոդելների ճշգրտությունը, կիրառելիությունը և վերջիններիս հնարավորությունները ՀՀ-ի համար: Արդյունքները ցույց են տալիս, որ խոր ուսուցման մեթոդներն ապահովում են բարձր ճշգրտություն և ավտոմատացման մակարդակ, միաժամանակ ունենալով որոշ սահմանափակումներ՝ կապված աղմուկի և ընդհանրացման խնդիրների հետ:

Բանալի բառեր. օդատիեզերական պատկերներ, խոր ուսուցում, մեքենայական ուսուցում, օբյեկտների հայտնաբերում, սեզմենտացիա, CNN, Mask R-CNN, Segment Anything Model

Ներածություն

Բարձր լուծաչափով օդատիեզերական պատկերները բարձացնում են տվյալների մշակման ավտոմատացման հնարավորությունները և ճշգրտությունը: Հատկապես կարևոր է տարբեր տիպի օբյեկտների՝ շենքերի, ճանապարհների, բնական կառույցների և այլ տարբերի հայտնաբերումն ու սեզմենտացիան, որոնք առանցքային նշանակություն ունեն տարածական վերլուծության համար:

Ավանդական մեթոդները (թվայնացում, սպեկտրալ դասակարգում և այլն) բարդ ֆունային և կառուցվածքային օբյեկտների դեպքում (շինություն, ճանապարհ և այլն) ունեն ցածր կիրառելիություն: Խոր ուսուցման մեթոդների զարգացումը, մասնավորապես, կոնվոլյուցիոն նեյրոնային ցանցերի (CNN) կիրառումը, զգալիորեն բարձրացրել է տարբեր օբյեկտների հայտնաբերման և մեկնաբանման ճշգրտությունը [1]: Ժամանակակից semantic segmentation մոտեցումները ներառում են նաև Fully Convolutional Networks (FCN) ճարտարապետությունը [2]: Օդատիեզերական պատկերների մշակման համար լայնորեն կիրառվում են նաև Pyramid Scene Parsing Network (PSPNet) մոդելները [3], որոնք արագորեն զարգանում են, սակայն ներկայումս ամենատարածված մոդելներն են DeepLab-երը, որոնք լայնորեն կիրառվում են իմաստային սեզմենտացիայի

(semantic segmentation) խնդիրներում [4], բայցևայնպես պատկերի ճանաչման կամ հայտնաբերման համար կիրառվում են նաև YOLO ընտանիքի մոդելները [5]:

Տարբեր մոդելներ, օրինակ Mask R-CNN, ապահովում է օբյեկտների պիքսելային մակարդակով սեգմենտացիա, իսկ մշակված Segment Anything Model (SAM) առաջարկում է առավել համապիտանի մոտեցում՝ տարբեր բնույթի օբյեկտների ճկուն առանձնացման համար:

Մինևս ժամանակ առկա են խնդիրներ, կապված տարբեր օբյեկտների բազմազանության, մոդելների ադապտացիայի և ադուկի նկատմամբ կայունության հետ, ինչը պայմանավորում է թեմայի արդիականությունը:

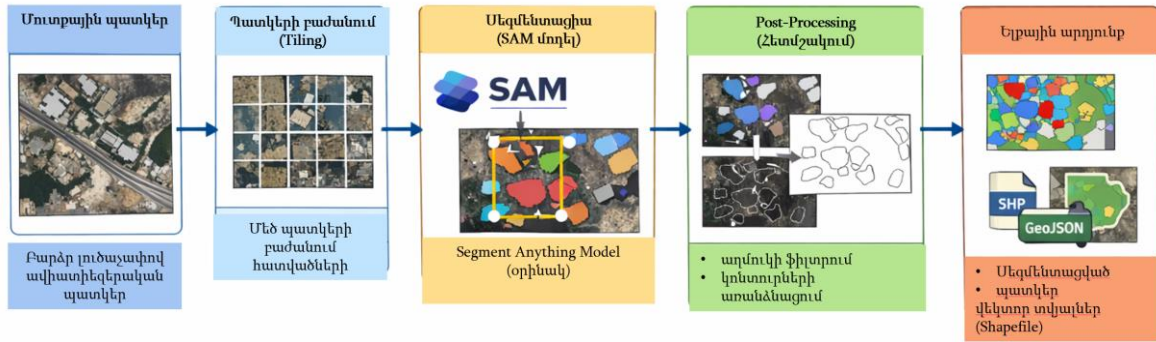
Հոդվածում ուսումնասիրվում են մեքենայական և խոր ուսուցման մեթոդների կիրառման հնարավորություններն օդատիեզերական պատկերներում տարբեր օբյեկտների հայտնաբերման և սեգմենտացիայի միջոցով: Դիտարկվում են ժամանակակից մոտեցումներ, մասնավորապես, CNN-հիմքով մոդելներ, Mask R-CNN և Segment Anything Model (SAM): Կատարված վերլուծությունը ցույց է տալիս, որ խոր ուսուցման մեթոդներն ապահովում են բարձր ճշգրտություն և ավտոմատացման մակարդակ՝ միաժամանակ պահպանելով որոշ սահմանափակումներ՝ կապված ընդհանրացման և բարդ տեսարանների հետ:

Սույն հետազոտության նպատակն է գնահատել մեքենայական և խոր ուսուցման մոդելների կիրառելիությունը բարձր լուծաչափով օդատիեզերական պատկերների ավտոմատ սեգմենտացիայի համար: Ուսումնասիրվել և համեմատվել են U-Net, Mask R-CNN և SAM մոդելները՝ գնահատելով դրանց ճշգրտությունն ու կիրառելիությունը ՀՀ-ում:

Մոդելների կիրառման համար բարդ ռեիեֆը, քարքարոտ մակերեսները և խիտ կառուցապատված տարածքները, ուղղակիորեն ազդում են ճշգրտության վրա: Առաջարկված մոտեցումը փորձարկվել է Արալեզ բնակավայրում՝ շեշտադրելով ոչ թե տարածքը, այլ իրականացված մեթոդը:

Նյութեր և մեթոդներ

Իրականացվել է բարձր լուծաչափով օդատիեզերական պատկերներում օբյեկտների ավտոմատ առանձնացում՝ հիմնված Meta AI-ի Segment Anything Model (SAM) սեգմենտացիոն մոդելի վրա: SAM-ը նախագծված է որպես promptable և class-agnostic մոդել՝ ապահովելով zero-shot փոխանցելիություն նոր պատկերային բաշխումների և խնդիրների նկատմամբ [6] (նկ. 1): Ինչպես նաև իրականացվել է բարձր լուծաչափով օդատիեզերական պատկերից օբյեկտների հայտնաբերմանը միտված տարաբանյութ մեթոդների համադրում՝ ներառելով պատկերների հատվածավորում (tiling), սեգմենտացիոն դիմակների (mask)-երի համադրում և արդյունքների հետմշակում:



Նկ. 1. Աշխատանքի իրականացման պարզ սխեմա

Սեգմենտացիայից հետո իրականացվել է դիմակների (mask) հետմշակում՝ աղմուկից մաքրելու նպատակով: Հետմշակման փուլը հնարավորություն է տվել նվազեցնել false positive արդյունքները և բարելավել վերջնական արդյունքի ճշգրտությունը:

Մուտքային պատկեր

Հետազոտության ընթացքում կիրառվել է ՀՀ տարածքի բարձր լուծաչափով օդատիեզերական պատկեր՝ ստացված UAV/օդալուսանկարահանման միջոցով: Պատկերի տարածական լուծաչափը կազմել է մինչև 5 սմ/պիքսել: Մուտքային տվյալները ներկայացված են GeoTIFF ձևաչափով և ներառում են RGB ալիքներ: Հետազոտվող տարածքը ներառում է շենք-շինություններ և բարդ կառուցվածքային տարրեր, որոնք օգտագործվել են սեգմենտացիոն մոդելների համեմատական գնահատման համար՝ հետագա փուլերում վեկտորացումը և տարածական վերլուծությունը ճշգրիտ իրականացնելու նպատակով: Սեգմենտացիայի համար կիրառվել է RGB ներկայացում՝ որպես մուտքային տվյալ (նկ. 2):



Նկ. 2. Բարձր լուծաչափով պատկեր (հատված հետազոտվող տարածքից)

Կիրառվել է կետային ամպ, որպեսզի ուսուցման տվյալները և վալիդացիոն տվյալների համեմատությունը լինի առավել ճշգրիտ (տվյալ դեպքում շենք - շինությունների համար): Հատկապես, որ կիրառված օրինակի դեպքում ուսուցման փուլը բացակայում է (նկ. 3):

Կետային ամպի տվյալները կիրառվել են որպես լրացուցիչ ստուգաչափական աղբյուր՝ շենքերի եզրագծերի և բարձրությունների համեմատական գնահատման նպատակով: Կետային ամպի կիրառությունը հնարավորություն է տվել նվազեցնել false positive արդյունքները և բարձրացնել շենքերի ավելի բարձր տեսանելիություն [7]:



Նկ. 3. Օդայուսանկարը և լիդար տվյալների համակցություն

GPU հիշողության սահմանափակումներից խուսափելու համար կիրառվել է պատկերի բաժանում հատվածների (tiling) սահող պատուհանի (sliding window) սկզբունքով՝ սահմանված tile չափով և նվազեցնելով overlay/overlap-ը՝ եզրային անհամապատասխանությունները [2]:

Նախնական մշակում

Մուտքային պատկերը (GeoTIFF) փոխարկվում է հաշվարկների համար հարմար պատկերային ներկայացման ձևով ($H \times W \times C$): Պատկերի համար կիրառվում է ինտենսիվության մասշտաբավորում, մասնավորապես, ուսումնասիրված տարածքի համար (16-bit/float ռաստերներ) կիրառվել է $[0, 255]$ միջակայքը: Հետագայում իրականացվել է պատկերի բաժանում հատվածների (tiling), որի նպատակը մեծ ծավալների հաշվարկային բարդությունն ու GPU հիշողության սահմանափակումները նվազեցնելն է: Tile -ի չափերն ըստ պատկերի ընտրվել են $512 \dots 512$ px, իսկ համադրումը (overlap)՝ $10 \dots 25$ %, որն օգնում է նվազեցնել եզրային կտրտվածության (edge effects) ազդեցությունը (նկ. 4) [6]:



Նկ. 4. Նախնական սեգմենտացիայի արդյունք

Ինչպես նշվեց վերևում, օդատիեզերական պատկերների սեգմենտացիայի համար լայնորեն կիրառվում են ինչպես դասական խոր ուսուցման մոդելներ, այնպես էլ նոր սերնդի foundation մոդելներ: Հետազոտության ընթացքում կիրառվել են հետևյալ մոդելները՝ U-Net, Mask R-CNN և Segment Anything Model (SAM): Մոդելներն ունեն ճարտարապետական, մաթեմատիկական և գործառնական տարբերություններ, որոնց միջոցով կարելի է տարբերակել արդյունքների ճշգրտությունը (աղ. 1, 2, 3):

Աղյուսակ 1

Սլզգորհիթմի ճարտարապետական տարբերություններ

Մոդել	Տիպ	Հիմնական գաղափար
U-Net*	Սեմանտիկ սեգմենտացիա	Encoder-decoder + skip connection
Mask R-CNN**	Instance սեգմենտացիա	Օբյեկտների հայտնաբերում + mask
SAM***	Foundation model	Prompt-based սեգմենտացիա

* - **U-Net** — դասական encoder-decoder կառուցվածք է, որը լավ է աշխատում փոքր dataset-ներով [8],

** - **Mask R-CNN** — օբյեկտների հայտնաբերում + mask, լավ է instance-level խնդիրների համար [9],

*** - **SAM** — prompt-driven, zero-shot, չի պահանջում վերապատրաստում կոնկրետ տվյալների վրա [6]:

Աղյուսակ 2

Այգորիթմի գործնական համեմատություն

Չափանիշ	U-Net	Mask R-CNN	Segment Anything Model (SAM)
Տվյալների պահանջ	Բարձր՝ պահանջում է մեծ ծավալի նշագրված (annotated) տվյալներ	Բարձր՝ պահանջում է ճշգրիտ instance-level annotation	Ցածր՝ չի պահանջում նախնական ուսուցում (zero-shot)
Ճշգրտություն	Բարձր՝ համապատասխան ուսուցման դեպքում	Բարձր՝ instance-level սեգմենտացիայի համար	Բարձր՝ ընդհանուր կիրառման դեպքում
Fine-tuning անհրաժեշտություն	Պարտադիր	Պարտադիր	Չի պահանջվում
Հաշվարկային արագություն	Միջին	Ցածր (հաշվարկային բարդության պատճառով)	Բարձր
Ընդհանրացման կարողություն (Generalization)	Սահմանափակ, կախված ուսուցման տվյալներից	Միջին	Բարձր
Կիրառելիություն օդատիեզերական պատկերներում	Արդյունավետ	Արդյունավետ	Բարձր արդյունավետություն

Աղյուսակ 3

Սեգմենտացիոն մոդելների համեմատական գնահատում

Մոդել	IoU	Precision	Recall	F1-score
U-Net	0,72	0,76	0,70	0,73
Mask R-CNN	0,78	0,81	0,75	0,78
SAM	0,82	0,84	0,80	0,82

Սեգմենտացիոն մոդելների գնահատման համար ընտրվել հատվածներ, որտեղ առկա են շենք-շինություններ, ճանապարհներ, ծառեր և այլն: Գնահատումն իրականացվել է reference segmentation տվյալների և մոդելների արդյունքների համեմատության միջոցով:

Համեմատական վերլուծության ընթացքում հաշվարկվել են IoU, Precision, Recall և F1-score ցուցանիշները: Գնահատման համար կիրառվել են հատվածավորված (tile-based) բարձր լուծաչափով օդատիեզերական պատկերներ, իսկ ստացված արդյունքները ներկայացնում են մոդելների հարաբերական արդյունավետությունը հետազոտվող տարածքի պայմաններում:

Ստացված արդյունքները ցույց են տալիս, որ SAM մոդելն ապահովում է համեմատաբար բարձր ճշգրտություն և ավելի լավ ընդհանրացման կարողություն՝ հատկապես բարդ կառուցվածք ունեցող օբյեկտների դեպքում: U-Net և Mask R-CNN մոդելները ևս ցուցաբերել են բավարար արդյունքներ, սակայն դրանց արդյունավետությունը մեծապես կախված է նախնական նշագրված տվյալների որակից և ծավալից:

Արդյունքներ և քննարկում

Տարբեր սեզմենտացիոն մոդելների կիրառումը ցույց է տալիս, որ դրանց արդյունավետությունը կախված է ինչպես տվյալների բնույթից, այնպես էլ առաջադրված խնդրի տեսակից:

U-Net մոդելը կարող է ապահովել բարձր ճշգրտություն, հատկապես լավ նշագրված տվյալների առկայության դեպքում, սակայն դրա կիրառումը սահմանափակվում է մեծ ծավալի annotation-ի անհրաժեշտությամբ [10]:

Mask R-CNN մոդելը հնարավորություն է տալիս իրականացնել օբյեկտների instance-level հայտնաբերում և սեզմենտացիա, ինչը կարևոր է առանձին օբյեկտների տարանջատման համար, սակայն այն պահանջում է զգալի հաշվարկային ռեսուրսներ և բարդ ուսուցման գործընթաց:

Segment Anything Model-ը (SAM) առանձնանում է իր ճկունությամբ և zero-shot կիրառելիությամբ՝ հնարավորություն տալով իրականացնել սեզմենտացիա առանց նախնական վերապատրաստման տվյալների [11]: Այն հատկապես արդյունավետ է բարդ ճարտարապետական տեսք ունեցող օբյեկտների և մեծ չափերով պատկերների համար, ինչպես նաև հարմար է մեծ չափերի պատկերների մշակման համար: Մինևս ժամանակ որոշ դեպքերում հնարավոր է ստանալ ավելորդ կամ աղմուկային mask-եր, որը հետագա փուլերում պահանջում է հետմշակում [12]:

Ընդհանուր առմամբ, կատարված հետազոտության արդյունքները ցույց են տալիս, որ առավել արդյունավետ է SAM մոդելի կիրառումը, որը, պահպանելով բարձր ճշգրտություն, ապահովում է նաև բարձր ընդհանրացման կարողություն և չի պահանջում նախնական ուսուցում:

Եզրակացություն

Սույն հետազոտության շրջանակում ուսումնասիրվել են օդատիեզերական պատկերների սեզմենտացիայի մեթոդները՝ կիրառելով U-Net, Mask R-CNN և Segment Anything Model (SAM) մոդելները: Մեթոդաբանությունը ներառել է տվյալների նախնական մշակում, սեզմենտացիա և հետմշակում:

Արդյունքները ցույց են տալիս, որ դասական մոդելներն ապահովում են բարձր ճշգրտություն՝ համապատասխան ուսուցման տվյալների առկայության դեպքում, մինչդեռ SAM մոդելը ցուցաբերում է բարձր արդյունավետություն առանց նախնական ուսուցման՝ ապահովելով լավ ընդհանրացում տարբեր պատկերային պայմաններում:

Խոր ուսուցման և foundation model-ների կիրառությունը հեռավորման պատկերների մշակման ոլորտում շարունակաբար ընդլայնվում է՝ ապահովելով ավելի բարձր ճշգրտություն, ավտոմատացման մակարդակ և տարբեր տարածական պայմաններում ընդհանրացման հնարավորություն [11]:

Այսպիսով, հետազոտության արդյունքները հաստատում են, որ SAM մոդելի կիրառումը հանդիսանում է արդյունավետ մոտեցում օդատիեզերական պատկերների ավտոմատ սեզմենտացիայի համար:

Ստացված համեմատական արդյունքները ցույց են տալիս, որ SAM մոդելը ապահովում է ավելի բարձր ընդհանրացման հնարավորություն և արդյունավետություն մեծածավալ օդատիեզերական պատկերների մշակման ընթացքում, ինչը հաստատում է foundation model-ների կիրառելիության բարձր ներուժը հեռագնման տվյալների ավտոմատ սեզմենտացիայի խնդիրներում [12]:

Գրականության ցանկ

- [1] **Y. LeCun, Y. Bengio, G. Hinton**, Deep Learning. Nature 521(7553) (2015) 436–444.
- [2] **J. Long, E. Shelhamer, T. Darrell**, Fully Convolutional Networks for Semantic Segmentation. CVPR, 2015.
- [3] **H. Zhao et al.**, Pyramid Scene Parsing Network. CVPR, 2017.
- [4] **L.-C. Chen et al.**, DeepLab: Semantic Image Segmentation with Deep Convolutional Nets. IEEE TPAMI, 2018.
- [5] **J. Redmon et al.**, YOLO9000: Better, Faster, Stronger. CVPR, 2017.
- [6] **A. Kirillov et al.**, Segment Anything. arXiv preprint arXiv:2304.02643, 2023.
- [7] **Y. Guo et al.**, Deep Learning for 3D Point Clouds: A Survey. IEEE TPAMI, 2020.
- [8] **O. Ronneberger, P. Fischer, T. Brox**, U-Net: Convolutional Networks for Biomedical Image Segmentation. MICCAI, 2015.
- [9] **K. He, G. Gkioxari, P. Dollár, R. Girshick**, Mask R-CNN. ICCV, 2017.
- [10] Y. Li et al., Deep Learning in Remote Sensing Image Interpretation: A Review. IEEE GRSM, 2019.
- [11] **S. Minaee et al.**, Image Segmentation Using Deep Learning: A Survey. IEEE TPAMI, 2021.
- [12] **L. Ma et al.**, Deep Learning in Remote Sensing Applications: A Meta-Analysis and Review. ISPRS Journal, 2019.

ОБНАРУЖЕНИЕ И ИНТЕРПРЕТАЦИЯ ОБЪЕКТОВ НА АЭРОКОСМИЧЕСКИХ ИЗОБРАЖЕНИЯХ С ПРИМЕНЕНИЕМ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ

Լուսինե Ներսեսովնա Եգիյան

ООО “Центр геопространственных технологий”, г. Ереван, РА

lusineyeghyan@gmail.com

Аэрокосмические изображения широко используются в различных областях (экологический мониторинг, градостроительство, управление чрезвычайными ситуациями и др.), однако автоматическое обнаружение объектов на изображениях остаётся сложной задачей. В данной статье рассматривается применение методов машинного и глубокого обучения с акцентом на их практическое использование и эффективность. Рассматриваются современные подходы, в частности модели CNN, Mask R-CNN и Segment Anything Model (SAM). Проводится экспериментальный анализ с оценкой точности, применимости и потенциала данных моделей для Республики Армения. Результаты показывают, что методы глубокого обучения обеспечивают высокую точность и уровень автоматизации, одновременно имея определённые ограничения, связанные с шумами и проблемами обобщения.

Ключевые слова: *аэрокосмическое изображение, глубокое обучение, машинное обучение, обнаружение объектов, сегментация, CNN, Mask R-CNN, Segment Anything Model*

OBJECT DETECTION AND INTERPRETATION IN AEROSPACE IMAGES USING DEEP LEARNING METHODS

Lusine Yeghiyan

“Center of Geospatial Technologies” LLC, Yerevan, RA

lusineyeghyan@gmail.com

Aerospace imagery is widely used in various fields (environmental monitoring, urban planning, disaster management, etc.); however, automatic object detection from images remains a challenging task. This article explores the application of machine learning and deep learning methods, focusing on their practical implementation and effectiveness. Modern approaches are considered, particularly CNN, Mask R-CNN, and Segment Anything Model (SAM). An experimental analysis is conducted to evaluate the accuracy, applicability, and potential of these models for the Republic of Armenia. The results show that deep learning methods provide high accuracy and a significant level of automation, while also having certain limitations related to noise and generalization issues.

Keywords: *aerospace imagery, deep learning, machine learning, object detection, segmentation, CNN, Mask R-CNN, Segment Anything Model*

Եղիյան Լուսինե Ներսեսի, (ՀՀ, ք. Երևան) – «Երկրատարածական տեխնոլոգիաների կենտրոն» ՍՊԸ, ԱՏՀ մասնաճյուղ, (+374) 77221835, lusineyeghyan@gmail.com

Егиян Лусине Нерсесовна, (РА, г. Ереван) – ООО «Центр геопространственных технологий», ГИС специалист, (+374)77221835, lusineyeghyan@gmail.com

Yeghiyan Lusine (RA, Yerevan) - “Centre of Geospatial Technologies” LLC, GIS specialist, (+374)77221835, lusineyeghyan@gmail.com

Ներկայացվել է՝ 03.04.2026թ.

Գրախոսվել է՝ 29.04.2026.

Ընդունվել է տպագրության՝ 30.04.2026թ.