«дрейфувати» на різних наборах даних, та обмежену вбудовану пояснюваність окремих виявлених аномалій. Надано рекомендації з конфігурування, перевірки на стійкість та прості інструменти для підтримки впровадження та збереження стабільної продуктивності в умовах, що змінюються.

Ключові слова: *виявлення аномалій; Isolation Forest; неконтрольоване навчання; калібрування порогу; багатовимірний аналіз; масштабована аналітика; інтерпретованість.*

УДК 004.8:004.93 **DOI: 10.18372/2073-4751.83.20547**

Козачук О.
orcid.org/0000-0003-3361-6197

ПОРІВНЯЛЬНИЙ АНАЛІЗ МОНОКУЛЯРНОЇ ОЦІНКИ ГЛИБИНИ СУПУТНИКОВИХ ДВОВИМІРНИХ ЗОБРАЖЕНЬ

Державний університет «Київський авіаційний інститут»

e-mail: oleksandrkozachukk@gmail.com

Вступ

Комерційні *3D*-джерела, такі як *Maxar Precision3D* [1], та візуалізації в *Google Maps/Earth* [2] надають високоякісні моделі рельєфу, але їх використання в академічних експериментах обмежене ліцензіями, доступом та охопленням або ж є складними. Відтворення *3D* форми об’єкта зі зображення меншої розмірності дає змогу отримувати додаткову просторову інформацію про сцену за мінімальних вимог до вхідних даних. Зростання обсягів супутникових зображень і доступності картографічних сервісів породжує запит на просту, відтворювану та зіставну оцінку глибини з одиночних *2D*-знімків. Традиційно завдання глибинного моделювання розв’язують за допомогою фотограмметрії та технологій дистанційного зондування, наприклад, аеро лазерного сканування, *LiDAR*, а також із залученням супутникових моделей висот рельєфу (*DEM/DSM*). Водночас комерційні *3D*-сервіси хоч і надають готові побудови рельєфу, але обмежені вартістю, ліцензійними умовами та територіальним покриттям, що ускладнює масштабоване наукове відтворення для окремих регіонів. Карти висот доповнюють дані про місцевість, однак їхня точність і просторове розрізнення варіюють залежно від сенсора та масштабу збору даних [3]. На рис. 1 представлено класичну *3D*-реконструкцію об'єкта, що попередньо потребує багаторакурсної зйомки з відомими параметрами камери. Помітно, що навіть трьох зображень не вистачає для побудови якісної *3D*-моделі.

Однак поява візуальних трансформерів [4] і великих наборів даних дала змогу оцінювати глибину з одного зображення.

Рис. 1. Класична реконструкція об’єкта спостереження

Попри це, більшість сучасних моделей монокулярної глибини навчалися на загально змістовних, синтетичних наборах даних *MPI-Sintel* [5], *Spring* [6] та інші, що призводить до доменного зсуву під час застосування до супутникових сцен: змінюються геометрія, текстури, масштаб, освітлення; до того ж еталонна метрична глибина для зображень *Google Maps* недоступна для прямої валідації. Тому метою роботи стало перенесення навчання із загальних наборів на домен супутникових знімків.

Запропоноване розв'язання передбачає: відбір моделі, що найкраще адаптується до відповідних знімків; навчання обраної моделі на відкритому супутниковому наборі даних з глибиною; валідацію якості після адаптації та порівняння.

Аналіз досліджень та постановка проблеми

Класичні методи побудови об'єкта чи структури місцевості пов'язані з фотограмметрією та активними сенсорами, зокрема *LiDAR*, а також на цифрові моделі рельєфу, об'єкта, тощо. Ці методи забезпечують високу метричну точність, однак потребують багаторакурсної зйомки з відомою геометрією камери, або ж високотехнологічних сенсорів.

На противагу, монокулярна або ж бінокулярна оцінка глибини зображення стала активно розвиватися упродовж останнього десятиліття. Ранні *CNN*-підходи показали можливість відновлення відносної геометрії сцени без багатокадрової [7], [8], [9]. Подальший прогрес забезпечили великі набори даних і перехресне навчання на неоднорідних даних [10], [11], що суттєво покращило узагальнюваність на незнайомих сценах. Паралельно з'явилися спеціалізовані архітектури й втрати [12], [13], [14], які підсилюють локальну деталізацію та стабілізують масштаб, зсув.

Ключовим етапом став перехід до візуальних трансформерів і *self-supervised* представлень [15], а в подальшому [16], які суттєво підвищили якість ознак. На цій базі побудовано низку сучасних глибинних моделей загального призначення: [17] з хорошим переносом між доменами; [18], [19], що масштабує навчання на великих даних, в яких відсутня попередня анотація і демонструє стабільні межі об'єктів та високу робастність; [20], яке комбінує сильний візуальний енкодер з високою різкістю контурів і практичною інтеграцією в екосистему.

Разом із цим, перенесення монокулярних моделей у супутниковий домен стикається із суттєвим доменним зсувом: інші масштаби та фокусні відстані, анізотропія текстур, повторювані структури забудови, варіативні кути зйомки, атмосферні ефекти.

Мета статті

Метою статті є розробка методів та засобів монокулярної оцінки глибини для супутникових *2D*-зображень з подальшою цільовою адаптацією найкращої моделі до домену ортофото. Результатом дослідження є навчена нейронна мережа, що оцінює глибину вхідного двовимірного знімку. Для досягнення поставленої мети необхідно вирішити наступні завдання:

1. Відбір архітектур нейромереж і уніфікація порівняння: на підставі сучасних публікацій відібрати репрезентативні архітектури *MDE*; реалізувати єдиний стандартизований замір показників для порівняння.
2. Порівняльний аналіз і вибір моделі: провести експерименти на супутникових знімках однакової роздільної здатності, обрати модель, що найкраще адаптується до знімків картографічних сервісів.
3. Цільова адаптація: виконати навчання обраної моделі на доменному наборі ортофото з еталонною глибиною.
4. Валідація після адаптації: провести тестування моделі на відповідних вхідних даних; кількісно оцінити приріст метрик і якісно проілюструвати зміни.

Запропонований науково-методичний апарат придатний для спрощеної *3D*-реконструкції зі *2D*-супутникових знімків і може слугувати універсальним інструментом для дослідників та аналітиків.

Виклад основного матеріалу дослідження

Для порівняльного аналізу було обрано три репрезентативні лінійки моделей монокулярної оцінки глибини, що демонструють різні підходи до побудови ознак та декодування глибини:

1. *Depth Pro* [20]. Трансформерна архітектура з візуальним енкодером

на самонавчених представленнях і спеціалізованим декодером глибини. Вирізняється високою різкістю контурів та стабільною геометрією глибини, що важливо для супутникових сцен із чіткими межами дахів, доріг і фасадів.

2. *Depth-Anything V2* [19]. Сімейство моделей із різною місткістю енкодера (*ViT-S/B/L*). *DA-V2* демонструє хорошу стійкість на відкритих даних і високу швидкодію.
3. *DPT-Large* [11]. Перевірений базовий орієнтир із гібридним декодером; добре переноситься між доменами завдяки попередньому навчанню на гетерогенних наборах даних.

Усім трьом підходам властиво мати трансформерний енкодер, однак вони відрізняються джерелами та обсягом попереднього навчання, організацією декодера й стратегіями нормалізації/масштабування глибини, що впливає на узагальненість у супутниковому домені. За результатами аналізу [20], *DepthPro* лідирує на частині наборів даних і в середньому показує кращі метрики відновлення глибини, див. рис. 2. Водночас жодна з розглянутих моделей не навчалася на супутникових знімках, що зумовлює доменний зсув. Тому необхідно провести оцінку й цільову адаптацію на *2D*-супутникових знімках.

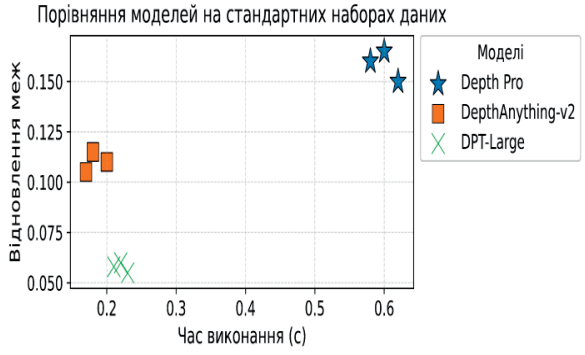


Рис. 2. Порівняння моделей на стандартних наборах даних. Набори даних, що використовувались для тренувань: *NUY, KITTI, ETH3D*

Нехай $\Omega = \{1,\dots,H\}x\{1,\dots,W\}x3$, де $H = W = 728$ - дискретна матриця пікселів вхідного кольорового зображення, що належить супутниковому знімку. Для проведення тестування моделей було сформовано набір даних з N знімків, де $N = 100$.

$$D = \{I_k\}_{k=1}^{N},\ I_k : \Omega \rightarrow \quad [1]$$

де D - початковий набір даних. Якщо вихідні файли не мають точного розміру $728x728x3$ пікселів, вони попередньо масштабуються до $HxWx3$ та лінійно нормуються до [0,1].

В роботі розглядається множина моделей:

$$M = \{DPro, DA\text{-}V2_{s,b,l}, DPT\text{-}Large\} \quad [2]$$

де індекси s, b, l - розміри шарів енкодера моделі, від меншого до більшого. Для оцінки якості відтворення моделі $m \epsilon M$, маємо предиктор:

$$f_{\Theta_m}: \quad [3]$$

де f_{Θ_m} - нейромережева модель. І для k-го знімка прогноз карти глибини:

$$D_{m,k} = f_{\Theta_m}(I_k) \epsilon R^{HxW} \quad [4]$$

Оскільки монокулярна глибина визначена з точністю до афінного перетворення, безеталонні метрики обчислюються на робастно нормованій карті

$$\hat{D}_{m,k} = clip\left(\frac{D_{m,k}(x) - Q_1(D_{m,k})}{Q99(D_{m,k}) - Q_1(D_{m,k}) + \varepsilon}, 0{,}1\right) \quad [5]$$

де $Q_{p\%}(D)$ - p-й перцентиль значень тензора D;

$clip(z, 0{,}1) = min\{1, max\{0, z\}\}$ - округлення значення z діапазону [0,1];

$\varepsilon > 0$ - для запобігання ділення на нуль.

Метрики оцінки якості відтворення глибинної карти рахуються з $\hat{D}_{m,k}(x)$, окрім частки валідних пікселів та ентропії розподілу глибини, вони рахуються з сирих даних карти $D_{m,k}(x)$. Для визначення якості відтвореного результату враховуються наступні метрики:

1. Оцінка частки валідних пікселів:

$$p_{val} = \frac{1}{|\Omega|} \vee \{x \epsilon \Omega : isfinite(D|m,k)\} \vee \quad [6]$$

де $\Omega \vee HxW$ - кількість пікселів, $x \epsilon R$.

2. Ентропія розподілу глибини

$$H_D(\hat{D}_{m,k}) = -\sum_{i=1}^{B} p_i log(p_i + \varepsilon) \quad [7]$$

де p_i - емпіричні ймовірності гістограми $\hat{D}_{m,k}$ з $B = 64$ - розбиття гістограми;

$\varepsilon > 0$ - для запобігання ділення на нуль.

3. Сумарна варіація на піксель:

$$TV_{l_1} = \frac{1}{\Omega \vee \sum_{x \in \Omega}} \quad [8]$$

де δ_x ,δ_y - похідні за Соболем.

4. Середній косинус кута між напрямами градієнтів:

$$C_\alpha = \frac{1}{\Omega \vee \sum_{x \in \Omega}(u_I(x), u_D(x))} \quad [9]$$

де $u_I = \frac{\nabla Y_k}{\vee \nabla Y_k \vee \ _2 + \varepsilon}$;

$u_D = \frac{\nabla \hat{D}}{\vee \nabla \hat{D} \vee \ _2 + \varepsilon}$.

5. Якість відтворення меж, *F1* міра:

$$F_1^{(p)} = \frac{2TP}{2TP + FP + FN}, \text{ для } E_D^{(p)}, E_I \quad [10]$$

де $E_D^{(p)}(x) = 1\{\vee \nabla \hat{D}_{m,k} |\ \ _2 \geq t_p\}$;

$t_p = P_{p\%}(\vee \nabla \hat{D}_{m,k} \vee \ _2)$;

$p \epsilon \{80,90\}$.

6. Медіанна нев'язка площини в патчах, локальна планарність:

Розбиваємо сцену на патчі $64x64$ з кроком розбиття 64. У кожному патчі апроксимуємо $\hat{D}$ площиною $aX + bY + c$ і рахуємо середню абсолютну нев'язку r.

$$R_{pl} = median_{\Pi} \frac{1}{\Pi \vee \sum_{(x,y) \in \Pi} \hat{D}_{m,k}(x,y) - -(Ax + bY + c) \vee} \quad [11]$$

де Π - неперекривні патчі $64x64$;

a, b, c - МНК-оцінка параметрів площини на патчі.

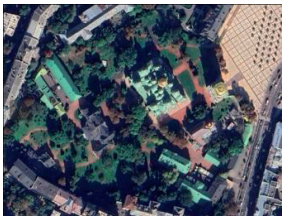
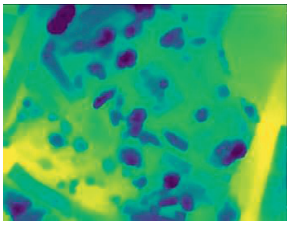
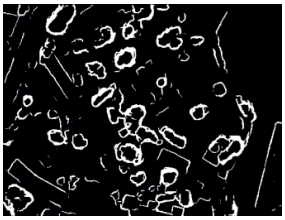
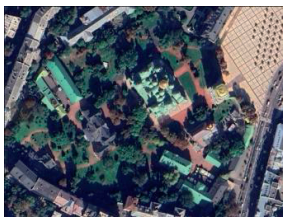
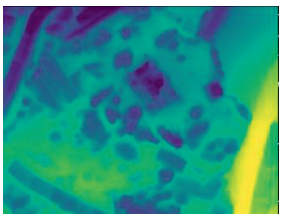
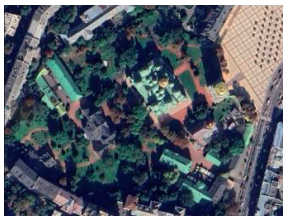
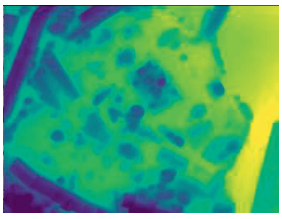
7. Часові характеристики

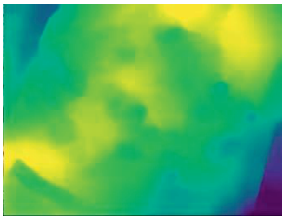
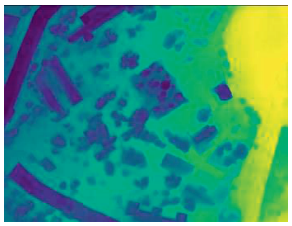
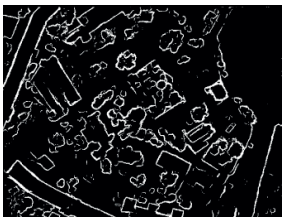
$$t_m = \frac{1}{N} \sum_{k=1}^{N} t_{m,k}, FPS_m = \frac{1}{t_m} \quad [12]$$

де $t_{m,k}$ - час роботи моделі m на зображенні I_k.

Розглянемо візуалізацію отриманих результатів за наведеними моделями m, див. табл 1.

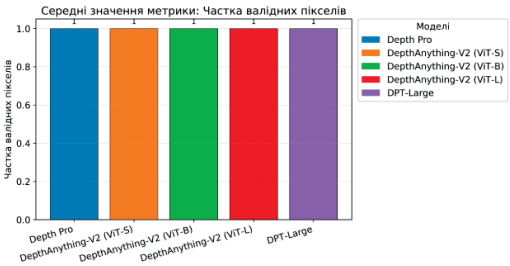
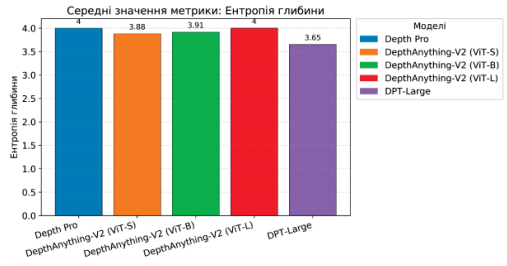
Таблиця 1. Візуалізація результатів для розглянутих моделей

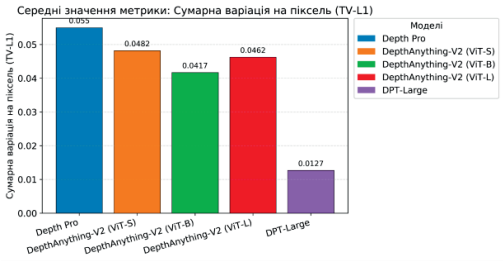
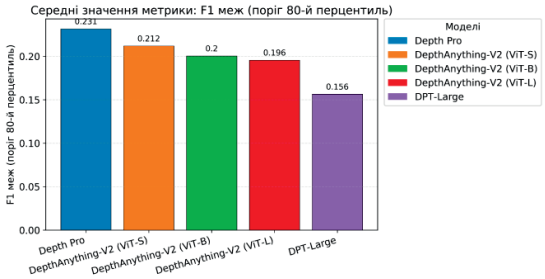
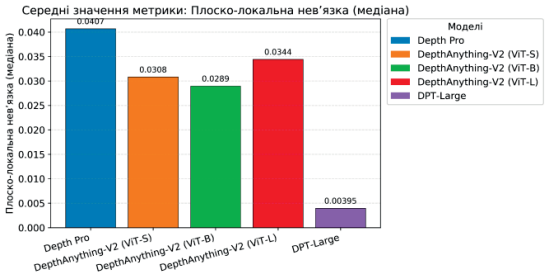
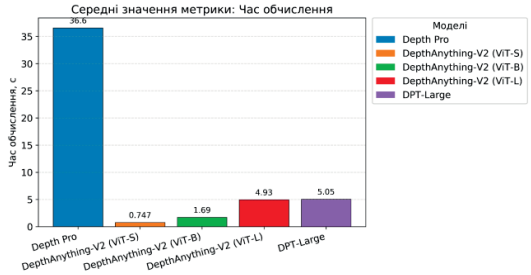
Модель	*Вхідне зображення*	*Оцінена карта глибини*	*Бінарна карта контурів за глибиною*
DA-V2-s			
DA-V2-b			
DA-V2-l			

Модель	*Вхідне зображення*	*Оцінена карта глибини*	*Бінарна карта контурів за глибиною*
DPT-l			
DP			

На основі візуальної оцінки можна сказати, що найкраще адаптовані до супутникових даних моделі: *DA-V2, l* та *DP*. Також наведено оцінку якості реконструкції у таблиці 2.

Таблиця 2. Підсумкові метрики якості

Метрика	Рисунок
Частка валідних пікселів	
Ентропія розподілу глибини	

Метрика	Рисунок
Сумарна варіація на піксель	
Середній косинус кута між напрямами градієнтів	
Якість відтворення меж, *F1* міра	
Медіанна нев’язка площини в патчах, локальна планарність	
Часові характеристики (на *CPU*)	

Відповідно до результатів наведених у таблиці 2, можна сказати, що найкраще з поставленою задачею справилась модель *DepthPro*, хоч і є повільнішою. Тому наступним етапом у дослідженні є донав-

чання даної моделі на супутникових даних *US3D* [21].

Тренування мережі проводилось протягом 20 епох із частковим заморожуванням частини енкодера. Історію навчання та метрики наведено на рисунку 3.

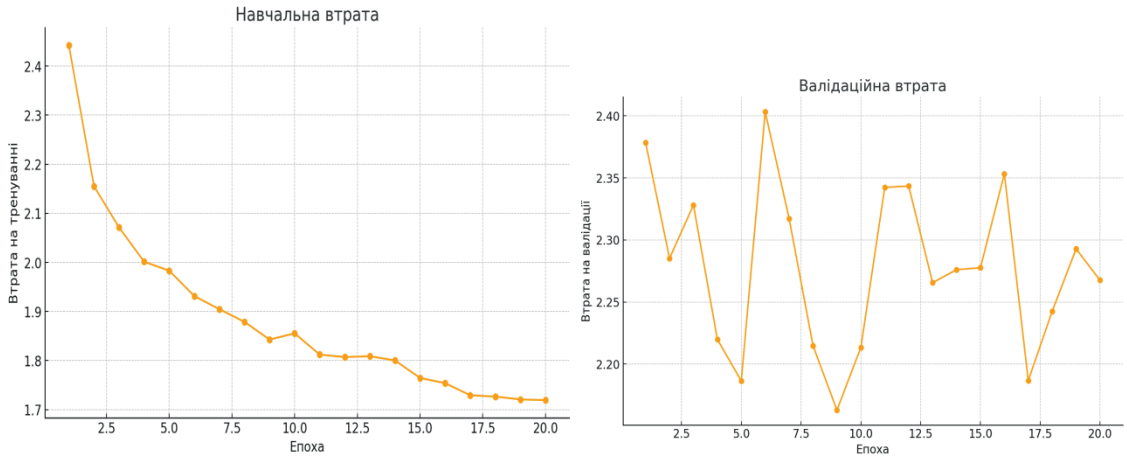
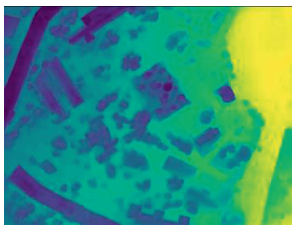
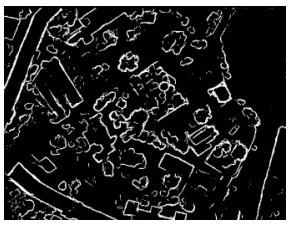
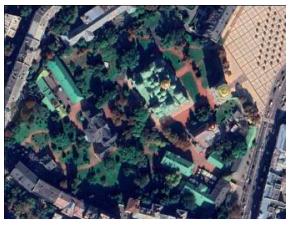
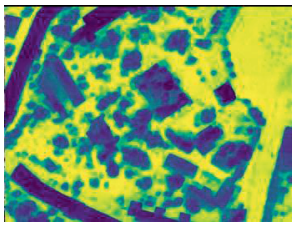


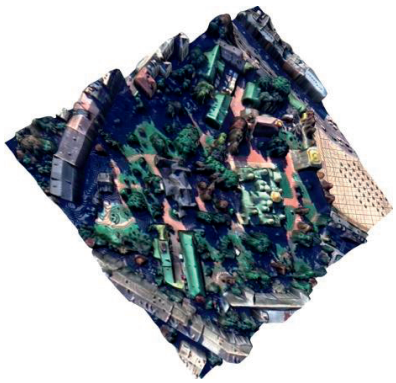
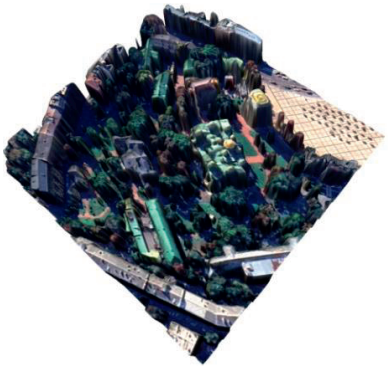
Рис. 3. Історія донавчання моделі

Бачимо, що мережа донавчилась на супутниковому наборі даних, тому наступним кроком є валідація отриманих результатів. Помітно, що після навчання мережа демонструє стабільне покращення, див таблицю 3. Наступним кроком є *3D* ілюстрація відновленої геометрії до навчання та після, див. таблицю 4.

Таблиця 3. Порівняння результатів *DepthPro* до та після донавчання

Модель	Вхідне зображення	Оцінена карта глибини	Бінарна карта контурів за глибиною
DP (до донавчання)			
DP (після донавчання)			

Таблиця 4. 3D порівняння результатів DepthPro до та після донавчання

Модель	Вхідне зображення	Тривимірна проєкція
DP (до донавчання)		
DP (після донавчання)		

Висновки

Таким чином, у цьому дослідженні розроблено та наведено метрики оцінки якості прогнозування моделі, а також проведено порівняльний аналіз сучасних моделей монокулярної оцінки глибини на супутникових сценах. Експерименти показали, що *DepthPro* пропонує найкращу узгодженість глибини з контурною структурою зображення, *Depth-Anything V2* (*ViT-S*) є найшвидшим за збереження прийнятної якості, а *DPT-Large* займає середню позицію, забезпечуючи оптимальний баланс між швидкістю та гладкістю карт глибини.

Удосконалення супутникового домену моделі *DepthPro* на відкритих ортофото з шарами висоти *US3D-JAX* зменшила валідаційну втрату та розширила динамічний діапазон відновленої глибини. У той же час спостерігається компроміс, який поєднує покращення локальної планарності та гладкості, а також незначну втрату різкості дрібних меж. Це демонструє чутливість до складу навчальної вибірки та підкреслює необхідність регуляризації, яка зберігає межі. Відтворювальні, прості в інтеграції та масштабовані.

Література

1. M. Technologies, «Precision3D Data Suite,» Maxar Technologies, [Онлайновий]. Available: https://www.maxar.com/maxar-intelligence/products/precision3d-data-suite. [Дата звернення: 30 Вересня 2025].

2. Google, «Google Maps,» Google, [Онлайновий]. Available: https://mapsplatform.google.com/maps-products/3d-maps/. [Дата звернення: 30 Вересня 2025].

3. P. L. Guth, A. Van Niekerk, C. H. Grohmann, J.-P. Muller, L. Hawker, I. V. Florinsky, D. Gesch, H. I. Reuter, V. Herrera-Cruz, S. Riazanoff, C. López-Vázquez, C. C. Carabajal, C. Albinet та S, «Digital Elevation Models: Terminology and Definitions,» *Remote Sensing,* т. 13, № 18, 2021.

A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G.

Heigold, S. Gelly, J. Uszkoreit та N. Houlsby, «An Image Is Worth 16×16 Words: Transformers for Image Recognition at Scale,» в *International Conference on Learning Representations (ICLR)*, 2021.

4. D. J. Butler, J. Wulff, G. B. Stanley та M. J. Black, «A Naturalistic Open Source Movie for Optical Flow Evaluation,» в *European Conference on Computer Vision (ECCV)*, 2012.

5. L. Mehl, J. Schmalfuss, A. Jahedi, Y. Nalivayko та A. Bruhn, «Spring: A High-Resolution High-Detail Dataset and Benchmark for Scene Flow, Optical Flow and Stereo,» в *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.

6. D. Eigen, C. Puhrsch та R. Fergus, «Depth Map Prediction from a Single Image Using a Multi-Scale Deep Network,» в *Advances in Neural Information Processing Systems (NeurIPS)*, 2014.

7. C. Godard, O. Mac Aodha та G. J. Brostow, «Unsupervised Monocular Depth Estimation with Left-Right Consistency,» в *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

8. C. Godard, O. Mac Aodha, M. Firman та G. J. Brostow, «Digging Into Self-Supervised Monocular Depth Estimation,» в *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019.

9. R. Ranftl, K. Lasinger, D. Hafner, K. Schindler та V. Koltun, «Towards Robust Monocular Depth Estimation: Mixing Datasets for Zero-Shot Cross-Dataset Transfer,» *IEEE Transactions on Pattern Analysis and Machine Intelligence,* т. 44, № 3, p. 1623–1637, 2022.

10. R. Ranftl, A. Bochkovskiy та V. Koltun, «Vision Transformers for Dense Prediction,» в *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021.

11. S. F. Bhat, I. Alhashim та P. Wonka, «AdaBins: Depth Estimation Using Adaptive Bins,» в *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.

12. W. Yin, J. Zhang, O. Wang, S. Niklaus, L. Mai, S. Chen та C. Shen, «Learning to Recover 3D Scene Shape From a Single Image,» в *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.

13. S. F. Bhat, R. Birkl, D. Wofk, P. Wonka та M. Müller, «ZoeDepth: Zero-shot Transfer by Combining Relative and Metric Depth,» 2023. [Онлайновий]. Available: https://doi.org/10.48550/arXiv.2302.12288. [Дата звернення: 30 Вересня 2025].

14. M. Caron, H. Touvron, I. Misra, H. Jégou, J. Mairal, P. Bojanowski та A. Joulin, «Emerging Properties in Self-Supervised Vision Transformers,» в *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021.

15. M. Oquab, T. Darcet, T. Moutakanni, H. Vo, M. Szafraniec, V. Khalidov, P. Fernandez, D. Haziza, F. Massa, A. El-Nouby, N. Ballas, W. Galuba, R. Howes, P.-Y. Huang, S.-W. Li та Misra, «DINOv2: Learning Robust Visual Features without Supervision,» 2023. [Онлайновий]. Available: https://doi.org/10.48550/arXiv.2304.07193. [Дата звернення: 30 Вересня 2025].

16. R. Ranftl, A. Bochkovskiy та V. Koltun, «Vision Transformers for Dense Prediction,» в *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021.

17. L. Yang, B. Kang, Z. Huang, X. Xu, J. Feng та H. Zhao, «Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data,» в *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

18. B. K. Z. H. Z. Z. X. X. J. F. H. Z. Lihe Yang, «Depth Anything V2,» 2024. [Онлайновий]. Available: https://arxiv.org/abs/2406.09414. [Дата звернення: 2025 September 2025].

A. D. H. G. M. S. Y. Z. S. R. R. V. K. Aleksei Bochkovskii, «Sharp Monocular Metric Depth in Less Than a Second,» 2024. [Онлайновий]. Available: https://arxiv.org/abs/2410.02073. [Дата звернення: 30 Вересня 2025].

19. G. /. US3D, «US3D Dataset – Urban Semantic 3D (Jacksonville & Omaha),» [Онлайновий]. Available: https://eod-grss-ieee.com/dataset-detail/bS9HOHBpaEJuMVJ2ZWV4Z01NTGNvZz09. [Дата звернення: 30 Вересня 2025].

Козачук О.О.

ПОРІВНЯЛЬНИЙ АНАЛІЗ МОНОКУЛЯРНОЇ ОЦІНКИ ГЛИБИНИ СУПУТНИКОВИХ ДВОВИМІРНИХ ЗОБРАЖЕНЬ

В умовах стрімкого зростання використання супутникових картографічних сервісів виникає потреба у простих і ефективних методах та засобах побудови 3D (тривимірні) реконструкції місцевості маючи на вхід лише 2D (двовимірні) знімки, які забезпечують відтворюваність і зіставність результатів у прикладних задачах міського аналізу й навігації. Монокулярні моделі прогнозування глибини є інструментом для формалізації таких процесів, оскільки дають змогу відновлювати відносну структуру сцени, подаючи на вхід мінімальний обсяг вхідних даних, що допомагає оцінювати рельєф і висотні контрасти об'єктів. Попри наявність сервісів, що дозволяють отримати тривимірне супутникове зображення, їх обмеженням є те, що вони або комерційними або, що не працюють на території України. Саме тому ставиться завдання у побудові такої моделі, що дозволить вирішити ці проблеми. У роботі розглядаються попередньо навчені моделі та їх прикладне застосування на супутникових зображеннях. Розглянуті моделі були навчені на різних наборах даних, тому проводилась оцінка якості реконструкції на попередньо навчених моделях, відбір за якістю та навчання на наборі даних супутникових зображень. Результати дослідження можуть бути використані для побудови методів навігації літальних апаратів, реконструкції, аналізу місцевості та дозволяють зробити простішу й швидшу 3D реконструкцію довільного об'єкта, хоча поступається за точністю ресурсомістким методам. Варто зауважити, що запропонована методика дозволяє впроваджувати результати на прикладні задачі іншого домену. Вона є відтворюваною, простою у використанні та придатною для подальшого масштабування на більші вибірки і задачі метричного калібрування.

***Ключові слова**: монокулярна оцінка глибини; супутникові зображення; метрики; DepthPro; Depth-Anything V2; DPT-Large; US3D; 3D-реконструкція; Google Maps.*

Kozachuk O.O.

COMPARATIVE ANALYSIS OF MONOCULAR DEPTH ESTIMATION IN SATELLITE TWO-DIMENSIONAL IMAGES

With the rapid growth in the use of satellite mapping services, there is a need for simple and effective methods and means of constructing 3D (three-dimensional) reconstructions of terrain using only 2D (two-dimensional) images as input, which ensure the reproducibility and comparability of results in applied urban analysis and navigation tasks. Monocular depth prediction models are a tool for formalizing such processes, as they allow the relative structure of a scene to be reconstructed with a minimum amount of input data, which helps to assess the relief and height contrasts of objects. Despite the availability of services that allow obtaining three-dimensional satellite images, their limitation is that they are either commercial or do not work in Ukraine. That is why the task is set to build such a model that will solve these problems. The paper considers pre-trained models and their application to satellite images. The models considered were trained on different datasets, so the quality of reconstruction was evaluated on pre-trained models, selected by quality, and trained on a dataset of satellite images. The results of the study can be used to develop methods for aircraft navigation, reconstruction, and terrain analysis, and allow for simpler and faster 3D reconstruction of arbitrary objects, although they are inferior in accuracy to resource-intensive methods. It should be noted that the proposed methodology allows the results to be applied to applied problems in other domains. It is reproducible, easy to use, and suitable for further scaling to larger samples and metric calibration problems.

***Keywords**: monocular depth estimation; satellite images; metrics; DepthPro; Depth-Anything V2; DPT-Large; US3D; 3D reconstruction; Google Maps.*