

DOI: 10.34185/1991-7848.itmm.2023.01.087

**СТРАТЕГІЇ ВИЗНАЧЕННЯ ОПТИМАЛЬНОЇ КІЛЬКОСТІ ВХІДНИХ ДАНИХ
ДЛЯ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ НА СУПУТНИКОВИХ ЗОБРАЖЕННЯХ**

Солдатенко Д.В., Гнатушенко Вік.В.

Український державний університет науки і технологій, м. Дніпро, Україна

При підготовці даних, які використовуються для навчання нейромережі, в деяких випадках можна зіштовхнутися з їх недоліком, або необхідності в симуляції різноманітних вторинних умов, наприклад, погіршення або зміна погоди, якщо ми кажемо про супутникові зображення. Збільшення кількості даних можливо завдяки використанню різноманітних методів аугментації даних, приклади деяких з них приведені нижче, а саме: відображення, поворот, та зміна масштабу [1-3].

Метою дослідження є визначення оптимальної кількості вхідних даних, за допомогою підбору оптимальних методів аугментації даних, які підходять для різних видів використовуваних класів. З цією метою проводиться серія експериментів, щоб дослідити вплив кількості вхідних даних на кілька показників продуктивності, включаючи точність, конвергенцію та узагальнення моделі.

Розпочнемо з одного з найпростіших в реалізації методів аугментації даних - відображення, який полягає в дзеркальному відображенні зображення вздовж вертикальної або горизонтальної осей.

Математично відображення зображення можна описати так: нехай $I(x, y)$ - початкове зображення з розмірами (H, W) , тоді відображене зображення $I'(x', y')$ з розмірами (H, W) можна отримати за допомогою формули (1), для горизонтального відображення:

$$I'(x', y') = I(W - x' - 1, y'), \quad (1)$$

де x' та y' - координати пікселя на відображеному зображенні, W - ширина початкового зображення. Ця формула говорить нам, що кожен піксель на горизонтально відображеному зображенні $I'(x', y')$ можна отримати, обертаючи початкове зображення $I(x, y)$ вздовж вертикальної осі.

Формула (2) для вертикального відображення зображення не сильно відрізняється від горизонтального і буде наступною:

$$I'(x', y') = I(x', H - y' - 1), \quad (2)$$

де (x', y') - нові координати пікселя, (x, y) - початкові координати пікселя, H - висота початкового зображення. Як і у першому випадку, з формулою горизонтального відображення, кожен піксель на вертикально відображеному

зображенні $I'(x',y')$ можна отримати, обертаючи початкове зображення $I(x,y)$ вздовж горизонтальної осі. Приклад використання відображення зображення можна побачити на Рис. 1.

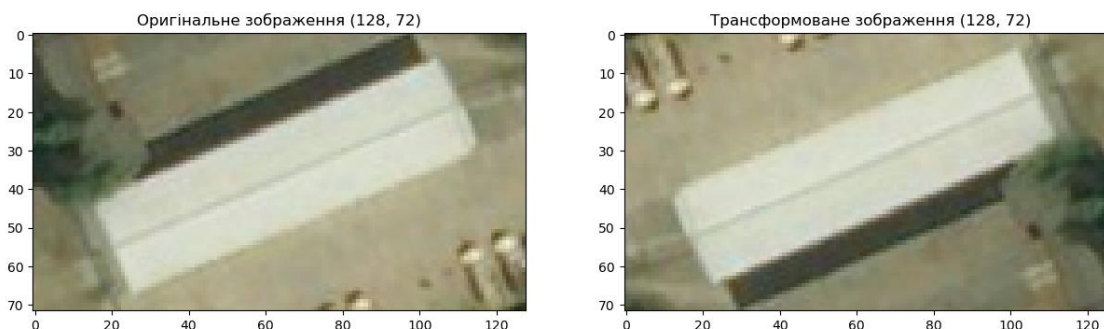


Рисунок 1 - Приклад оригінального та відображеного зображень

Наведені вище формули відображення використовується для зображень з одним каналом (чорно-білих) або зображень з кількома каналами (RGB, наприклад). У випадку зображень з кількома каналами, ці формули потрібно застосовувати до кожного каналу окремо.

Наступний з методів аугментації даних - поворот. Цей метод також використовується для збільшення розміру навчальної вибірки та покращення здатності моделі до узагальнення, шляхом створення нових зображень з різним кутом огляду.

Математично поворот зображення можна описати так: нехай $I(x,y)$ буде зображенням з координатами пікселів (x,y) , де x - номер стовпця, а y - номер рядка. Нехай $I'(x',y')$ буде повернутим зображенням з координатами пікселів (x',y') , де x' - номер стовпця, а y' - номер рядка.

Для обчислення повороту зображення на кут θ за годинниковою стрілкою навколо центру зображення ми можемо використовувати наступні формули (3-4):

$$x' = (x - x_c)\cos\theta - (y - y_c)\sin\theta + x_c \quad (3)$$

$$y' = (x - x_c)\sin\theta + (y - y_c)\cos\theta + y_c \quad (4)$$

де (x_c, y_c) є координатами центру зображення, (x,y) - координати пікселя на зображенні, а θ - кут обертання в радіанах. Формули (3-4) враховують, що пікселі повинні бути перенесені на нові координати (x',y') , щоб отримати нове зображення, обернуте на кут θ . Приклад повороту зображення можна побачити на Рис. 2.

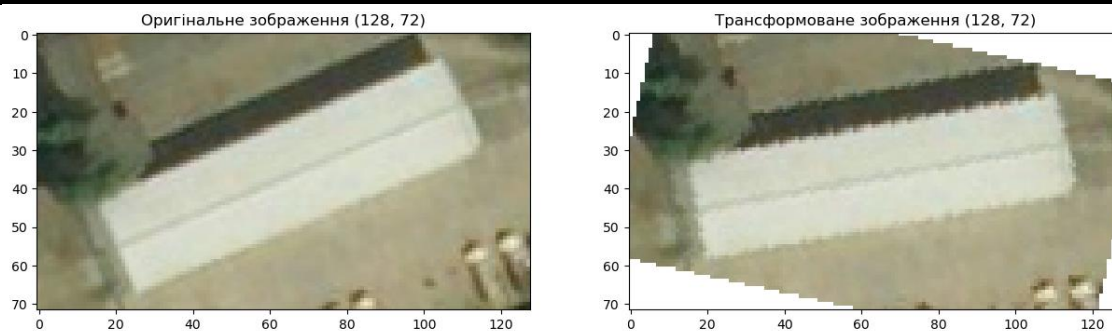


Рисунок 2 - Приклад оригінального та повернутого зображень

Важливо пам'ятати, що після повороту зображення ми можемо втратити частину інформації. Якщо ми повертаємо зображення на більше ніж 90 градусів, ми можемо втратити деякі деталі, такі як текст чи обличчя на зображенні, але у більшості випадків це не стосується супутникових знімків. Тому, при застосуванні методу повороту, ми повинні бути уважні, щоб не втратити важливу інформацію на зображенні.

Зміна масштабу (або ресайзінг) зображення - це метод аугментації даних, що полягає в зміні розміру зображення без зміни його вмісту. Цей метод також може бути корисним для збільшення розміру даних або для адаптації зображення до потрібного розміру, другий варіант більше підходить якщо потрібно зробити вхідні дані однакового розміру.

Математично зміну масштабу зображення вдвічі можна описати так, використовуючи формули білінійної інтерполяції: нехай $I(x, y)$ - початкове зображення з розмірами (H, W) , а $I'(x', y')$ - змінене зображення з розмірами $(H/2, W/2)$. Для того, щоб отримати значення пікселя в точці (x', y') , нам потрібно взяти значення пікселів у чотирьох сусідніх точках (x_1, y_1) , (x_1, y_2) , (x_2, y_1) та (x_2, y_2) , де $(x_1, y_1) = (\lfloor x' \rfloor, \lfloor y' \rfloor)$, $(x_1, y_2) = (\lfloor x' \rfloor, \lceil y' \rceil)$, $(x_2, y_1) = (\lceil x' \rceil, \lfloor y' \rfloor)$ та $(x_2, y_2) = (\lceil x' \rceil, \lceil y' \rceil)$. Значення пікселів у цих точках можна обчислити за допомогою формули (5) лінійної інтерполяції:

$$I'(x', y') = ((x_2 - x') (y_2 - y') I(x_1, y_1) + (x' - x_1) (y_2 - y') I(x_2, y_1) + (x_2 - x') (y' - y_1) I(x_1, y_2) + (x' - x_1) (y' - y_1) I(x_2, y_2)) / ((x_2 - x_1) (y_2 - y_1)), \quad (5)$$

де $\lfloor \cdot \rfloor$ - округлення до меншого цілого, а $\lceil \cdot \rceil$ - округлення до більшого цілого. Приклад зміну масштабу зображення, але зі збереженою висотою, для більшої наочності, можна побачити на Рис. 3.



Рисунок 3 - Приклад оригінального та зображення зі зміненим масштабом

Висновок. На основі проведених експериментах зроблено висновок, що збільшення кількості вхідних даних загалом покращує продуктивність моделі до точки насичення, після якої подальше збільшення не призводить до значних покращень і навіть може призвести до перенавчання нейромережі. У цьому дослідженні пропонується кілька стратегій для визначення цієї точки насичення та пом'якшення наслідків перенавчання. Результати, отримані в цьому дослідженні, можуть допомогти у розробці ефективніших і результативних моделей для розпізнавання супутникових зображень.

Література

1. Cubuk, Ekin D. and Zoph, Barret and Mane, Dandelion and Vasudevan, Vijay and Le, Quoc V. AutoAugment: Learning Augmentation Strategies From Data. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 113-123.
2. Xie, Qizhe and Dai, Zihang and Hovy, Eduard and Luong, Thang and Le, Quoc. Unsupervised Data Augmentation for Consistency Training. Advances in Neural Information Processing Systems, 2020, pp. 6256-6268.
3. Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training gans. In Advances in neural information processing systems, pages 2234–2242, 2016.

INVESTIGATING STRATEGIES FOR DETERMINING THE OPTIMAL INPUT DATA AMOUNT FOR OBJECT RECOGNITION IN SATELLITE IMAGERY

Soldatenko Dmytro, Hnatushenko Viktoria

Abstract. Satellite imagery recognition is a vital application of computer vision, with numerous potential use cases in fields such as disaster management, land use, and urban planning. This study aims to identify the optimal amount of input data required for training the neural network for satellite imagery recognition. To this end, a series of rigorous experiments are conducted to investigate the impact of input data quantity on several performance metrics, including accuracy,

convergence, and model generalization. The findings reveal that an increase in input data quantity generally improves model performance up to a saturation point, beyond which further increases do not lead to significant improvements and may even lead to overfitting. The research proposes several strategies for identifying this saturation point and mitigating the effects of overfitting. The results gained from this research may help in designing more effective models for satellite imagery recognition.

Keywords: neural network, image recognition, satellite imagery, data processing, data augmenting, artificial intelligence.

References

1. Cubuk, Ekin D. and Zoph, Barret and Mane, Dandelion and Vasudevan, Vijay and Le, Quoc V. AutoAugment: Learning Augmentation Strategies From Data. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 113-123.
2. Xie, Qizhe and Dai, Zihang and Hovy, Eduard and Luong, Thang and Le, Quoc. Unsupervised Data Augmentation for Consistency Training. Advances in Neural Information Processing Systems, 2020, pp. 6256-6268.
3. Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training gans. In Advances in neural information processing systems, pages 2234–2242, 2016.