

*В.М. Синєглазов (д.т.н, професор), І.О. Юденко, М.І. Бондаренко
(Національний Авіаційний Університет, Україна)*

Підвищення ефективності обробки зображень за рахунок використання згорткових автоенкодерів для видалення шумів

Дана стаття присвячена використанню згорткового автоенкодера для видалення шумів з зображення. Було досліджено архітектуру згорткового автоенкодера та перевірено його можливості на прикладах зашумлених зображень.

Використання згорткових автоенкодерів для видалення шумів.

На сьогоднішній день велике значення має обробка зображень отриманих за допомогою БПЛА. Складність поставленої проблеми полягає в наявності шумів, які впливають на якість отриманих зображень.

Для видалення шуму із зображень традиційно використовують різні фільтри, серед яких можна відзначити гаусівський, медіанний, білатеральний, нелінійний адаптивний [1; 2]. Останні роки для видалення шуму активно намагаються застосовувати моделі глибокого навчання [3–5], зокрема згорткові автоенкодери [6–8]. Дослідження їх можливостей для розв'язання даної задачі є актуальним і має великий практичний потенціал.

З метою зменшення впливу шумів на зображеннях класично використовують фільтрацію (наприклад, медіанний, гаусівський фільтри, тощо) [1; 2]. При цьому в процесі видалення шумів за допомогою фільтрів важливо дотримуватися балансу між зменшенням шуму та збереженням важливих деталей. Занадто агресивна фільтрація може призвести до втрати деталей і зміни характеру зображення.

В процесі польоту на отримані зображення можуть впливати наступні види шумів [1; 2]:

1. Гаусівський шум – це найпоширеніший вид шуму, який моделюється за допомогою гаусівського розподілу.

2. Шум Пуассона – він моделюється за допомогою розподілу Пуассона. Він імітує ефекти освітлення або електронного шуму на зображеннях.

3. Шум «сіль і перець» – це шум, при якому деякі пікселі приймають найвищі значення кольорової компоненти (сіль) або найнижче (перець). Він імітує втрату або дефекти пікселів, які можуть виникнути в результаті помилок під час збору або передачі зображень.

4. Шум Лапласа – даний тип шуму моделюється за допомогою розподілу Лапласа і дозволяє імітувати різні артефакти на зображеннях.

5. Змішаний шум – це комбінація шумів гаусівського і «сіль та перець».

6. Шум квантування – вид шуму, що виникає при конвертації аналогового сигналу в цифровий формат шляхом квантування.

Перед обробкою цих зображень для певних цілей необхідно видалити шуми, щоб отримати кращі результати. З метою покращення якості зображення пропонується використовувати штучні нейронні мережі.

Для розв'язання поставленої задачі пропонується використати згортковий автоенкодер. Ця модель в основному складається з вхідного шару, шару згортки, шару згортки, шару деконволюції та вихідного рівня. Кілька невеликих і лінійно з'єднаних блоків згорткового автокодувальника (CDA) [9; 10] використовуються на згортковому рівні та намагаються усунути це погіршення якості вхідного зображення. Структура згорткового автокодувальника (CDA) проілюстрована на рисунку 1.

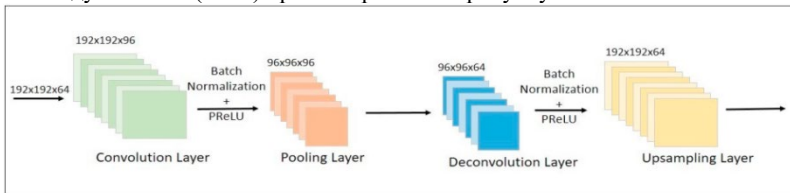


Рис. 1. Структура блоку Convolutional Denoising Autoencoder

Блок CDA включає чотири внутрішні рівні, які називаються шаром згортки, рівнем об'єднання, рівнем деконволюції та рівнем підвищення дискретизації. У цій моделі пакетна нормалізація та функція активації Parametric Rectified Linear Unit (PReLU) розгортаються між рівнем згортки та рівнем об'єднання, а також між рівнем деконволюції та рівнем підвищення дискретизації відповідно. Вихід CDA з'єднаний з його входом. Шари згортки з'єднані тут із шарами підвищення дискретизації прямими з'єднаннями. Це допомагає прискорити процес навчання.

На рисунку 2 показана загальна запропонована архітектура моделі. Він складається з вхідного шару, за яким слідує шар згортки [11]. Для обробки зображення використовується серія згорткових автокодерів із зменшенням шуму. Наприкінці два деконволюційних шари звикли повертати знешумлене зображення. З'єднання пропуску використовуються для зв'язування входу та виходу одного й того ж автокодувальника шумозаглушувача. Тут використовуються десять згорткових блоків автоматичного кодування шумів.

Використання таких пропускових з'єднань у запропонованій нами моделі має дві мети. По-перше, більш глибока гілка згладжує інформацію про кольор більш реально, тоді як деякі дрібні деталі зображення можуть бути втрачені. Через це шару деконволюції може бути важко відновити зображення. Таким чином, пропуск підключення допомагає відновити деякі дрібніші деталі зображення. По-друге, з'єднання пропуску також дають переваги зворотному поширенню, що може значно полегшити навчання глибокої нейронної мережі. Функціональні зображення, передані пропусковими з'єднаннями, передають багато деталей зображення, що сприяє деконволюції для реконструкції покращеного чистого зображення. З глибоких гілок загальна модель може ефективно розпізнавати згладжування кольорів за допомогою з'єднання пропуску.

Оцінювання якості моделей згорткових автоенкодерів проводилося із використанням наступних метрик [4]:

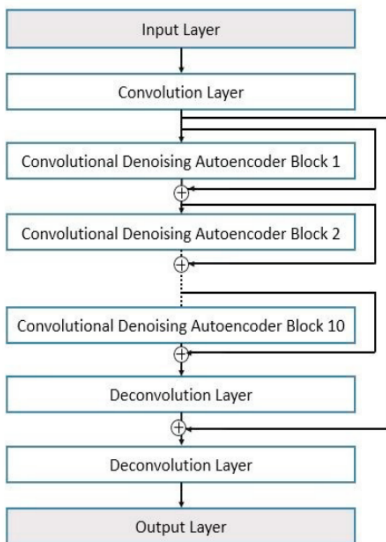


Рис. 2. Візуалізація запропонованої глибокої нейронної архітектури

1. Функція втрат. Було обрано MSE (Mean Squared Error), яка створюється на основі різниці між вихідними даними, відтвореними автоенкодером, і справжніми вхідними даними. Вона вимірює середньоквадратичну відстань між кожним пікселем вихідного та відповідним пікселем вхідного зображення.

2. PSNR (пікове співвідношення сигнал-шум). PSNR вимірює якість відновленого зображення, порівнюючи його з оригінальним зображенням. Метрика обчислюється як відношення максимально можливої потужності сигналу до середньоквадратичної помилки між оригінальним та відновленим зображенням. Вищі значення PSNR вказують на кращу якість відновленого зображення.

3. SSIM (індекс структурної схожості). Ця метрика вимірює схожість структури оригінального та відновленого зображень, враховуючи не лише яскравість і контраст, а й структурні елементи зображення. Значення SSIM лежить в області від -1 до 1, де 1 означає повну схожість між зображеннями.

В якості показників продуктивності для вимірювання продуктивності запропонованої системи використовуються PSNR (пікове відношення сигнал/шум) [12; 13] та SSIM (індекс структурної подібності) [14], які показані в рівняннях 1, 2 і 3. PSNR оцінює пікове значення співвідношення сигнал/шум між вихідним зображенням і реконструйованим зображенням, де високе значення PSNR вказує на хорошу якість реконструйованого зображення. SSIM використовується для вимірювання перцептивної різниці між оригінальним зображенням і реконструйованим/очищеним зображенням, де значення SSIM «1» вказує на те, що вихідне зображення ідентичне реконструйованому зображенню, а «< 1» вказує на те, що оригінальне зображення далеке від реконструйованого зображення. Значення PSNR і SSIM для різних стандартних варіантів проілюстровано в таблиці 1.

$$PSNR = 10 \cdot \left(\frac{MAXIMUM_I^2}{MeanSquareError} \right) \quad (1)$$

Де $MAXIMUM_I$ — найбільше ймовірнісне значення зображення.

$$MeanSquareError = \frac{1}{mn} \sum_{p=0}^{m-1} \sum_{q=0}^{n-1} [I(p, q) - R(p, q)]^2 \quad (2)$$

Для навчання згорткового автоенкодера можна використовувати наступні оптимізатори: середньоквадратичне розповсюдження (RMS Prop),

адаптивний імпульс (Adam), адаптивний імпульс Nesterov (Nadam), адаптивний максимальний об'єднання (Adamax) [15].

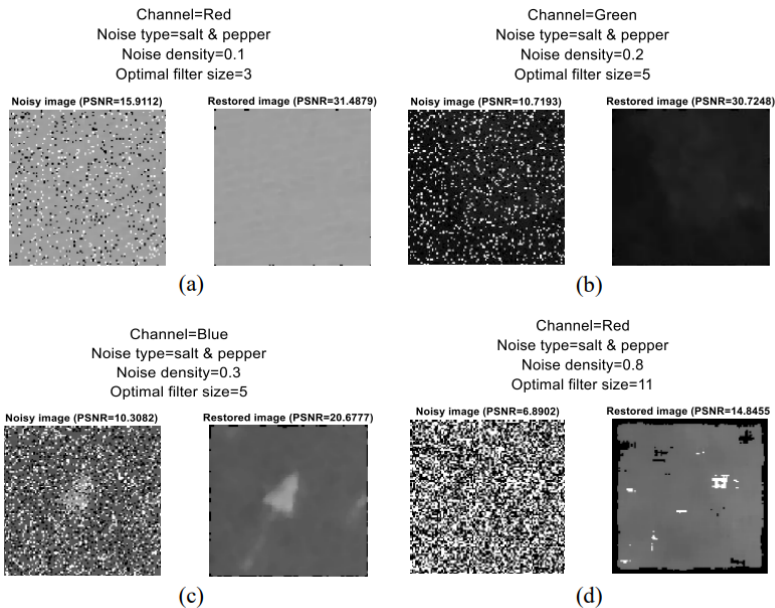


Рис. 4. Зашумлені та відновлені зображення на вибраних зразках для: (а) поверхні даху великого залу, (б) ландшафтної рослинності, (с) поверхні дороги та (д) поверхні даху балкона.

Висновок

У роботі запропоновано підхід для усунення шумів зображень оснований на використанні згорткових автоенкодерів. Проведені дослідження показали, що використання даного підходу дозволяє покращити якість зображення: відновити дуже зашумлене, розмите або навіть пошкоджене зображення без початкового зразка, яке буде майже ідентичне істині.

Список літератури

1. Gonzalez R., Woods R. Digital Image Processing, 4th Ed. – Pearson, 2017. – 1019 p. – Режим доступу: <https://dl.icdst.org/pdfs/files4/01c56e081202b62bd7d3b4f8545775fb.pdf>
2. Zeliski R. Computer Vision: Algorithms and Applications, 2nd Ed. – Springer, 2022. – 1206 p. – Режим доступу: <https://zeliski.org/Book/>

3. Elad M., Kowar B., Vaksman G. Image Denoising: The Deep Learning Revolution and Beyond -- A Survey Paper [Электронный ресурс] // SIAM Journal on Imaging Sciences. – 2023. – Vol. 16(3). – P. 1594-1654. – Режим доступа: <https://arxiv.org/pdf/2301.03362.pdf>
4. Chintan D. Image De-noising Using Deep Learning [Электронный ресурс]. – 2021. – Режим доступа: <https://pub.towardsai.net/image-de-noising-using-deep-learning-1a8334c81f06>
5. Solomon S. Image Denoising using Deep Learning [Электронный ресурс]. – 2021. – Режим доступа: <https://medium.com/analytics-vidhya/image-denoising-using-deep-learning-dc2b19a3fd54>
6. Chaudhary K. Denoising AutoEncoders can reduce noise in images [Электронный ресурс]. – 2023. – Режим доступа: <https://medium.com/game-of-bits/denoising-autoencoders-can-reduce-noise-in-images-5b74753eaf97>
7. Goyal C. Image Denoising using AutoEncoders - A Beginner's Guide to Deep Learning Project [Электронный ресурс]. – 2023. – Режим доступа: <https://www.analyticsvidhya.com/blog/2021/07/image-denoising-using-autoencoders-a-beginners-guide-to-deep-learning-project/>
8. Kuo C. Convolutional Autoencoders for Image Noise Reduction [Электронный ресурс]. – 2019. – Режим доступа: <https://towardsdatascience.com/convolutional-autoencoders-for-image-noise-reduction-32fcc9fc1763>
9. Larrue, Tara, et al. "Denoising Videos with Convolutional Autoencoders." (2018).
10. Jain, Viren, and Sebastian Seung. "Natural image denoising with convolutional networks." Advances in neural information processing systems. 2009.
11. U. Ojha, U. Adhikari and D. K. Singh, "Image annotation using deep learning: A review." In: International Conference on Intelligent Computing and Control (I2C2), Coimbatore, 2017, pp. 1-5
12. Huynh-Thu, Quan, and Mohammed Ghanbari. "Scope of validity of PSNR in image/video quality assessment." Electronics letters. 44.13 (2008): 800-801.
13. Ansari, Mohd & Kurchaniya, Diksha & Dixit, Manish. (2017). "A Comprehensive Analysis of Image Edge Detection Techniques." International Journal of Multimedia and Ubiquitous Engineering. 12. 1-12
14. Hasan, M. and El-Sakka. "Improved BM3D image denoising using SSIM-optimized Wiener filter." M. J Image Video Proc. 2018: 25
15. Yaqub, M., Feng, J., Zia, M., Arshid, K., Jia, K., and Rehman, Z., "State-of-the-art CNN optimizer for brain tumor segmentation in magnetic resonance images", Brain Sciences, 10: 1–19, (2020).