

Одеський національний університет імені І. І. Мечникова
Факультет математики, фізики та інформаційних технологій
Кафедра оптимального керування і економічної кібернетики

Кваліфікаційна робота

на здобуття ступеня вищої освіти «магістр»

«Методи просторової та часової інтерполяції
послідовностей зображень»

«Methods of spatial and temporal interpolation of image
sequences»

Виконав: здобувач денної форми навчання
спеціальності 113 Прикладна математика
Освітня програма «Прикладна математика»
Мороз Дмитро Володимирович

Керівник: доктор фіз.-мат. наук, доц. Кічмаренко О.Д. _____
Рецензент: канд. фіз.-мат. наук, доц. Страхов Є. М.

Рекомендовано до захисту:

Захищено на засіданні ЕК № _____

Протокол засідання кафедри

Протокол № ____ від _____ 2022 р.

№ ____ від _____ 2022 р.

Оцінка _____ / _____ / _____

Завідувач кафедри

Голова ЕК

Одеса — 2022 р.

ЗМІСТ

ВСТУП.....	2
РОЗДІЛ 1 ОГЛЯД ПРЕДМЕТНОЇ ОБЛАСТІ	6
1.1. Підвищення роздільної здатності на основі методів інтерполяції ..	7
1.1.1. Бікубічна інтерполяція.....	8
1.1.2. Метод CONTOUR STENCIL	9
1.1.3. Інтерполяційний поліном Віттекера — Шеннона.....	10
1.1.4. Методи на основі кратно-масштабних апроксимацій (КМА).....	11
1.1.5. Методи штучного інтелекту в задачі інтерполяції зображень....	13
1.2. Кадрова інтерполяція	14
1.2.1. Кадрове усереднення.....	15
1.2.2. Оцінка руху	15
1.2.3. Інтерполяція на основі штучного інтелекту	17
РОЗДІЛ 2 МОДЕЛІ ШТУЧНОГО ІНТЕЛЕКТУ В ЗАДАЧІ ПРОСТОРОВОЇ ІНТЕРПОЛЯЦІЇ ЗОБРАЖЕНЬ	18
2.1. Супер роздільна здатність зображень	18
2.2. GAN алгоритми машинного навчання	19
2.3. Модель SRGAN	20
2.4. Модель ESRGAN та її покращення	21
РОЗДІЛ 3 МОДЕЛІ ШТУЧНОГО ІНТЕЛЕКТУ В ЗАДАЧІ ЧАСОВОЇ ІНТЕРПОЛЯЦІЇ ПОСЛІДОВНОСТІ ЗОБРАЖЕНЬ	23
3.1. Часова інтерполяція послідовності зображень	23
3.2. Модель TOFlow	24
3.3. Модель FLAVR.....	24
3.4. Метод штучного інтелекту для часової інтерполяції послідовності зображень CAIN	25
РОЗДІЛ 4 ПОРІВНЯЛЬНИЙ АНАЛІЗ АЛГОРИТМІВ.....	27
4.1. Результати просторової інтерполяції зображень.....	27
4.2. Результати часової інтерполяції послідовності зображень	30
ВИСНОВКИ	31
Список літератури.....	33

ВСТУП

Візуальна інформація є одним з найголовніших елементів прийняття рішень у сучасному світі. Процес отримання (реєстрації або зйомки) зображень стикається з низкою проблем, починаючи від неоднакової роздільної здатності, варіацій формату, нерівномірного освітлення, спотворень і шумів. Іншими зовнішніми проблемами є варіації в орієнтації та контрасті. У більшості випадків ці проблеми вимагають попередньої обробки зображень – препроцесінга, який, в свою чергу, спонукає до розробки нових та покращення існуючих алгоритмів обробки зображень і відео. Більшість алгоритмів покращення зображень, як правило, є простими і прийнятними, в залежності від сфери застосування зображень чи відео. Відрізняються вони, здебільшого, продуктивністю і якістю результату.

Розвиток інформаційних технологій зумовив важливість цифрових методів обробки зображень. Найпростішими прикладами такої обробки може бути зміна контрасту, яскравості, різкості, згладжування, корекція кольору. Реалістичність зображень відіграє велику роль в прийнятті рішень, де ці зображення застосовуються. Наприклад, це дуже важливо при розробці тренажерів, прив'язуванні до місцевості в реальних умовах, миттєвій ідентифікації об'єктів, тощо. Цифрову обробку зображень можна умовно поділити на два напрямки роботи, таких як:

1. Перетворення зображення для подальшої обробки, розпізнавання, виділення контурів, тощо.
2. Поліпшення якості зображень, поліпшення візуального сприйняття людиною чи алгоритмів розпізнавання.

Існує кілька напрямків для покращення зображень: інтерполяція, розтягування контрастного діапазону, вирівнювання гістограм, стискання динамічного діапазону інтенсивності пікселів. На сьогоднішній день

розроблено різні методи для відновлення зображень, фільтрації, сегментації та відстеження об'єктів на основі диференціальних рівнянь з частинними похідними (PDE). З використанням PDE пропонується інваріантність щодо класичних методів разом із переосмисленням традиційних методів, таких як згортка, фільтрація та морфологічні операції розширення чи ерозії за новою об'єднуючою структурою [1]. Ці методи значно покращили математичне моделювання, зв'язок із фізичними явищами та наближення до геометрії задачі, а також розпізнавання форми, фільтрацію, що зберігає структуру, та сегментацію об'єктів у більш міцній та інтуїтивній структурі. Використання середніх, медіанних, гаусівських та інших фільтрів допомагає зменшити шум шляхом згладжування зображення і, отже, пом'якшення країв. Для вирішення цієї проблеми використовуються методи покращення зображення на основі PDE, які ґрунтуються на припущенні, що інтенсивність освітлення на краях змінюється, як геометричний тепловий потік.

Одним із методів покращення візуального сприйняття є збільшення роздільної здатності. Для збільшення роздільної здатності розроблено широке коло математичних моделей і методів. Найпростішими з них є методи інтерполяції.

Існують різні методи інтерполяції зображень при масштабуванні, найбільш відомі з них – поліноміальні методи інтерполяції. При масштабуванні бажано не погіршувати якість зображення, тобто зберігати діапазон яскравості, форму та межі об'єктів. Інтерполяція супроводжується різними артефактами, такими як розмиття кордонів, блоковість, а при зменшенні зображень можливий елайзинг, який часто проявляється у вигляді хвилі біля меж об'єктів.

Існує низка проблем, для вирішення яких потрібні зображення з високою роздільною здатністю. Висока роздільна здатність означає велику щільність пікселів, а отже, і більш точну деталізацію.

Зі зростанням популярності методів штучного інтелекту, за останні два десятиліття ентузіасти почали використовувати моделі машинного навчання для обробки зображень. Одним з таких алгоритмів є ESRGAN (Enhanced Super Resolution Generative Adversarial Network). В основі цього алгоритму лежать дві моделі: одна генерує зображення, а друга визначає – справжнє воно чи ні. ESRGAN порівнює згенероване зображення з реальним і намагається визначити, чи одне зображення реалістичніше за інше. За словами дослідників, цей тонкий підхід змушує алгоритм генерувати деталі з чіткішими краями і більш реалістичними текстурами.

Іншим методом покращення візуального сприйняття є збільшення кількості кадрів відео.

Існує кілька методів для часової інтерполяції послідовності зображень, з них найпоширенішими є класичні методи. Але останнім часом науковці почали створювати моделі штучного інтелекту для розв'язання цієї задачі. Одним з таких методів є CAIN (Channel Attention Is All You Need). Основною складовою мережі є PixelShuffle – операція реорганізації компонування зображення шляхом об'єднання зображення з декількома різними змінними-перемикачами, та генерації зображень з пониженою вибіркою, що відповідають окремим змінним-перемикачам.

Об'єктом дослідження роботи є методи просторової та часової інтерполяції зображень.

Предметом дослідження є алгоритм штучного інтелекту для підвищення роздільної здатності зображень ESRGAN та метод штучного інтелекту для збільшення кількості кадрів відео CAIN.

Метою роботи є вивчення методології алгоритмів ESRGAN та CAIN, реалізація та покращення моделі, і порівняльний аналіз з іншими алгоритмами для підвищення реалістичності сприйняття інформації в послідовності зображень.

Для досягнення поставленої мети визначені наступні **завдання**:

- огляд методів просторової та часової інтерполяції послідовності зображень;
- огляд методів для пошуку найкращого алгоритму щодо просторової інтерполяції зображень на основі метрики PSNR;
- реалізація та покращення спільної моделі ESRGAN-CAIN та розробка ефективного алгоритму;
- порівняльний аналіз алгоритму ESRGAN з відомими алгоритмами інтерполяції зображень;
- психо-візуальний аналіз результатів алгоритму CAIN.

Апробація результатів. За результатами дослідження опубліковані тези [27] на XI Міжнародну науково-практичної конференцію (ПІКТ – 2022), Кічмаренко О.Д., Мороз Д.В. «Аналіз алгоритмів просторової інтерполяції послідовностей зображень» // Проблеми інформатики та комп’ютерної техніки – м. Чернівці, 10–13 лист. 2022. Чернівці: Черн. нац. ун-т, 2022. С. 28 – 30.

ВИСНОВКИ

Розглянуто покращену модель ESRGAN для покращення зображень з метою збільшення їх роздільної здатності з заміненим дискримінатором на U-Net, зі спектральною нормалізацією. Цей метод показав кращі результати, в порівнянні з іншими моделями, з точки зору якості вихідного зображення. Але, дана модель може бути в подальшому покращена з точки зору обчислювальної складності для роботи в режимі реального часу на менш потужних комп'ютерах. Також на сьогодні не існує простих способів пошуку оптимальних значень гіперпараметрів навчання, що є критичним для часу навчання і якості результату.

Найкраще, з психо-візуальної точки зору, модель справляється з покращенням синтетичних, тобто, нереальних зображень. Тому можна зробити припущення, що доцільно модифікувати деякі блоки оригінальної моделі ESRGAN. Для цього в подальшому пропонується додавання генератора шумового стохастичного процесу на вході мережі генератора. Очікується, що отримані зображення будуть мати більш реалістичні текстурні складові.

Іншою проблемою, яку було розв'язано з допомогою методу CAIN, є якість об'єктів зареєстрованих зображень/відео і плавність між перемиканням кадрів, яка є основою чіткого і швидкого розпізнавання об'єкту психо-візуальною системою оператора-навідника.

Однак, оцінити результати часової інтерполяції послідовності зображень обчислювальним шляхом досить складно, тому віддано перевагу психо-візуальному методу оцінки.

Ще однією проблемою методу CAIN є швидкість генерації проміжних кадрів, яка експоненційно зростає в залежності від роздільної здатності послідовності зображень.

Зображення, в залежності від конкретної задачі, мають топологію, яка має враховувати властивості об'єкту – просторово-часову орієнтацію, масштаб, геометричні параметри об'єкта, включно з координатами, кутовим положенням, лінійними розмірами, відстанню та іншими. А це ускладнює розв'язання подібних задач ще більше при інтерпретації динамічних візуальних сцен через високу розмірність простору ознак та наявність геометричних перетворень над об'єктом. Тому в подібних випадках застосовується стиск простору ознак методом витягування інтегральних та інваріантних до геометричних перетворень параметрів зображень. Це важливо при розпізнаванні, наприклад, не тільки друкованих і рукописних букв, а і профілів ракет, літаків і танків та багатьох інших об'єктів.

Список літератури

1. G. Aubert. and P. Kornprobst, “Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations 2nd ed. Berlin, Germany: Springer-Verlag, 2006.
2. Izonin I., Tkachenko R., Peleshko D., Vatiuk, D. Метод збільшення роздільної здатності зображень на основі штучних нейронних мереж.// Вісник Львівського державного університету безпеки життєдіяльності, 11, 2018 — с.47-56.
3. Koo Yido, Wonchan Kim. An image resolution enhancing technique using adaptive sub-pixel interpolation method.// IEEE Transactions on Consumer Electronics.45(1), 1999; с.118-123
4. Paul Cockshott W, Sumitha L. Balasuriya, Irwan Prasetya Gunawan, Paul Siebert J. Image enhancement using vector quantisation-based interpolation.// Proceedings of the British Machine Vision Conference, University of Warwick, 2007.
5. Jagadeesh P. Image resolution enhancement based on edge directed interpolation using dual tree-complex wavelet.// Proceedings of International Conference on Recent Trends in Information Technology, 3-5 June Chennai, Tamil Nada.: 2011; с.759-763
6. R. Keys. Cubic convolution interpolation for digital image processi- ng.// IEEE Transactions on Acoustics, Speech, and Signal Processing. 29 (6): 1153–1160.
7. P. Getreuer. Image Interpolation with Contour Stencils — 2011.
8. A. J. Jerri. The shannon sampling theorem—Its various extensions and applications: A tutorial review// Proc. IEEE, vol. 65, no. 11, pp. 1565-1596, Nov. 1977.
9. Moroz, V. V. Image resampling using generalized series WhittakerKotelnikov-Shannon.// Eastern-European Journal of Enterprise Technologies, 2(10(62)), 63–66. — 2013

10. Mallat, S. A theory for multiresolution signal decomposition: the wavelet representation// *IEEE Pattern Anal. and Machine Intell.*, vol. 11, no. 7, pp. 674–693.— 1989
11. Daubechies, Ingrid. *Ten lectures on wavelets.*// Philadelphia, PA :Society for Industrial and Applied Mathematics — 1992
12. Chao Dong, Chen Change Loy, Kaiming He, Xiaoou Tang. *Image SuperResolution Using Deep Convolutional Networks.*// *IEEE Transactions on Pattern Analysis and Machine Intelligence* — 2016.
13. He Kaiming, Zhang Xiangyu, Ren Shaoqing, Sun, Jian. *Deep Residual Learning for Image Recognition.*// *Proc. Computer Vision and Pattern Recognition (CVPR), IEEE.*,— 2020.
14. Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, Kyoung Mu Lee. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2017*, pp. 136-144.
15. Honggang Chen, Xiaohai He, Linbo Qing, Yuanyuan Wu, Chao Ren, Ce Zhu. *Real-World Single Image Super-Resolution: A Brief Review*,— 2021.
16. Schawinski Kevin, Zhang Ce, Zhang Hantian, Fowler Lucas, Santhanam Gokula Krishnan. *Generative Adversarial Networks recover features in astrophysical images of galaxies beyond the deconvolution limit.*// *Monthly Notices of the Royal Astronomical Society: Letters.* 467 (1): L110–L114— 2017
17. Wang Zheng Wei, She Qi, Ward Tomas E. *Generative Adversarial Networks in Computer Vision: A Survey and Taxonomy.*— Association for Computing Machinery New York NY United States — 2019
18. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. *Deep Residual Learning for Image Recognition*, 2015, [online]
19. Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, Wenzhe Shi. *Photo-Realistic Single Image*

- SuperResolution Using a Generative Adversarial Network//EEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- 20.K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In International Conference on Learning Representations (ICLR), 2015.
 - 21.Huang, Gao; Liu, Zhuang; Van Der Maaten, Laurens; Weinberger, Kilian Q. (2017). Densely Connected Convolutional Networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI: IEEE. pp. 2261–2269
 22. Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Chen Change Loy, Yu Qiao, Xiaoou Tang. ESRGAN: Enhanced SuperResolution Generative Adversarial Networks//ECCV 2018 workshop, 2018.
 23. Edgar Schonfeld, Bernt Schiele, Anna Khoreva. A U-Net Based Discriminator for Generative Adversarial Networks// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 8207-8216
 24. Tianfan Xue, Baian Chen, Jiajun Wu, Donglai Wei, William T. Freeman. Video Enhancement with Task-Oriented Flow//International Journal of Computer Vision (IJCV), 127(8):1106-1125, 2019017
 25. Tarun Kalluri, Deepak Pathak, Manmohan Chandraker, Du Tran. FLAVR: Flow-Agnostic Video Representations for Fast Frame Interpolation; [online]
 26. Choi, M., Kim, H., Han, B., Xu, N., Lee, K. M. Channel Attention Is All YouNeed for Video Frame Interpolation.// Proceedings of the AAAI Conference on Artificial Intelligence, 34(07), 10663-10671.
 27. Кічмаренко О.Д., Мороз Д.В.. «Аналіз алгоритмів просторової інтерполяції послідовностей зображень» // Проблеми інформатики та комп'ютерної техніки – м. Чернівці, 10 – 13 лист. 2022. Чернівці: Черн. нац. ун-т, 2022. С. 28 – 30.