

Є.С. Сулема, Б.С. Топчієв

ІНТЕЛЕКТУАЛЬНА КОЛОРИЗАЦІЯ ЗОБРАЖЕНЬ ЗА ДОПОМОГОЮ ГЕНЕРАТИВНИХ ЗМАГАЛЬНИХ МЕРЕЖ

Анотація. Статтю присвячено застосуванню генеративних змагальних нейронних мереж на прикладі розв'язку задачі інтелектуальної колоризації зображень. В основі розробленої програмної системи для інтелектуальної колоризації лежить робота двох конкуруючих згорткових нейронних мереж: мережі-генератора та мережі-дискримінатора. На етапі підготовки даних пропонується виконувати перетворення зображень з колірної моделі RGB у модель YUV або LAB, що дозволяє отримати кращий результат колоризації. Також у статті розглянуто можливість переходу від оброблення зображень до оброблення мультимодальних даних та запропоновано подавати вхідні мультимодальні дані про об'єкти на зображенні у вигляді мультимодальної багаторівневої онтологічної моделі.

Ключові слова: нейронні мережі, інтелектуальна колоризація, сегментація зображень, оброблення зображень, оброблення мультимодальних даних.

Вступ. Інтелектуальна колоризація є процесом зміни кольорів певних об'єктів на зображенні та їх елементів за допомогою математичних алгоритмів. Галузь застосування інтелектуальної колоризації включає широке коло задач, пов'язаних з обробленням архівних зображень, художнім обробленням фотографій та відео, створенням реалістичних візуальних моделей тощо. Існують два основних підходи до інтелектуальної колоризації. У першому підході використовуються стандартні методи, такі як накладання додаткових шарів на виділені області на зображенні, зміна показників яскравості, насиченості та відтінків [1]. Ці методи потребують від користувача активної участі та уміння локалізувати об'єкти, які необхідно перефарбувати, переналаштовувати тіні, яскравість та експозицію на зображенні, щоб перетворене зображення виглядало природно. Другий підхід ґрунтується на технології машинного навчання. Так, у роботі [2] запропоновано використовувати методи глибинного навчання для колоризації зображень та методи двосторонньої

фільтрації для їх постоброблення. В роботі [3] запропонований метод колоризації, який ґрунтується на використанні глибинних згорткових генеративних змагальних мереж, що передбачають як навчання з вчителем, так і навчання без вчителя.

Оскільки другий підхід до вирішення задачі інтелектуальної колоризації дозволяє виконувати автоматизоване оброблення зображень, то він є більш перспективним, а задача розроблення нових та вдосконалення існуючих методів такого класу є актуальною. У цій статті представлено спосіб колоризації зображень, який ґрунтується на використанні генеративних змагальних нейронних мереж.

Постановка задачі. У цьому дослідженні вирішувалась задача проектування системи нейронних мереж, яка виконує колоризацію зображень згідно із заданою користувачем схемою розфарбовування об'єктів на вхідному зображенні.

Архітектура системи. В основі розробленої програмної системи для інтелектуальної колоризації лежить робота двох конкуруючих згорткових нейронних мереж [4]. Перша мережа є мережею-генератором, вона виконує оброблення вхідного зображення відповідно до мети оброблення, яка задається користувачем. Задача цієї мережі полягає у тому, щоб згенерувати дані, максимально подібні до оригінальних даних, що використовуються під час навчання. Друга мережа є мережею-дискримінатором, вона оцінює якість результату шляхом визначення різниці між зображенням-зразком, що задає бажаний результат оброблення, та зображенням, яке генерує перша мережа. Такий підхід забезпечує високу якість навчання обох мереж та, відповідно, високу якість отриманих результатів.

Окрім власно зображення, що потребує колоризації, вхідними даними для розробленої програмної системи є маска, яка задає об'єкти на зображенні, що потребують колоризації, та бажані кольори. Вихідними даними є зображення, яке є копією вхідного зображення, колоризованого відповідно до заданої маски.

Основною задачею, що виконується під час оброблення вхідного зображення, є його сегментація. Для вирішення цієї задачі можуть бути

застосовані декілька видів архітектур згорткових нейронних мереж, з яких основними є U-Net [5] та FPN [6]. В цьому дослідженні використовується мережа U-Net, оскільки вона дозволяє отримати візуально якісніший результат. Архітектура мережі U-Net складається з шляху звуження для визначення основних особливостей зображення та симетричного йому шляху розширення, який дозволяє здійснити точну локалізацію.

Навчання мережі виконується методом стохастичного градієнтного спуску [5]. При цьому можливе навчання на невеликій кількості зображень. Основною перевагою цієї мережі є швидкість оброблення даних.

Мережа-дискримінатор, що реалізує бінарний класифікатор, застосовується для оцінювання якості роботи мережі-генератора.

Процес тренування зводиться до таких основних кроків:

1. На вхід системи подається зображення.
2. Зображення оброблюється мережею-генератором, на виході якої отримується нове, відредаговане зображення.
3. На вхід мережі-дискримінатора подається зображення з мережі-генератора та реальне зображення.
4. Ці два зображення по черзі аналізуються мережею-дискримінатором. Для реального зображення очікуваним результатом є True, для згенерованого – False. Якщо мережа-дискримінатор видає невірний результат, то вона починає проходити додаткові цикли навчання. Якщо результат є вірним, то додаткові цикли навчання проходить мережа-генератор.

Розроблена програмна систем реалізує архітектуру з п'ятьма шарами, з яких чотири шари є шарами згортки та один шар використовується для об'єднання. У майбутньому планується збільшити кількості шарів до семи.

Експериментальна частина дослідження. Для виконання навчання системи нейронних мереж фундаментально важливим є правильний підбір даних. В першу чергу необхідно забезпечити наявність великого об'єму даних. Тому для навчання було обрано декілька великих наборів даних, а саме, ImageNet, який містить близько 1 мільйону зображень [7], та COCO, що складається з приблизно 200 тисяч зображень [8]. Другим важ-

ливим аспектом є релевантність даних. У цій задачі даними мають виступати фотографії розміром 512 на 512 пікселів чи більше, які мають містити певні довільні об'єкти, що будуть сегментуватись за допомогою нейронної мережі.

Вхідними даними для розробленої програмної системи є зображення, яке потребує колоризації, та маска, яка задає бажані кольори та схему їх застосування до зображення.

В процесі навчання було виявлено, що колірна модель RGB [9, 10], яка за замовчуванням використовується для подання графічних даних зображень, не дозволяє отримати вичерпну інформацію про кольори цілих об'єктів, а скоріше є корисною при роботі з окремо узятими пікселями. Це пояснюється тим, що у колірній моделі RGB відтінок кольору визначається комбінацією трьох колірних компонент, кожна з яких визначає інтенсивність певного базового кольору (червоного, зеленого та синього). Проте, доцільним є подання відтінку кольору з явним виділенням компоненти, що визначає загальну яскравість зображення у відтінках сірого кольору (greyscale image). Тому при підготовці даних для тренування було виконано перетворення колірної моделі RGB у колірну модель YUV [9, 10] та колірну модель LAB [10].

Колірна модель YUV визначає відтінок кольору простір як комбінацію яскравісної компоненти Y та двох кольорорізницевих компонент U та V. Переведення зображення з простору RGB у простір YUV відбувається наступним чином:

$$Y = 0.299 * R + 0.587 * G + 0.114 * B \quad (1)$$

$$U = 0.492 * (B - Y) \quad (2)$$

$$V = 0.877 * (R - Y) \quad (3)$$

Колірна модель CIELAB визначає відтінок кольору аналогічно моделі YUV: компонента L визначає яскравість, компонента A визначає колірну складову у діапазоні від зеленого до червоного та компонента B визначає колірну складову у діапазоні від синього до жовтого.

Проведені у дослідженні експерименти показали, що обидві колірні моделі можуть бути застосовані для перетворення кольору з моделі RGB. Під час тестування системи було проведено навчання мереж на на-

борі, що містив 100 тисяч зображень розміром 512 x 512 пікселів. При цьому тренування проводилось на наборах даних ImageNet та COCO як окремо, так і разом. На рис. 1. наведено приклад результату роботи розробленої програмної системи.



Рисунок 1 - Приклад результату колоризації зображення

Отримані результати колоризації зображень можуть бути оцінені суб'єктивно, оскільки об'єктивні способи оцінювання реалістичності колоризованих зображень відсутні. Суб'єктивна оцінка дозволяє стверджувати, що запропонований спосіб інтелектуальної колоризації дозволяє отримати результат належної якості. При цьому оброблення вхідних даних відбувається у режимі, близькому до реального часу.

Перехід до мультимодальних даних. Задача інтелектуальної колоризації зображень є лише однією із широкого кола задач, пов'язаних із застосуванням нейронних мереж. При вирішенні цієї задачі оброблюються дані лише однієї модальності – графічні дані. Проте, наразі активно ведуться дослідження [11, 12], які передбачають оброблення нейронними мережами даних кількох модальностей, що дозволяє, наприклад, підвищити якість класифікації об'єктів спостереження.

На етапі підготовки даних перед їх аналізом за допомогою нейронних мереж мультимодальні дані доцільно представляти у вигляді мультимодальної багаторівневої онтологічної моделі (рис. 2) [13].

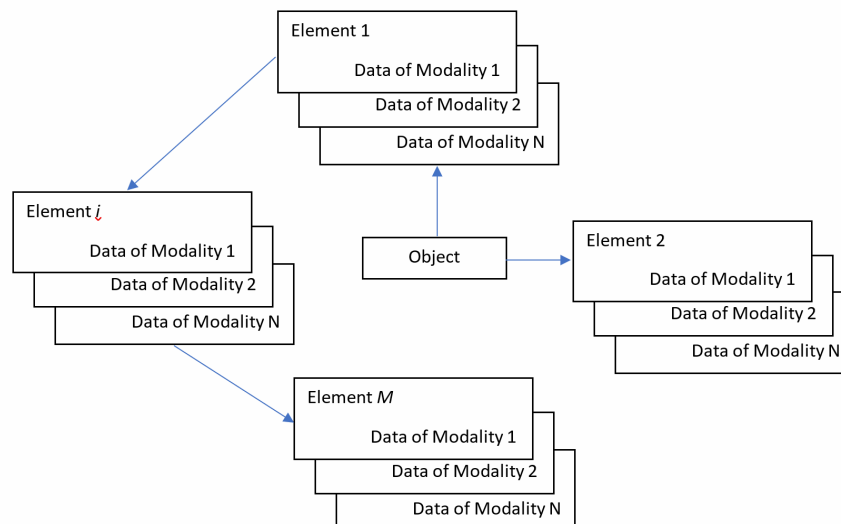


Рисунок 2 - Мультиmodalна багаторівнева онтологічна модель [13]

Багаторівнева онтологічна модель включає макро- та мікро-рівень подання даних. На макро-рівні відбувається декомпозиція об'єктів на зображенні із утворенням семантичної ієрархії між ними. На мікро-рівні кожен об'єкт визначається наборами даних всіх модальностей, які є релевантними для цього об'єкту.

Такий спосіб подання мультиmodalних даних дозволить спростити подальше їх оброблення нейронною мережею за рахунок ієрархічної структури опису об'єктів.

Висновки. Запропонований спосіб колоризації зображень дозволяє досягти значного зниження часу оброблення вхідних даних. Спроектвана архітектура виконує оброблення одного зображення у режимі, близькому до реального часу, в той час, як способи колоризації на основі оптимізації [14] вимагають для отримання аналогічних результатів близько 30-45 секунд. На відміну від інших підходів, колоризація зображень на основі генеративних змагальних мереж не потребує значної участі користувача у обробленні зображень та має потенціал для вдосконалення архітектури програмної системи оброблення зображень. У дослідженні, представленому у цій статті, було спроектовано базову архітектуру програмної системи інтелектуальної колоризації та проведено тестування. У подальшому планується удосконалити архітектуру та набори даних для поліпшення якості колоризації та часових характеристик системи ней-

ронних мереж.

Крім того, у статті розглянуто можливість переходу від оброблення даних однієї модальності (зображень) до оброблення мультимодальних даних та запропоновано використовувати мультимодальну багаторівневу онтологічну модель на етапі підготовки даних перед їх аналізом за допомогою нейронних мереж, що дозволить спростити роботу з даними.

ЛИТЕРАТУРА / ЛІТЕРАТУРА

1. “Adobe Photoshop: Match, replace, and mix colors”. – 2017. – Режим доступу : <https://helpx.adobe.com/photoshop/using/matching-replacing-mixing-colors.html>
2. Z. Cheng, Q. Yang, B. Sheng, “Deep colorization” // Proceedings of IEEE International Conference on Computer Vision. – 2015. – pp. 415-423.
3. Junghoon Seo, Taewon Yoon, and Jinwoo Kim, “One-to-one Example-based Automatic Image Coloring Using Deep Convolutional Generative Adversarial Network” // Journal of Advances in Information Technology. – 2017. – Vol. 8, No. 2. – pp. 80-85.
4. Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, “Generative Adversarial Networks” // Advances in Neural Information Processing Systems. – 2014. – pp. 1-9.
5. Olaf Ronneberger, Philipp Fischer, Thomas Brox. “U-Net: Convolutional Networks for Biomedical Image Segmentation” // Medical Image Computing and Computer-Assisted Intervention. – 2015. – pp. 234-241.
6. Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, Serge Belongie, “Feature Pyramid Networks for Object Detection” // Computer Vision and Pattern Recognition. – 2017. – pp. 2117-2125.
7. ImageNet: Summary and Statistics. – 2010. – Режим доступу : <http://image-net.org/about-stats>
8. COCO Dataset. – 2019. – Режим доступу : <http://cocodataset.org>
9. Michal Podpora, Grzegorz Paweł Korbias, Aleksandra Kawala-Janik, “YUV vs RGB – Choosing a Color Space for Human-Machine Interaction” // Computer Science and Information Systems. – 2014. – pp. 29-34.
10. Noor A. Ibraheem, Mokhtar M. Hasan, Rafiqul Z. Khan, Pramod K. Mishra, “Understanding Color Models: A Review” // ARPN Journal of Science and Technology. – 2012. – Vol. 2, No. 3. – pp. 265-275.

11. Ivanov I., Sopov E., Panfilov I., “Comprehensive approach for solving multimodal data analysis problems based on integration of evolutionary, neural and deep neural network algorithms” // IOP Conf. Series: Materials Science and Engineering. – 2018. – pp. 1-10.
12. Shishir Kumar, Prabhat Mahanti, Su-Jing Wang, "Intelligent computational techniques for multimodal data" // Multimedia Tools and Applications. – 2019. – Vol. 78. – pp. 23809–23814.
13. Ye. Sulema, I. Dychka, O. Sulema, “Multimodal Data Representation Models for Virtual, Remote, and Mixed Laboratories Development”//Lecture Notes in Networks and Systems.–Springer Cham, 2018. – Vol. 47, pp. 559-569.
14. Anat Levin, Dani Lischinski, Yair Weiss. “Colorization using Optimization” // Proceedings of ACM SIGGRAPH. – 2004. – pp. 689–694.

REFERENCES

1. “Adobe Photoshop: Match, replace, and mix colors”, 2017, <https://helpx.adobe.com/photoshop/using/matching-replacing-mixing-colors.html>
2. Z. Cheng, Q. Yang, B. Sheng, “Deep colorization”, Proceedings of IEEE International Conference on Computer Vision, 2015, pp. 415-423
3. Junghoon Seo, Taewon Yoon, and Jinwoo Kim, “One-to-one Example-based Automatic Image Coloring Using Deep Convolutional Generative Adversarial Network”, Journal of Advances in Information Technology, 2017, Vol. 8, No. 2, pp. 80-85.
4. Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, “Generative Adversarial Networks”, Advances in Neural Information Processing Systems, 2014, pp.1-9.
5. Olaf Ronneberger, Philipp Fischer, Thomas Brox. ”U-Net: Convolutional Networks for Biomedical Image Segmentation”, Medical Image Computing and Computer-Assisted Intervention, 2015, pp. 234-241.
6. Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, Serge Belongie, “Feature Pyramid Networks for Object Detection”, Computer Vision and Pattern Recognition, 2017, pp. 2117-2125.
7. ImageNet: Summary and Statistics, 2010, <http://image-net.org/about-stats>
8. COCO Dataset, 2019, <http://cocodataset.org>
9. Michal Podpora, Grzegorz Paweł Korbas, Aleksandra Kawala-Janik, “YUV vs RGB – Choosing a Color Space for Human-Machine Interaction”, Computer Science and Information Systems, 2014, pp. 29-34.

10. Noor A. Ibraheem, Mokhtar M. Hasan, Rafiqul Z. Khan, Pramod K. Mishra, "Understanding Color Models: A Review", ARPN Journal of Science and Technology, 2012, Vol. 2, No. 3, pp. 265-275.
11. Ivanov I., Sopov E., Panfilov I., "Comprehensive approach for solving multimodal data analysis problems based on integration of evolutionary, neural and deep neural network algorithms", IOP Conf. Series: Materials Science and Engineering, 2018, pp. 1-10.
12. Shishir Kumar, Prabhat Mahanti, Su-Jing Wang, "Intelligent computational techniques for multimodal data", Multimedia Tools and Applications, 2019, Vol. 78, pp. 23809–23814.
13. Ye. Sulema, I. Dychka, O. Sulema, "Multimodal Data Representation Models for Virtual, Remote, and Mixed Laboratories Development", Lecture Notes in Networks and Systems, Springer Cham, 2018, Vol. 47, pp. 559-569.
14. Anat Levin, Dani Lischinski, Yair Weiss. "Colorization using Optimization" Proceedings of ACM SIGGRAPH, 2004, pp. 689–694.

Received 12.11.2019.

Accepted 15.11.2019.

Інтелектуальна колоризація зображень за допомогою генеративних змагальних мереж

Статтю присвячено застосуванню генеративних змагальних нейронних мереж на прикладі розв'язку задачі інтелектуальної колоризації зображень, що виконується за допомогою двох конкуруючих згорткових нейронних мереж. На етапі підготовки даних пропонується виконувати перетворення зображень з колірної моделі RGB у колірну модель YUV або LAB. Також у статті розглянуто можливість переходу від оброблення зображень до оброблення мультимодальних даних та запропоновано подавати вхідні мультимодальні дані у вигляді мультимодальної багаторівневої онтологічної моделі.

Image intelligent colorization using generative adversarial networks

The paper is devoted to the use of generative adversarial neural networks on the example of solving the problem of image intelligent colorization. Intelligent colorization is the process of changing colors of certain objects and details in an image by using mathematical algorithms. Existing solutions to this problem use standard methods, such as overlaying additional layers on highlighted areas in the image, changing the brightness, saturation, and hue levels. However, all of these methods require active involvement and user skills to locate the objects which need to be repainted, and then also need to adjust other metrics such as shadows, brightness, and exposure on the image to make the converted image look natural. However, machine learning technologies enable automated image processing. The paper presents the software system for intelligent colorization based on the work of two adversarial neural networks: the network-generator and the network-discriminator. The first network is a generator network; it performs input image processing in accordance with the purpose set by the user. The objective of this network is to generate data that is as similar as possible to the original data used in training. The second network is a discriminating network; it evaluates the quality of the result by determining the difference between the sample image that sets the desired processing result and

the image that the first network generates. This approach provides high quality training for both networks and, accordingly, high quality of the results. At the stage of data preparation, it is proposed to perform image conversion from RGB color model to YUV or LAB color model that allows to achieve better colorization result. The paper also considers the possibility of moving from image processing to multimodal data processing and proposes to represent input multimodal data about objects on an image in the form of a multimodal multilevel ontological model.

Сулема Евгения Станиславовна - к.т.н., доцент кафедри програмного забезпечення комп'ютерних систем Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського».

Топчиев Борис Сергеевич - магістрант кафедри програмного забезпечення комп'ютерних систем Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського».

Сулема Євгенія Станіславівна – к.т.н., доцент кафедри програмного забезпечення комп'ютерних систем Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Топчієв Борис Сергійович – магістрант кафедри програмного забезпечення комп'ютерних систем Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Sulema Yevgeniya - candidate of technical sciences, associate professor of the computer systems software department of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

Topchiiev Boris - graduate student of the Department of Computer Systems Software of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.