

ПРОБЛЕМА ОЦІНЮВАННЯ ГЕНЕРАТИВНИХ ЗМАГАЛЬНИХ МЕРЕЖ

Вінницький національний технічний університет

Анотація

У даній роботі представлено огляд сучасних методів оцінювання результатів генерування генеративних змагальних мереж (ГЗМ), проаналізовано особливості, переваги та недоліки цих методів, а також показано важливість вирішення проблеми оцінювання ГЗМ для їх подальшого розвитку.

Ключові слова: генеративна змагальна мережа, ГЗМ, оцінювання, вірність, різноманітність.

Abstract

The article presents an overview of modern evaluation methods of generation results of generative adversarial networks (GANs). It also shows the analysis of the features, advantages, and disadvantages of those evaluation methods, as well as the importance of solving the problem of GAN evaluation for their further development.

Keywords: generative adversarial network, GAN, evaluation, fidelity, diversity.

Сьогодні розвиток у галузі генеративних моделей швидко набирає оберти. Одним із популярних видів таких моделей є генеративні змагальні мережі (ГЗМ) [1]. Наразі такі моделі здатні: генерувати високоякісні зображення людей та тварин; створювати фотореалістичні зображення з сегментаційних малюнків, де кольорами показані елементи сцени, такі як вода, трава, скелі та інше; переносити стиль одних зображень на інші, як, наприклад, перетворення коней на зебр, та навпаки.

Незважаючи на бурхливий розвиток ГЗМ, відкритим та все більш актуальним є питання оцінювання результатів їх роботи. Якщо взяти, наприклад, звичайну задачу класифікації, то ми можемо легко оцінити якість роботи моделей на тестовій вибірці даних, для якої ми маємо вхідні дані і правильні відповіді. Це стосується й інших задач навчання з вчителем: регресії, сегментації, детекції тощо. Для генеративних моделей вхідні дані – це вектор випадкового шуму, і ми не можемо знати, яке зображення повинна отримати генеративна модель або які кольори пікселів повинні бути в тому чи іншому місці на картинці.

Зазвичай для оцінювання такого роду моделей використовується візуальний аналіз даних, що були згенеровані моделлю, експертом конкретної предметної області, який може зробити висновок, наскільки гарними та відповідними виходять зображення. Але цей підхід має великі ризики щодо компетенцій експертів, до того ж у деяких областях навіть експерт не здатен точно ідентифікувати коректність зображення.

Існують і автоматичні способи оцінювання генеративних моделей. Найбільш відомими є Inception Score [2] та Fréchet Inception Distance [3]. Ці підходи отримали свою назву завдяки тому, що вони використовують готову модель нейронної мережі Inception, натреновану фахівцями з машинного навчання компанії Google на відомому наборі даних ImageNet, для формування певного набору ознак для вхідних зображень.

Inception Score використовує значення з останнього шару Inception мережі, які вказують на імовірність приналежності вхідного зображення до одного з тисячі класів набору даних ImageNet. Цей підхід базується на двох критеріях:

- вірність (fidelity) – як добре окремі зображення, згенеровані моделлю, відносяться до певного класу (наприклад, наскільки модель «впевнена», що зображення представляє собою собаку або літак);
- різноманітність (diversity) – як добре згенеровані моделлю зображення розподіляються за окремими класами (тобто якщо модель генерує лише один клас з тисячі, то вона не є різноманітною).

Але у процесі оцінювання описаного вище підходу реальні зображення ніякої участі не беруть. Зі

свого боку Fréchet Inception Distance дозволяє оцінити, наскільки схожими є багатовимірні нормальні розподіли реальних та згенерованих зображень. Оскільки порівняння пікселів зображень вочевидь не має сенсу, у даному підході Inception мережа дозволяє для кожного зображення отримати вектори абстрактних ознак, для яких визначається середнє та коваріація, котрі надалі беруть участь у розрахунку близькості розподілів. Зазвичай вектори абстрактних ознак вилучаються з передостанніх шарів моделі, оскільки ці шари описують важливі з точки зору моделі ознаки вхідних зображень.

Хоча описані методи дозволяють оцінити моделі за різними критеріями (вірність та різноманітність генерування, схожість з реальними даними), але вони мають один суттєвий недолік – прив'язку до мережі Inception, що навчена на специфічному наборі даних ImageNet і може не в повній мірі описувати інші вибірки, як, наприклад, знімки комп'ютерної томографії або мікроскопічні зображення клітин. А тому вектори ознак, що використовують ці підходи, можуть бути недостатньо інформативними для оцінювання. Для методу Fréchet Inception Distance також необхідно враховувати, що він працює зі статистичними показниками: середнім та коваріацією вибірки, але не всі дані мають нормальний розподіл.

Висновки

Забезпечення якісного генерування зображень є важливою та актуальною проблемою. Незважаючи на те, що в своїй основі генеративні змагальні мережі мають дискримінатор (або критика), що відповідає за адекватність роботи генератора і намагається визначити, які фотографії є реальними, а які згенерованими, часто бувають ситуації, що і генератор, і дискримінатор не звертають уваги на дуже вагомні ознаки реального набору даних. Наприклад, генератор може відмінно навчитись створювати обличчя людей, але мало уваги приділятиме фону, в результаті чого на зображеннях можуть з'являтися неприродні фрагменти чи артефакти, або ж генератор може навчитись створювати зображення клітин організмів, але випадково додавати на них більше або менше певних органел. Для уникнення таких випадків необхідним є створення нових методів контролю і оцінювання роботи ГЗМ.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio "Generative Adversarial Networks," arXiv: 1406.2661 [stat.ML], Jun. 2014.
2. Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., and Chen, X. Improved techniques for training gans. In Advances in Neural Information Processing Systems, pp. 2234–2242, 2016.
3. Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, Sepp Hochreiter, "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium," arXiv:1706.08500 [cs.LG], Jun. 2017.

Ісаснков Ярослав Александрович — аспірант, факультет комп'ютерних систем і автоматики, Вінницький національний технічний університет, Вінниця, e-mail: oiuygl@gmail.com

Мокін Олександр Борисович — д-р техн. наук, професор, професор кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, Вінниця, e-mail: abmokin@gmail.com

Isaienkov Yaroslav — graduate student, Faculty of Computer Systems and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail: oiuygl@gmail.com

Mokin Oleksandr — Dr. Sc. (Eng.), Professor of the Department of System Analysis and Information Technologies, Vinnytsia National Technical University, Vinnytsia, e-mail: abmokin@gmail.com