DOI: https://doi.org/10.36910/6775-2524-0560-2023-52-13
УДК 004.054
Прозур Віталій Олександрович, аспірант
https://orcid.org/0009-0000-2996-483X
Інститут Прикладного Системного Аналізу, НТУУ «КПІ» ім. Ігоря Сікорського, м. Київ, Україна

АНАЛІЗ ВИДІВ ГЕНЕРАТИВНИХ ЗМАГАЛЬНИХ МЕРЕЖ

Прозур В.О. Аналіз видів генеративних змагальних мереж. У статті проаналізовано види генеративних змагальних мереж RNNGAN, WGAN, BiGAN. Розкрито їх структуру та компоненти. Зазначається, що у сучасних дослідженнях в галузі штучного інтелекту та обробки сигналів широко використовується підхід, що базується на використанні рекурентних нейронних мереж (RNN) та генеративних змагальних мереж (GAN). Однією з інноваційних концепцій в цій області є модель, відома як Recurrent Neural Network GAN. Запропоновано архітектуру мережі та математично представлено функцію втрат генератора та дискримінатора. Представлено модель Recurrent Conditional GAN, яка створена з метою генерації медичних даних, що є актуальним завданням у сучасних областях медичного дослідження та діагностики. В цьому підході використовується комбінація рекурентних нейронних мереж (RNN) та умовних генеративних змагальних мереж (cGAN). Запропоновано архітектуру мережі та математично представлено функцію втрат генератора та дискримінатора. Розкрито принципи мережі TimeGAN. Наголошується, що алгоритм TimeGAN включає в себе структуру, яка використовує елементи звичайних навчальних підходів GAN без учителя, а також підхід з учителем. Пропонується повна архітектура моделі TimeGAN та опис математичних функцій. Проаналізовано Bidirectional Generative Adversarial Networks (BiGAN), які представляють собою вид генеративних змагальних мереж, який включає кодувальник, додатково до звичайних компонентів генератора та дискримінатора, який перетворює реальні дані у латентний простір, в який вводиться генератор, фактично виконуючи обернену функцію порівняно з генератором. Зазначається, що тренування генеративно-змагальних мереж (GAN) представляє собою досить складне завдання. Існує можливість, що моделі можуть не збігтися до оптимального стану. Представлений аналіз підкреслює багатообіцяючі перспективи та різноманітність описаних підходів, що можуть сприяти подальшому розвитку галузей, де вони застосовуються, від медицини до мистецтва та інженерії.
Ключові слова: змагальна мережа, аналіз, алгоритм, нейронна мережа, дискримінатор, генератор, дані.

Prozur Vitalii. Analysis of types of generative competitive networks. The article analyzes the types of generative competitive networks RNNGAN, WGAN, BiGAN. Their structure and components are disclosed. It is noted that in modern research in the field of artificial intelligence and signal processing, an approach based on the use of recurrent neural networks (RNN) and generative adversarial networks (GAN) is widely used. One of the innovative concepts in this area is a model known as Recurrent Neural Network GAN. The architecture of the network is proposed and the loss function of the generator and discriminator is presented mathematically. The Recurrent Conditional GAN model is presented, which was created for the purpose of generating medical data, which is an urgent task in the modern fields of medical research and diagnostics. This approach uses a combination of recurrent neural networks (RNNs) and conditional generative adversarial networks (cGANs). The architecture of the network is proposed and the loss function of the generator and discriminator is presented mathematically. The principles of the TimeGAN network are disclosed. It is emphasized that the TimeGAN algorithm includes a framework that uses elements of conventional unsupervised GAN training approaches as well as a tutored approach. A complete architecture of the TimeGAN model and a description of the mathematical functions are offered. Bidirectional Generative Adversarial Networks (BiGAN) are analyzed, which are a type of generative adversarial networks that includes an encoder, in addition to the usual generator components and a discriminator, which transforms real data into a latent space into which the generator is input, actually performing the inverse function compared to the generator. It is noted that training generative-competitive networks (GAN) is a rather difficult task. There is a possibility that the models may not converge to the optimal state. The presented analysis highlights the promising prospects and diversity of the described approaches, which can contribute to the further development of the fields where they are applied, from medicine to art and engineering.
Key words: adversarial network, analysis, algorithm, neural network, discriminator, generator, data.

Вступ та постановка проблеми. З появою генеративних змагальних мереж (GAN) у 2014 році, що були розроблені Яном Гудфелоу та його співавторами, відбувся значний прогрес у полі покращень та модифікацій даного алгоритму.

Заснований на концепції двох конкуруючих нейронних мереж, GAN призначає одну мережу для генерації даних, а іншу – для оцінки їх автентичності. Ця взаємодія створює внутрішній механізм змагання, що сприяє навчанню генеративної мережі (генератор) створювати дедалі більш реалістичні дані, в той час як дискримінативна мережа (дискримінатор) вчиться краще відрізняти справжні дані від створених.

Аналіз останніх досліджень і публікацій. Наукова база щодо теми дослідження доволі широка. Так Я. О. Ісаєнков та О. Б. Мокін [1] здійснили аналіз генеративних моделей глибокого

навчання та особливостей їх реалізації на прикладі WGAN. Авторами представлено особливості будови, навчання та сфери застосування генеративних моделей глибокого навчання. До основних завдань таких модель відносяться генерування даних (зображень, музики, текстів, відео), перенесення стилів з одних даних на інші, поліпшення якості даних, їх кластеризація, пошук аномалій тощо. Науковці зазначають, що результати роботи генеративних моделей, окрім поширених розважальних цілей, можуть використовуватися як: додаткові дані для навчання інших моделей машинного навчання, джерела нових ідей для творчих професій, інструменти анонімізації чутливих даних тощо.

Робота [2] присвячена 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.

Стосовно неконтрольованих генеративних моделей варто відмітити роботу [3]. Авторами розглянуто використання алгоритму гармонійного пошуку в нейронних мережах для покращення виявлення шахрайства в банківській системі. З'ясовано, що, хоча дана модель має перевагу у спроможності до навчання на основі минулої поведінки, є труднощі в тривалій обробці великої кількості нейронних мереж. Також наведено етапи реалізації моделі. Крім того, проаналізовано моделювання виявлення шахрайства з кредитними картками на базі використання двох типів моделей: під наглядом і без нагляду. До моделей під наглядом віднесено логістичну регресію, K найближчі сусіди, екстремальне підвищення градієнта. Серед неконтрольованих генеративних моделей розглянуто однокласну опорну векторну модель, обмежену модель Больцмана, генеративно-змагальну мережу.

У статті [4] розглянуто основні шляхи використання нейронних мереж та методів машинного навчання різних типів у комп'ютерних відеоіграх. Машинне навчання та нейромережі – гарячі теми в багатьох технологічних галузях. Одна з них – створення комп'ютерних ігор, де нові інструменти використовуються для того, щоб зробити ігри цікавіше. Ремастерінг і модифікації ігор нейронними мережами стали новим трендом. Проводяться дослідження з корекції кольору та світла, анімації персонажів у реальному часі та керування їхньою поведінкою. Розглянуто основні типи нейронних мереж, які можуть навчатися таким функціям.

Із зарубіжних авторів варто відмітити роботи таких науковців як: Т. Каррас, С. Лайне, М. Айттала, Дж. Хеллстен, Й. Лехтінен і Т. Айла [5], М. Пасіні [6], Є. Сулема, І. Дичка, О. Сулема [7], Хорн К. Дж., Булл Л., Гіфтед-Пезуа О. [8], Шаббір А., Шабір М., Джавед А. Р., Чакраборті К., Різван М. [9], Д. Кроче, Г. Кастеллуччі та Р. Базілі [10], К. Се, Дж. Ван, З. Чжан, З. Рен, А. Юйле [11], Л. Лю, Ю. Лу, М. Ян, Цюй, Жу Чжу та Х. Лі [12], Д. Якубовіц, Р. Жир'єс [13], Х. Зенаті, К. С. Фу, Б. Лекуа, Г. Манек та В. Р. Чандрасекар [14]. та інших.

Однак незважаючи на масштабність наукових досліджень питання актуальності даної роботи не викликає сумнівів.

Постановка завдання. Метою дослідження є аналіз різних видів GAN (такі як RNNGAN, WGAN, BiGAN, тощо), який розкриє їхню структуру та компоненти.

Викладення основного матеріалу дослідження. В умовах сьогодення, генеративні моделі стають все більш популярними та виходять за рамки наукового пізнання. Їх застосування формується на базі широкого спектру задач. Розглянемо основні моделі генеративних змагальних мереж для усвідомлення їх структури та компонентів.

C-RNNGAN

Recurrent Neural Network GAN. У сучасних дослідженнях в галузі штучного інтелекту та обробки сигналів широко використовується підхід, що базується на використанні рекурентних нейронних мереж (RNN) та генеративних змагальних мереж (GAN). Однією з інноваційних концепцій в цій області є модель, відома як Recurrent Neural Network GAN.

Ця архітектура нейронної мережі поєднує в собі властивості RNN, які дозволяють моделювати та узагальнювати послідовність даних, та GAN, які забезпечують можливість генерації нових даних. Важливо зазначити, що початково RNN GAN створена автором для аналізу аудіосигналів як часових

рядів, де кожен відтинок аудіо сприймається як точка в часі, а послідовність таких точок утворює основу для створення музичних фрагментів.

Архітектура C-RNNGAN

Генератор – це RNN, а дискримінатор – двонаправлений RNN, завдяки чому дикримінатор може сприймати контекст послідовності в обох напрямках. Загальну структуру мережі можна побачити на рисунку 1.

Функція втрат C-RNNGAN

Функція втрат генератора:

$$L_G = \frac{1}{m}\sum_{i=1}^{m} \log\left(1 - D\left(G\left(z^{(i)}\right)\right)\right)$$

Функція втрат дискримінатора:

$$L_D = \frac{1}{m}\sum_{i=1}^{m} \left[-\log D(x^{(i)}) - \log\left(1 - D\left(G\left(z^{(i)}\right)\right)\right)\right]$$

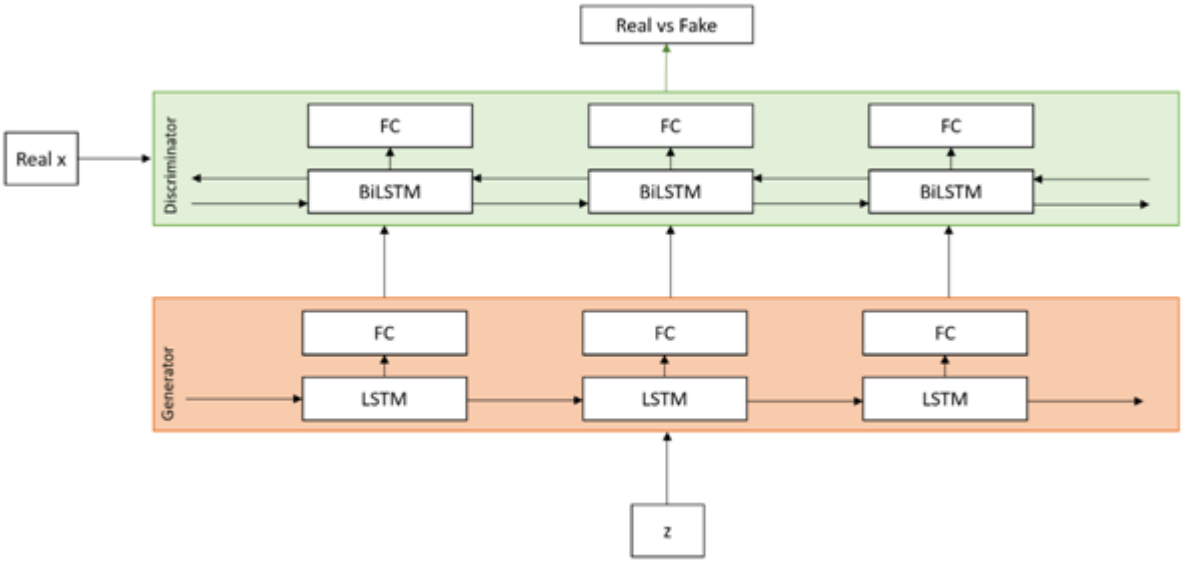

Рисунок 1 – Архітектура C-RNNGAN

RCGAN

Представлено модель Recurrent Conditional GAN, створена з метою генерації медичних даних, що є актуальним завданням у сучасних областях медичного дослідження та діагностики. В цьому підході використовується комбінація рекурентних нейронних мереж (RNN) та умовних генеративних змагальних мереж (cGAN), що дозволяє не тільки створювати реалістичні медичні дані, але і контролювати їхні характеристики залежно від вхідних параметрів.

Рекурентні нейронні мережі дозволяють моделювати залежності у часових послідовностях, що є важливим для медичних даних, де часові та просторові фактори можуть мати визначальний вплив на клінічні показники. Умовні GAN додають до процесу генерації додатковий рівень контролю, дозволяючи враховувати конкретні характеристики пацієнтів або клінічні сценарії.

Архітектура RCGAN

Архітектурно, модель RCGAN відрізняється від RNN-GAN. Незважаючи на використання RNN LSTM у RCGAN, дискримінатор відрізняється своєю односпрямованою структурою, тому результати G не передаються як вхідні дані на наступний часовий крок.

Додатково, відзначається, що RCGAN приймає вхідний вектор c як умовний вектор для послідовності та оптимізується відносно цього вектора. Генеральну конфігурацію моделі наведено на рисунку 2.

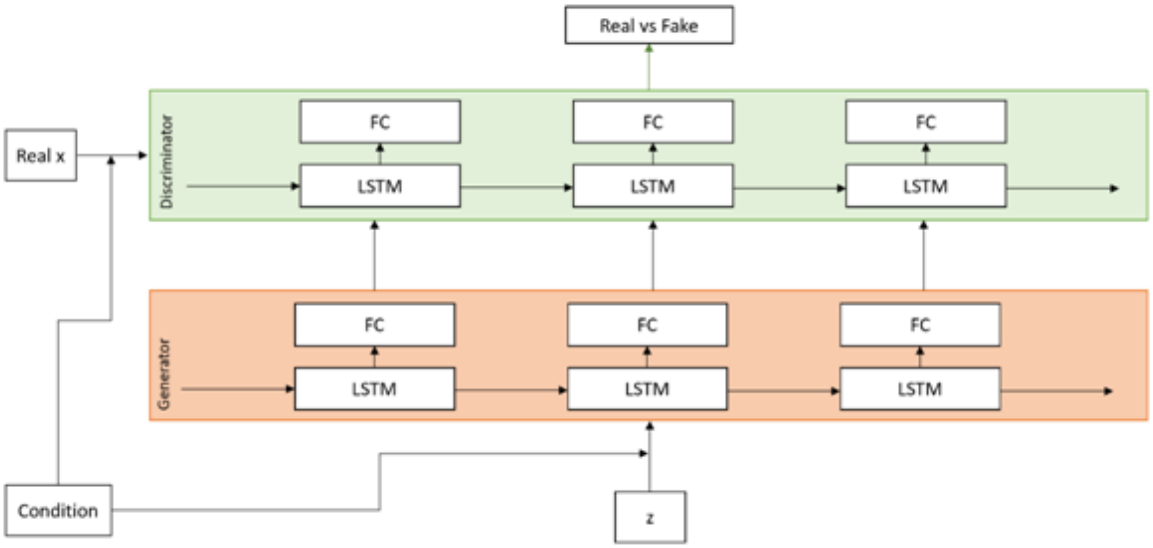

Рисунок 2 – Загальна архітектура RCGAN та її розширена версія

Функція втрат RCGAN

Функція втрат генератора:

$$L_G = \frac{1}{m}\sum_{i=1}^{m} \log\left(1 - D\left(G\left(z^{(i)}\right)\right)\right)$$

Функція втрат дискримінатора:

$$L_D = \frac{1}{m}\sum_{i=1}^{m} \left[-\log D(x^{(i)}) - \log\left(1 - D\left(G\left(z^{(i)}\right)\right)\right)\right]$$

TimeGAN

Основною ціллю автора було створення моделі, яка здатна відтворювати взаємозв'язки в послідовностях, одночасно зберігаючи всі статистичні характеристики цих послідовностей. Алгоритм TimeGAN включає в себе структуру, яка використовує елементи звичайних навчальних підходів GAN без учителя, а також підхід з учителем. Шляхом поєднання неконтрольованої мережі GAN з контрольованою авторегресійною моделлю, дана архітектура націлена на генерацію часових рядів, які зберігають їхню індивідуальну часову динаміку.

Архітектура TimeGAN

Архітектурно, ця модель складається із п'яти основних складових: генератора, дискримінатора, супервайзера, інкодера та декодера.

Ця модель дотримується концепції латентного простору GAN, що дозволяє використовувати інкодер та декодер для трансформації вхідних даних у цей код, а завдяки присутності супервайзера, досягається можливість навчання з учителем, що сприяє прискоренню процесу навчання моделі.

Кожна з окремих компонент цієї моделі реалізована за допомогою рекурентних нейронних мереж з використанням комірок GRU.

Повна архітектура моделі TimeGAN наведена на рисунку 3.

Функції втрат TimeGAN

Ця модель має 3 функції втрат: функцію втрат декодування (reconstruction loss), функцію без та з наглядом (unsupervised та supervised loss).

Reconstruction loss:

$$\mathscr{L}_R = MSE[x, De(E(x))]$$

Unsupervised loss:

$$\mathscr{L}_U = \frac{1}{m}\sum_{i=1}^{m}\left[-\log D(x^{(i)}) - \log\left(1 - D\left(G(z^{(i)})\right)\right)\right],$$

Supervised loss:

$$\mathscr{L}_S = MSE[E(x), G(z)]$$

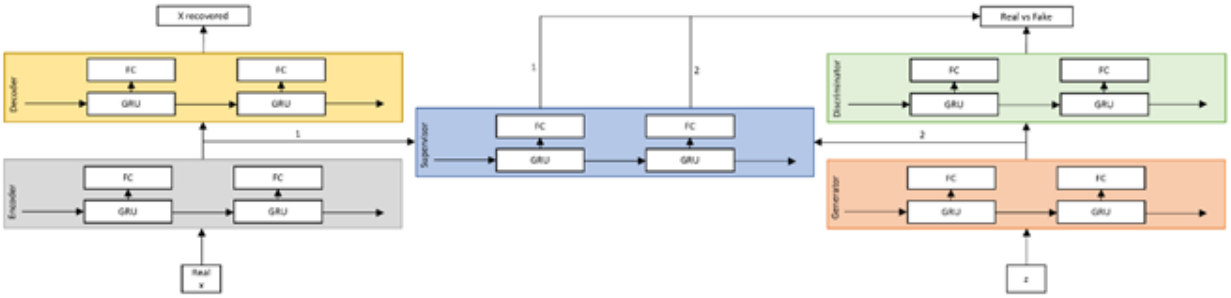

Рисунок 3 – Архітектура моделі TimeGAN

BiGAN

Bidirectional Generative Adversarial Networks (BiGAN) представляють собою вид генеративних змагальних мереж, який включає кодувальник, додатково до звичайних компонентів генератора та дискримінатора, який перетворює реальні дані у латентний простір, в який вводиться генератор, фактично виконуючи обернену функцію порівняно з генератором.

Архітектура BiGAN

Таким чином, архітектура BiGAN розширює звичайний GAN, включаючи кодер E(x) до генератора G(z). У цьому підході дискримінатор опрацьовує не лише вхідні дані (x – реальні дані та G(z) – згенеровані дані), а також і інформацію з латентного простору (E(x) – вихід кодера, z – вхід генератора), що призводить до роботи з парами (x, E(x)) – для реальних даних, та (G(z), z) – для згенерованих даних. Загальну архітектуру можна побачити нижче.

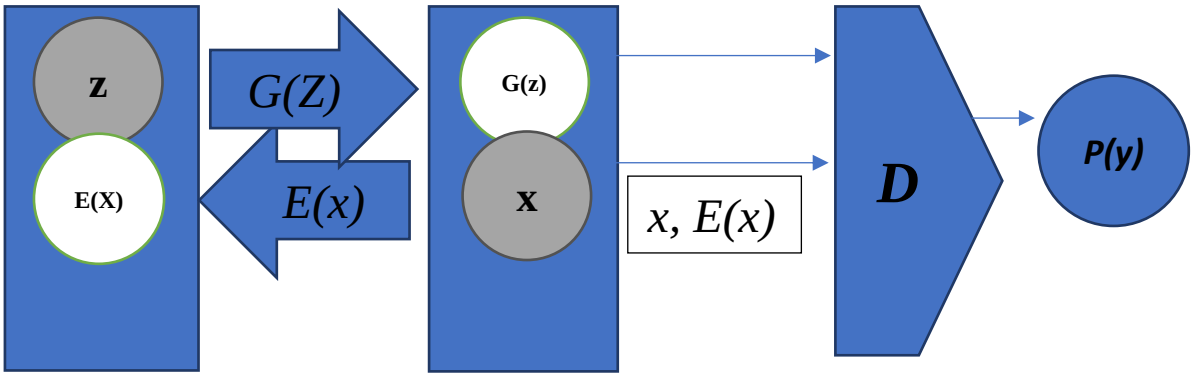

Рисунок 4 – Загальна архітектура BiGAN

Функція втрат BiGAN

Мета навчання BiGAN, може бути сформована як задача мінімаксу:

$$\min_{G,E}\max_{D} V(D,E,G)$$

де

$$V(D,E,G) \coloneqq \mathbb{E}_{x\sim p_x}\underbrace{\left[\mathbb{E}_{z\sim \rho_{E(\cdot|x)}}[logD(x,z)]\right]}_{logD(x,E(x))} + \mathbb{E}_{z\sim p_z}\underbrace{\left[\mathbb{E}_{x\sim \rho_{G(\cdot|z)}}\left[log\left(1-D(x,z)\right)\right]\right]}_{log\left(1-D(G(z),z)\right)}$$

Wasserstein GAN

Тренування генеративно-змагальних мереж (GAN) представляє собою досить складне завдання. Існує можливість, що моделі можуть не збігтися до оптимального стану.

Основна мета, яка лежить в основі підходу GAN, полягає в тому, щоб генератор почав перетворювати випадковий шум у тип даних, який ми бажаємо імітувати. Для досягнення цієї мети, ідея порівняння подібності між двома розподілами є надзвичайно важливою в рамках GAN. Серед найбільш широко використовуваних метрик для порівняння розподілів можна виділити:

JS Дивергенція (Йенсен – Шеннон) –

$$D_{JS}(p|q) = \frac{1}{2}DKL\left(p\left|\frac{p+q}{2}\right.\right) + \frac{1}{2}D_{KL}\left(q\left|\frac{p+q}{2}\right.\right)$$

Розбіжність KL(Kullback – Leibler) –

$$D_{KL}(p|q) = \int_x p(x)\log fracp(x)q(x)dx. D_{KL}$$

дорівнює нулю, коли $p(x)$ дорівнює $q(x)$.

Розбіжність JS обмежена 0 і 1, і, на відміну від розбіжності KL, симетрична і плавна, а отже її використання зазвичай є більш доцільним.

Функція втрат WGAN

WGAN же натомість використовує відстань Вассерштайна:

$$\$W\left(p_r,p_g\right) = frac1Ksup_{|f|L\ leqK}mathbbEx\ simp_r[f(x)] - mathbbE_{x\ simp_g}[f(x)]\$$$

як функцію втрат. Порівняно з розбіжностями KL та JS, метрика Вассерштайна дає плавний показник (без різких стрибків розбіжностей). Це робить його набагато більш придатним для створення стабільного навчального процесу під час градієнтного спуску.

Крім того, порівняно з KL та JS, відстань Вассерштайна майже скрізь диференціюється. Як ми знаємо, під час зворотного поширення ми диференціюємо функцію втрат, щоб створити градієнти, які, в свою чергу, оновлюють ваги. Тому наявність диференційованої функції втрат є досить важливою.

Висновки. В даній статті були детально проаналізовані різні види генеративних змагальних мереж, включаючи WGAN, C-RNNGAN, RCGAN, TIMEGAN та BiGAN. Кожен з цих підходів приносить унікальні інновації у галузі генерації та моделювання даних.

WGAN вирішує проблему стійкості навчання GAN із застосуванням функції втрати Вассерштейна. C-RNNGAN використовує рекурентні мережі для генерації послідовних даних, що дозволяє досягти вражаючої якості в музичному контенті. RCGAN пропонує новітній підхід, де рекурентні мережі використовуються для аналізу та генерації медичних даних.

TIMEGAN розширює можливості у генерації часових рядів, узгоджуючи неконтрольовану GAN з контрольованою авторегресійною моделлю. Нарешті, BiGAN розширює стандартну архітектуру GAN, додавши кодувальник для отримання зворотних перетворень даних.

Цей аналіз підкреслює багатообіцяючі перспективи та різноманітність даних підходів, що можуть сприяти подальшому розвитку галузей, де вони застосовуються, від медицини до мистецтва та інженерії.

Список бібліографічного опису

1. Ісаєнков Я. О., Мокін О. Б. Аналіз генеративних моделей глибокого навчання та особливостей їх реалізації на прикладі WGAN. Вісник Вінницького політехнічного інституту. 2022. № 1. С. 82-94. DOI 10.31649/1997-9266-2022-160-1-82-94

2. Сулема Є. С., Топчієв Б. С. Інтелектуальна колоризація зображень за допомогою генеративних змагальних мереж. «Системні технології» «System technologies», 2019. № 5 (124). С. 94-103. DOI 10.34185/1562-9945-5-124-2019-09
3. Аналіз математичних моделей протидії банківським кібершахрайствам / Кузьменко О. В., Яровенко Г. М., Скринька Л.О. // Вісник СумДУ. Серія «Економіка», 2022. № 2'. С. 111-120. DOI:10.21272/1817-9215.2022.2-13
4. Сеніва К. Р. Способи використання нейронних мереж та машинного навчання в комп'ютерних іграх. Вісник Хмельницького національного університету, 2021. №2 (295). С. 97-100. DOI 10.31891/2307-5732-2021-295-2-97-100.
5. T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and Improving the Image Quality of StyleGAN," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 8107-8116. http://doi.org/10.1109/CVPR42600.2020.00813.
6. M. Pasini, "MelGAN-VC: Voice Conversion and Audio Style Transfer on arbitrarily long samples using Spectrograms," in arXiv e-prints, 2019. [Online]. Available: https://arxiv.org/pdf/1910.03713.pdf.
7. 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.
8. Horna C. J., Toro L., Regalado-Pezua O. Silver bank: Vulnerability and risks during cyberattacks. Emerald Emerging Markets Case Studies. 2022. Vol. 12, no. 1. P. 1-33. DOI: 10.1108/EEMCS-02-2021-0034.
9. Shabbir A., Shabir M., Javed A. R., Chakraborty C., Rizwan M. Suspicious transaction detection in banking cyber-physical systems. Computers and Electrical Engineering. 2022. Vol. 97. DOI: 10.1016/j.compeleceng.2021.107596.
10. D. Croce, G. Castellucci, and R. Basili, "GAN-BERT: Generative Adversarial Learning for Robust Text Classification with a Bunch of Labeled Examples," in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, pp. 2114–2119. https://doi.org/10.18653/v1/2020.acl-main.191.
11. Mitigating Adversarial Effects Through Randomization. / C. Xie, J. Wang, Z. Zhang, Z. Ren, A. Yuille // Proceedings of the International Conference on Learning Representations, Toulon, France, 24–26 April 2017. 2017. P. 1–16. https://doi.org/10.48550/arXiv.1711.01991.
12. L. Liu, Y. Lu, M. Yang, Q. Qu, J. Zhu, and H. Li, "Generative Adversarial Network for Abstractive Text Summarization," in arXiv e-prints, 2017. [Online]. Available: https://arxiv.org/pdf/1711.09357.pdf.
13. Jakubovitz D. Improving DNN Robustness to Adversarial Attacks using Jacobian Regularization / D. Jakubovitz, R. Giryes // Proceedings of the European Conference on Computer Vision, Munich, Germany, 8-14 Sept. 2018. 2018. P. 1-16. DOI: https://doi.org/10.48550/arXiv.1803.08680.
14. H. Zenati, C. S. Foo, B. Lecouat, G. Manek, and V. R. Chandrasekhar, "Efficient GAN-Based Anomaly Detection," in arXiv e-prints, 2018. [Online]. Available: https://arxiv.org/pdf/1802.06222.pdf.

References

1. T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and Improving the Image Quality of StyleGAN," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 8107-8116. http://doi.org/10.1109/CVPR42600.2020.00813.
2. M. Pasini, "MelGAN-VC: Voice Conversion and Audio Style Transfer on arbitrarily long samples using Spectrograms," in arXiv e-prints, 2019. [Online]. Available: https://arxiv.org/pdf/1910.03713.pdf.
3. 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.
4. Horna C. J., Toro L., Regalado-Pezua O. Silver bank: Vulnerability and risks during cyberattacks. Emerald Emerging Markets Case Studies. 2022. Vol. 12, no. 1. P. 1-33. DOI: 10.1108/EEMCS-02-2021-0034.
5. Shabbir A., Shabir M., Javed A. R., Chakraborty C., Rizwan M. Suspicious transaction detection in banking cyber-physical systems. Computers and Electrical Engineering. 2022. Vol. 97. DOI: 10.1016/j.compeleceng.2021.107596.
6. D. Croce, G. Castellucci, and R. Basili, "GAN-BERT: Generative Adversarial Learning for Robust Text Classification with a Bunch of Labeled Examples," in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, pp. 2114–2119. https://doi.org/10.18653/v1/2020.acl-main.191.
7. Mitigating Adversarial Effects Through Randomization. / C. Xie, J. Wang, Z. Zhang, Z. Ren, A. Yuille // Proceedings of the International Conference on Learning Representations, Toulon, France, 24–26 April 2017. 2017. P. 1–16. https://doi.org/10.48550/arXiv.1711.01991.
8. L. Liu, Y. Lu, M. Yang, Q. Qu, J. Zhu, and H. Li, "Generative Adversarial Network for Abstractive Text Summarization," in arXiv e-prints, 2017. [Online]. Available: https://arxiv.org/pdf/1711.09357.pdf.
9. Jakubovitz D. Improving DNN Robustness to Adversarial Attacks using Jacobian Regularization / D. Jakubovitz, R. Giryes // Proceedings of the European Conference on Computer Vision, Munich, Germany, 8-14 Sept. 2018. 2018. P. 1-16. DOI: https://doi.org/10.48550/arXiv.1803.08680.
10. H. Zenati, C. S. Foo, B. Lecouat, G. Manek, and V. R. Chandrasekhar, "Efficient GAN-Based Anomaly Detection," in arXiv e-prints, 2018. [Online]. Available: https://arxiv.org/pdf/1802.06222.pdf.