

ДОСЛІДЖЕННЯ НЕЙРОМЕРЕЖЕВИХ МЕТОДІВ ЗІСТАВЛЕННЯ ЗОБРАЖЕНЬ ТА ЇХ ТЕКСТОВИХ АНОТАЦІЙ

Рябова Н. В., Потапов Д. С.

Харківський національний університет радіоелектроніки, Харків, Україна

Генеративне моделювання на цей час є одним з найбільш перспективних напрямків розвитку глибинного навчання та тем штучного інтелекту, що викликають жваві дискусії та обговорення. Галузями застосування генеративного моделювання є такі творчі галузі як: створення зображень, текстів, музики [1].

Найбільш актуальними прикладами генеративних моделей глибинного навчання є варіаційні автокодировщики, генеративно-змагальні мережі, моделі типу кодер-декодер.

Метою даної роботи є дослідження архітектур та методів генеративного моделювання з метою виявлення найбільш творчих та ефективних у навчанні моделей. У якості об'єкту дослідження розглядається перетворення тексту на зображення.

Слід зазначити, що створення реалістичного зображення, яке відповідає заданому текстовому опису, вимагає обробки нечіткої та неповної інформації у вигляді природно-мовних текстів. Предметом дослідження є генеративно-змагальні мережі (Generative Adversarial Networks, GAN) надзвичайно корисні для створення реалістичних зображень. Підходи на основі GAN продемонстрували великий успіх у створенні зображень високої роздільної здатності з фотореалістичними деталями на основі текстових описів, семантичної сегментації та графів сцен.

Для створення таких зображень, пропонується проста, але ефективна генеративно-змагальна мережа. Вона розкладає текст на зображення на два етапи.

Перший це – Stage-I GAN: накреслює примітивну форму та основні кольори об'єкта, обумовлені заданим текстовим описом, і малює макет фону з випадкового вектора шуму, що дає зображення з низькою роздільною здатністю.

Другий це – GAN Stage-II: виправляє дефекти в низькій роздільній здатності зображення з етапу-I і доповнює деталі об'єкта знову прочитавши текстовий опис, створивши фотореалістичне зображення високої роздільної здатності [2].

Список літератури

1. Foster D. Generative Deep Learning. Teaching Machines to Paint, Write, Compose and Play. O'Reilly Media, Inc. 2019. 330 p.
2. I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. C. Courville, and Y. Bengio. Generative Adversarial Nets. In NIPS, 2014.