УДК 339, 004.8

Марціяш Г. Я., спец. вищ. кат., Сербін В., Смага І. - ст. гр. КН-423
*Відокремлений структурний підрозділ «Тернопільський фаховий коледж Тернопільського національного технічного університету імені Івана Пулюя», Україна*

# ВИКОРИСТАННЯ НЕЙРОННИХ МЕРЕЖ ДЛЯ ГЕНЕРАЦІЇ РЕАЛІСТИЧНОГО КОНТЕНТУ: ПРОБЛЕМИ ТА МЕТОДИ ЇХ ВИРІШЕННЯ

Martsiyash H., Serbin V., Smaha I.
*Separate Structural Subdivision «Ternopil Professional College of Ternopil Ivan Puluj National Technical University»*

# USE OF NEURAL NETWORKS TO GENERATE REALISTIC CONTENT: PROBLEMS AND METHODS OF THEIR SOLUTION

Ключові слова: нейронні мережі, штучний інтелект, генерація зображень, діпфейки

Нейронні мережі – це інноваційна технологія, яка імітує роботу людського мозку та здатна до самостійного навчання без необхідності створення окремих алгоритмів для кожного завдання. Їх використання у сфері генерації зображень відкриває широкі можливості для автоматизації творчих процесів, економії часу та ресурсів, а також робить візуальне мистецтво доступним для людей без спеціальних навичок. Завдяки здатності зберігати узгодженість стилю, кольору та якості, нейромережі ефективно справляються зі створенням складних графічних елементів, зокрема фотореалістичних облич. Це стає проблемою при необхідності розпізнати дійний контент.

Сьогодні майже кожен, хто має смартфон та український додаток Reface, може замінити своє обличчя або обличчя друзів на чиєсь із знаменитостей. Загалом поняття діпфейків розуміється, як методика синтезу зображень різних людей за допомогою штучного інтелекту. Вона використовується для накладання вже існуючих відео чи зображень на початкові. Здебільшого, аби створити оманливе уявлення про когось або щось, поширення чорного піару та дезінформації тощо. Технологія діпфейків здатна, як розважити людину так і створити для неї проблеми [1]:

- Діпфейки можуть бути використані для створення фальшивих новин, політичних заяв, відео та іншого контенту, що можуть поширювати дезінформацію через громадськість.
- Діпфейки можуть бути використані для створення вигаданих відео або фотографій, які порушують приватність людей. Це може мати серйозні наслідки для особистого та професійного життя людей, оскільки їхні зображення можуть бути використані в шкідливих або неправдивих контекстах.
- Діпфейки можуть бути використані для шахрайства, фінансових афер, викрадення ідентичності та інших форм кіберзлочинності. Вони можуть допомогти зловмисникам отримати доступ до особистих даних [2].

Не дивно, але боротися з діпфейками можна лише за допомогою штучного інтелекту. Штучний інтелект вже допомагає виявляти підроблені відео, але багато існуючих систем виявлення мають серйозну слабкість: вони найкраще працюють для

знаменитостей, тому що вони можуть тренуватися годинами вільно доступних відеоматеріалів [1].

Нижче подано приклад використання нейронної мережі для створення діпфейку у вигляді серії знімків (рис. 1).

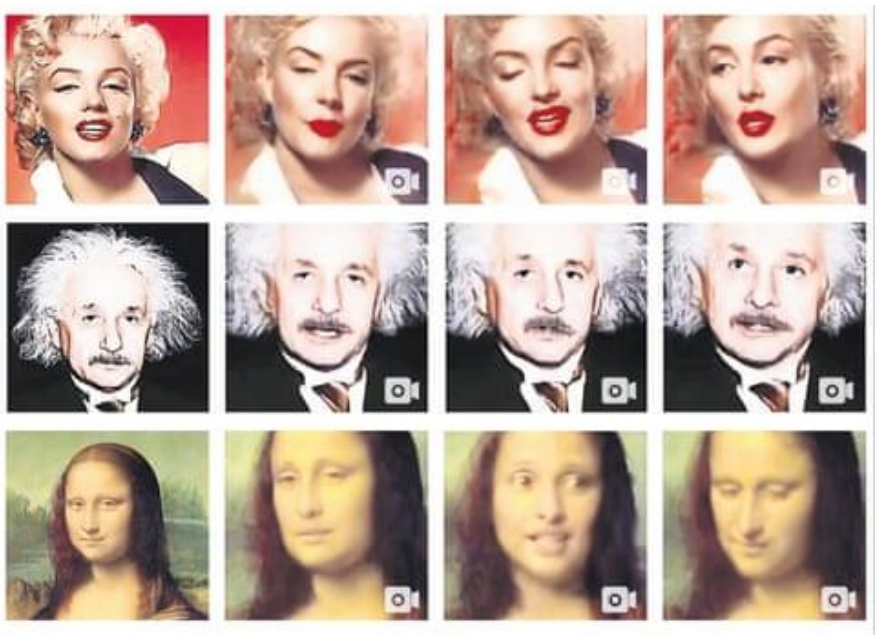

Рисунок 1 - Створення діпфейків [3]

Також для прикладу можна додати історії у яких учасники конкурсів, використовуючи нейронні мережі обманюють глядачів та журі. Одна із таких історій розповідає про німця, який подав на фотоконкурс зображення під назвою The Electrician. Знімок зображує двох жінок, нібито знятих в епоху перших днів винаходу технології фотографії [4] (рис. 2).

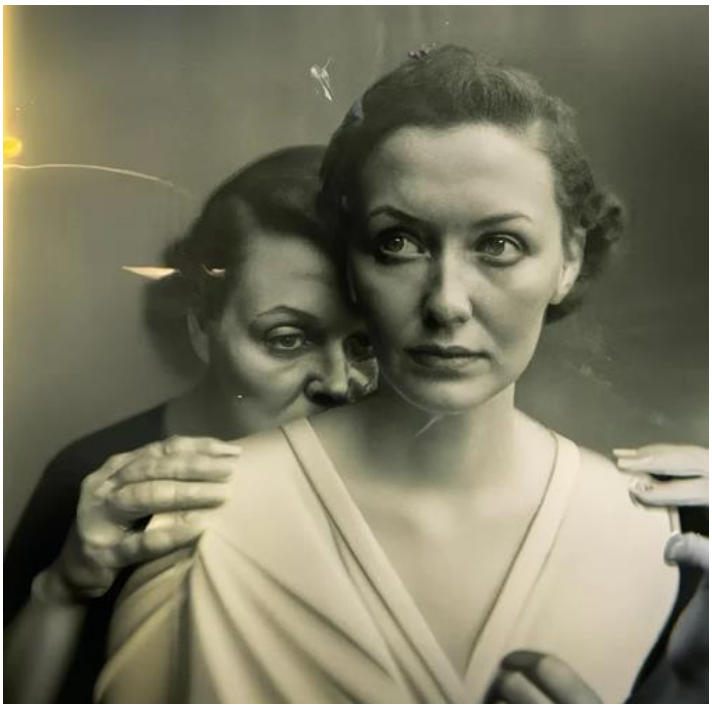

Рисунок 2 - Згенероване зображення

Дане зображення виглядає реалістичним, проте ним не є. Для таких випадків необхідно обробляти джерела інформації та перевіряти ознаки наявності використання штучного інтелекту. Щоб запобігти поширенню фотореалістичних зображень, створених нейронними мережами з маніпулятивною метою, та вберегти себе від дезінформації, варто дотримуватись таких порад:

– Використовувати інструменти перевірки зображень. Реверсивний пошук зображень (наприклад, через Google Images ) дозволяє виявити, чи публікувалося вже це фото, і в якому контексті. Спеціалізовані сервіси на базі ШІ, як-от Deepware Scanner або Sensity, допомагають виявити діпфейки.

– Звертати увагу на "ознаки" AI-зображень. Навіть найреалістичніші зображення можуть мати помилки: нереалістичні або асиметричні риси обличчя, дивні форми рук, вух, очей, розмиті або спотворені елементи фону, неприродне освітлення або відсутність тіней.

– Не поширювати неперевірений контент. Навіть жартома чи «просто цікаво» – будь-яке поширення може мати наслідки.

– Бути обережними з особистими фото.Не публікувати надто багато своїх чітких фото в соцмережах – їх можуть використати для створення діпфейків.

**Висновок:** Таким чином, нейронні мережі стали потужним інструментом для генерації фотореалістичного контенту, що активно використовується як у розважальних, так і в комерційних цілях. Попри значні переваги, зокрема доступність і швидкість створення зображень, їх широке використання породжує низку загроз – від поширення дезінформації до порушення приватності та кіберзлочинності.

Ситуацію ускладнює те, що частина користувачів свідомо експлуатує такі технології задля отримання вигоди або популярності, зокрема у соціальних мережах. Це підкреслює важливість розвитку етичних норм, інструментів ідентифікації штучно створеного контенту та відповідальності за його використання.

Література.

1. Needfredom.org.ua. Діпфейки як розпізнати та захиститися – [Електронний ресурс] – Режим доступу: https://surl.li/akkuyx
2. JustTalk.com.ua. Проблеми та виклики пов’язані зі збором електронних доказів – [Електронний ресурс] – Режим доступу : https://surl.li/mxlktp
3. TheGuardian.com. How do we do about deepfake video – [Електронний ресурс] – Режим доступу: https://surli.cc/elftab
4. LivePravda.com.ua. Зображення створене за допомогою ШІ, виграло престижний конкурс – [Електронний ресурс] – Режим доступу : https://surl.li/ezctin