

ШТУЧНИЙ ІНТЕЛЕКТ, ЯК ІНСТРУМЕНТ ПОЛІПШЕННЯ ЦИФРОВИХ ЗОБРАЖЕНЬ

Пестрецов Д. Р.

Національний авіаційний університет, м. Київ
Науковий керівник — Бобарчук О. А., к.т.н., завідувач кафедри КММТ

Анотація. Розглянуті програмні методи поліпшення зображень, які базуються на застосуванні штучного інтелекту від компанії Google. Модель машинного навчання перетворює фотографію з низькою роздільною здатністю на детальну фотографію з високою роздільною здатністю, і потенційні застосування цього варіюються від відновлення старих сімейних фотографій до покращення медичних зображень.

Ключові слова: цифрові зображення, поліпшення якості, програмне забезпечення, штучний інтелект.

Технології поліпшення фотографій у фільмах і телевізійних шоу часто висміюють як неймовірне, але дослідження в області виявлення дефектів фотографій насправді все більше й більше відносяться до наукової фантастики. Прикладом є остання технологія масштабування фотографій *Google AI*.

У запису під назвою «Генерація зображень високої точності з використанням моделей дифузії», опублікованому в блозі *Google AI* (і поміченому *DPR*), дослідники *Google* із підрозділу *Brain Team* розповідають про нові досягнення, які вони здійснили в галузі суперроздільності зображень.

Модель машинного навчання перетворює фотографію з низькою роздільною здатністю на детальну фотографію з високою роздільною здатністю, і потенційні застосування цього варіюються від відновлення старих сімейних фотографій до покращення медичних зображень.

Google досліджує концепцію під назвою «дифузійні моделі», яка вперше була запропонована в 2015 році, але донедавна поступалася місцем сімейства моделей глибокого вивчення під назвою «глибокі генеративні моделі». Під час опитування компанія виявила, що результати, отримані завдяки цьому новому підходу, випереджають існуючі технології.

Перший підхід називається *SR3* або *Super-Resolution via Repeated Refinement*.

«*SR3* — це модель дифузії з надвисокою роздільною здатністю, яка приймає зображення з низькою роздільною здатністю та створює відповідне зображення високої роздільної здатності з чистого шуму», — пише *Google*. «Модель навчена процесу спотворення зображення, під час якого шум поступово додається до зображення високої роздільної здатності, поки не залишиться лише чистий шум.

«Потім процес відбувається в зворотному напрямку, починаючи з чистого шуму та поступово видаляючи шум, щоб досягти цільового розподілу за допомогою вказівок вхідного зображення з низькою роздільною здатністю».

Було виявлено, що *SR3* добре працює при масштабуванні портретів і природних зображень. Якщо використовується для масштабування облич у 8 разів, «коефіцієнт

плутанини» становить майже 50 %, тоді як існуючі методи досягають лише 34 %, що свідчить про те, що результати справді фотореалістичні.

Після того, як в *Google* засвідчили, наскільки ефективним був *SR3* у масштабуванні фотографій, компанія пішла далі, використовуючи другий підхід під назвою *CDM*, модель класової умовної дифузії.

«*CDM* — це модель умовної дифузії класу, навчена на даних *ImageNet* для створення природних зображень високої роздільної здатності», — пише *Google*. «Оскільки *ImageNet* є складним набором даних з високою ентропією, ми створили *CDM* як каскад кількох моделей дифузії. Цей каскадний підхід включає об'єднання кількох генеративних моделей з різними роздільними здатностями: одна дифузійна модель, яка генерує дані з низькою роздільною здатністю, за якою слідує послідовність дифузійних моделей із надвисокою роздільною здатністю *SR3*, які поступово збільшують роздільну здатність створеного зображення до найвищої роздільної здатності».

Google опублікував набір прикладів фотографій із низькою роздільною здатністю, збільшених каскадом. Фотографію 32×32 можна збільшити до 64×64 , а потім до 256×256 . Фотографію 64×64 можна збільшити до 256×256 , а потім до 1024×1024 .

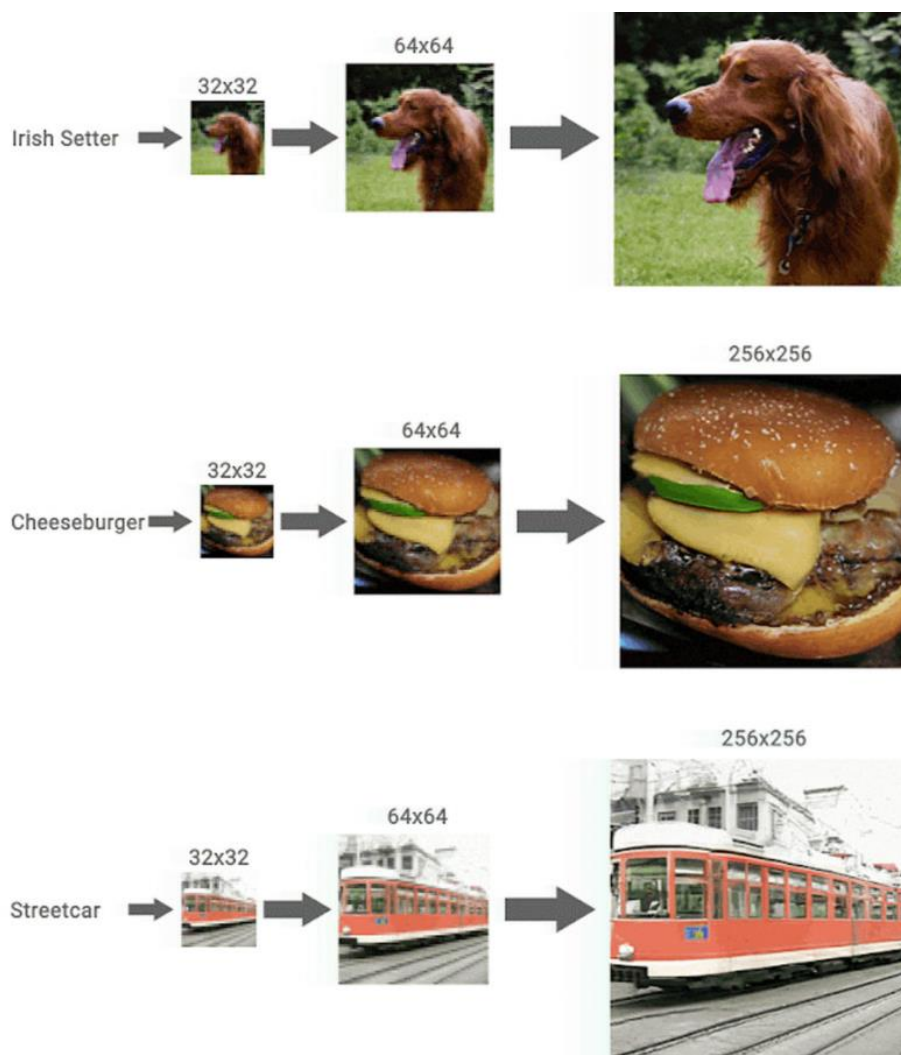


Рис. 1. Збільшення каскадом

«За допомогою *SR3* і *CDM* ми підвищили продуктивність дифузійних моделей до найсучаснішого рівня за суперроздільністю та тестами генерації *ImageNet*, що залежать від класу», — пишуть дослідники *Google*. «Ми раді продовжити тестування меж дифузійних моделей для широкого спектру проблем генеративного моделювання».

Результати вражаючі, і остаточні фотографії, незважаючи на деякі помилки, на перший погляд більшості глядачам, швидше за все, здадуться оригінальними фотографіями.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Cascaded Diffusion Model – <https://cascaded-diffusion.github.io>
2. High Fidelity Image Generation Using Diffusion Models – <https://ai.googleblog.com/2021/07/high-fidelity-image-generation-using.html?m=1>
3. Image Super-Resolution via Iterative Refinement – <https://iterative-refinement.github.io>