

Создана модель ИИ, которая намеренно портит учебные изображения

Исследователи из Техасского университета и Калифорнийского университета в Беркли нашли решение: обучать ИИ на намеренно испорченных изображениях. Их модель «Ambient Diffusion» берет высококачественные фотографии и искажает их с помощью сильного пиксельного маскирования.

Удивительно, но ИИ все равно учится генерировать реалистичные изображения. Вспомните картины импрессионистов — суть есть, но детали размыты. Это исключает риск нарушения авторских прав, поскольку «ИИ не может напрямую запомнить и воспроизвести оригиналы», отмечают учёные.

Это открывает двери не только в искусство. Представьте, как можно обучать ИИ на медицинских снимках, не нарушая конфиденциальность пациентов за счет размытия чувствительных деталей. Исследователи даже намекнули на возможность применения в научных областях, таких как визуализация черных дыр.

Модель не идеальна, но это значительный шаг, пишут учёные.