Contribution ID: **80** Type: **not specified**

Робастый метод для детекции машинносгенерированных изоюражений

Sunday, 18 May 2025 16:42 (12 minutes)

Тезис:
В связи с улучшением качества машиносгенерированных изображений становится очень сложно отличать реальное изображение от сгенерированных. Существующие на данный момент решения имеют низкую обобщающую способность. В этой статье рассматриваются разные модели, в том числе несвязанные с нейронными сетями. Также используется вся существующая информацию и модели, для подбора наилучшего решения. Дополнительно строится модель, которая сначала проверяет метод генерации, потом уже использует конкретную модель для этого метода генерации. Помимо этого, используются методы графических редакторов, на основе искусственного интеллекта.
Аннотация:
В современном мире в связи с развитием генераторов изображений человеческому глазу стало уже слишком сложно отличать настоящие изображение от машиносгенерированное. Ещё сложнее человеку отличить реальное изображение от реального, но с использованием графического редактора. В связи с доступностью этих сервисов стали очень распространены разные виды мошенничества, использующие машиногенерацию. Таким образом задача детекции машинносгенерированных изображений стала очень важна.
На данный момент не существует общего подхода к решению этой задачи, устойчивого относительно появления новых моделей. Например, появление диффузионных моделей генерации изображений свело сущесвтующие на тот момент методы к точности около 60 процентов. Таким образом, существующие на данный момент методы имеют низкую обобщающую способность. Актуальные научные статьи на эту тему можно поделить на три типа: построение устойчивой модели с помощью добавления новых типов генерации в фазу обучения, решение задачи с помощью методов, не использующих AI (с помощью классических методов и рассмотрения спектра света), создание новых более мощных датасетов для данный задачи.
AI-модели обучается на всё более новых и новых датасетах, включая в себя новые способы генерации, создаются способы онлайн-обучения, что улучшает постепенно качество, но концептуально не отличается от предыдущих методов и не обеспечивает устойчивость в случае, если появится более инновационный метод генерации. До появления диффузионных моделей высокое качество показывал метод, рассматривающий спектр по Фурье. Но на диффузионных моделях не показывает уже высокого качества.
Таким образом, в этой статье проводится попытка объединить существующие методы и найти новый способ детекции машиносгенерированных изображений. Новизна заключается в объединении методов и построении модели, предполагающей сначала тип генерации, а потом проверяющей на генерацию сгенерировано ли изображение уже непосредственно с предположением определенного типа генерации. Преимущество этого подхода заключается в подборе оптимальной модели для конкретного класса генирации, проблема заключается в высокой цене ошибки: если произойдет ошибка в предсказании класса генерации, то будет использоваться заведомо плохо подходящая модель
В качестве векторизатора мы используем предобученный, который используется во множестве разных исследований для разных целей и задач, в том числе используется в качестве векторизатора для задач классификации.

Primary authors: KILINKAROV, Georgii; Mr ANDREW, Grabovoy

Presenter: KILINKAROV, Georgii

Session Classification: 18-Машинное обучение и нейросети

Track Classification: Машинное обучение и нейросети