Секция «Круглый стол для школьников «Постигая аудит»»

Почему атаки на систему распознавания изображений представляют опасность?

Научный руководитель – Тарасов Иван Владимирович

Семёнова Виктория Андреевна
E-mail: tiwell23@rambler.ru

ГБОУ «Школа №556», Москва, Днепропетровская ул., 33А

Аннотация. Искусственный интеллект в области распознавания изображений активно развивается, но данная система очень уязвима к кибератакам.

Ключевые слова: распознавание изображений, нейросети, искусственный интеллект, киберпреступность, кибератаки, защита данных.

Why are attacks on the image recognition system dangerous?

Annotation. Artificial intelligence in the field of image recognition is actively developing, but this system is very vulnerable to cyber attacks.

Keywords: image recognition, neural networks, artificial intelligence, cybercrime, cyber attacks, data protection.

Введение

Система распознавания изображений является развивающейся во многих отраслях науки в оплате, в технике, в медицине несмотря на свою уязвимость, она все равно используется почти в любом устройстве, но проблема состоит в том, что взлом такой системы может повлечь за собой огромную опасность для наших данных, особенно биометрических. Но и в других отраслях самоуправляемый транспорт, постановка диагнозов, спасение людей может нести опасность не только для наших данных, но и для жизни.

Актуальность данной темы в настоящее время очень востребована, искусственный интеллект в области распознавания изображений активно развивается, но, к сожалению, данная система очень уязвима к кибератакам, из-за которых может произойти утечка данных пользователей и нарушение алгоритма системы распознавания изображений.

Гипотеза. Существуют методы защиты данных на начальном этапе - разработки нейронной сети распознавания изображений к кибератакам.

Цель. Защита персональных данных пользователей.

Задачи.

1) Классификация атак на систему распознавания изображений.
2) Способы защиты нейронных сетей.

Что представляет из себя система распознавания изображений?

Распознавание изображений с помощью компьютерного зрения - это информационная технология, созданная для получения и понимания фотографий реального мира и их преобразования в цифровую информацию для дальнейшей обработки и анализа.

Традиционный метод распознавания изображений происходит в несколько этапов:

I этап. Происходит корректировка контраста, яркости и обрезка изображения под фиксированный размер.

II этап. Осуществляется извлечение признаков изображения на объектах.

III этап. Определяется классификация объектов, сегменты, выделенные в прошлом этапе, классифицируются, по выбранным разработчиком признакам.

IV этап. Распознавание образов на основе базы данных системы и выдает нам нужный результат.

Кибератаки на систему распознавания изображений могут нести самые различные угрозы: от получения доступа к цифровым кошелькам на чужих телефонах, до успешного получения доступа к местам с повышенной степенью защиты, например, отели, бизнес-центры, больницы или для получения биометрических данных пользователя). Любые устройства контроля доступа, которые заменяют охранников на систему распознавания лиц потенциально подвержены риску.

Условно атаки на системы компьютерного зрения можно разделить на 3 группы:

1) алгоритм обучения или классификации;
2) на приложение;
3) на сенсоры.

Классификация атак

Изображения:

- состязательные атаки - атаки представляют собой искажения, которые заставляют систему сомневаться в результате и выдавать заведомо не верный результат.
- бэкдор атаки - атаки, производимые на процедуру обучения, при таком обучении закладываются примеры, которые в дальнейшем должны позволить управлять нейронной сетью.

Приложения:

- закладки (нейронные сети) - используются для работы компьютерного зрения, представляют собой весьма сложные структуры, поведение которых иногда достаточно сложно предсказать.

На написание собственного алгоритма работы сети может уйти много времени. Сегодня для решения этой проблемы есть специализированные библиотеки и даже готовые сети, которые нужно только обучить на своих данных. В такую сеть может быть заложено определенное поведение. Значение может повлиять на поворот руля в автомобиле или распознание лица с камеры наблюдения.

Виды сценариев, угрожающих безопасности пользователя

Кража модели:

- атака, которая скорее всего будет затрагивать непосредственно создателя алгоритма.

Существует 2 вида кражи:

1) Онлайн — атакующему нужно воспроизвести саму сеть на основании входных и выходных данных.
2) Офлайн — анализ проводится над приложением или устройством с применением подходов обратной разработки.

Атаки на сенсоры:

Класс атак, который предполагает физические повреждения или искажение данных, которые регистрируют сенсоры.

Какие существующие методы защиты информации можно применить к системам распознавания изображений?

На данный момент нейронные сети очень уязвимы к атакам, например, создав одну угрозу и распространив ее в широких кругах, ее можно будет использовать снова, потому что библиотека система распознавания изображений находится в открытом доступе на разных языках программировании - Open CV.

Полученные во время взлома данные могут использоваться в совершенно разных целях:

- создание фальсифицированного видео с реальными лицами.

- взлом данных системы таким образом, чтобы произошла подмена одного объекта на другой.

- разблокировке с помощью распознавания лиц, при взломе любой сможет воспользоваться атакованными устройствами

- нарушить работу самоуправляемого транспорта, за счет неправильного восприятия дорожных знаков.

Во многом эти проблемы решаются шифрованием и заменой файлов. Если информация хранится в базе данных, но при атаке с заменой изображения, если вы не обезопасили свою систему в начале работы, её придется устанавливать заново. Также защитить или проверить данные можно с помощью искусственного интеллекта, но проблема состоит в том, что этот же искусственный интеллект можно использовать против защиты.

Возможные пути развития безопасности нейронных сетей

Нейронные сети продолжат развиваться и в том числе смогут стать необходимой частью почти любого приложения, соответственно и повышение безопасности возможно. Развитие безопасности на всех этапах обработки данных из реального цифровые, а может использовать весь потенциал нейронных сетей. Развитие возможно несколькими способами, но, к сожалению, они имеют ограниченный потенциал.

Проведённые нами опросы подтвердили широкие охваты данной функции. Также, как оказалось оплатой/разблокировкой телефона с помощью распознавания лица в основном, используют пользователи IOS. И правда стоит заметить, что Face ID от Apple на данный момент является одним из сложных для взлома. Рис. 1. Используете ли вы систему распознавания лиц?

Во-первых, можно увеличить количество камер распознавания лиц, тогда в базах данных будет больше информации о реальных лицах, следовательно появится возможность распознавать, хотя бы предположительно фальсификацию, аналогично в других отраслях (медицине, транспорте). Во-вторых, хранить базы данных либо только в преобразованном виде, либо ограничить доступ к базам данным до минимума. Рис. 2. Какая операционная система у вашего смартфона?

Его камера проецирует несколько десятков тысяч точек на лице для создания «карты лица», а затем преобразует полученную информацию в математическое представление и сравниваются с базой данных о лице. Таким образом, эту систему становится тяжело взломать, так как, во-первых, данные никогда не покидают iPhone пользователя, а, во-вторых, все несколько десятков тысяч точек сохраняются лишь в математическом представлений, что также затрудняет взлом.

Другие системы пока только стремятся к этому и используют FaceLock, используя только снимки лица. Система распознавания используется не только при разблокировке

в наших устройствах, благодаря этой функции повышается эффективность поиска, систематизация наших фотографий.

Вывод. Нехватка новых наборов данных, высокая стоимость разработки, зависимость от старых систем и тенденция просто адаптировать старые наборы данных - все это усугубляет проблему, поэтому стоит не только повышать безопасность систем, но и учиться создавать полностью новые системы распознавания изображений, такие системы продолжают активно внедрятся в нашу жизнь.

Источники и литература

1) https://text.ru/rd/aHR0cHM6Ly9kZWVwZmFrZWNoYWxsZW5nZS5jb20vMjAyMS8xMS8yNi8xMTU0My8%3D [Дата обращения: 14.10.2022 в 16:01].
2) https://text.ru/rd/aHR0cHM6Ly9oYWJyLmNvbS9ydS9wb3N0LzU2Nzc2NC8%3D [Дата обращения: 16.10.2022 в 18:03].
3) https://habr.com/ru/company/neuronet/blog/592119/ [Дата обращения: 16.10.2022 в 19:47].
4) https://habr.com/ru/company/otus/blog/549780/?ysclid=l9lw438bja516216540 [Дата обращения: 26.10.2022 в 13:09].

Иллюстрации

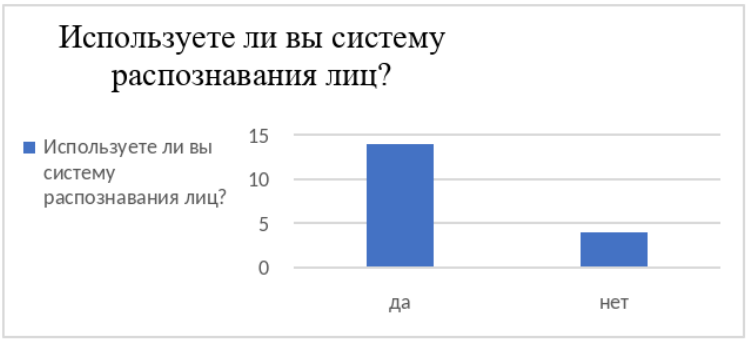

Рис. 1. Используете ли вы систему распознавания лиц?

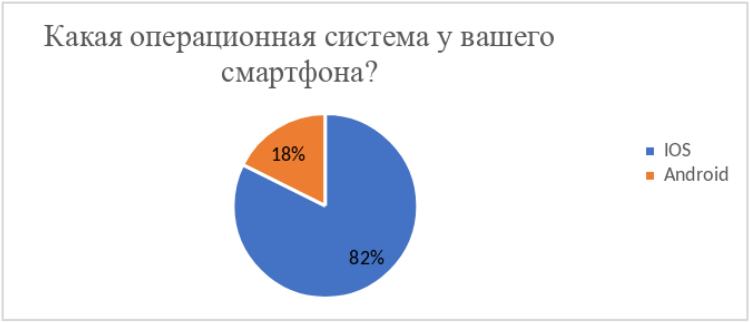

Рис. 2. Какая операционная система у вашего смартфона?