методов социальной инженерии. В первую очередь, фишинг-атаки контентно-центрированы и использование стандартных систем предотвращения утечек информации может оказать огромную услугу в раннем обнаружении и блокировке подобных попыток изменения чувствительных данных. Средства контроля съемных носителей и внешних устройств могут защитить интеллектуальную собственность организации только в том случае, если кто-то из ее сотрудников стал случайной жертвой подброшенного устройства с интересным содержимым.

Однако, самым эффективным методом будет программа обучения вопросам информационной безопасности. Сотрудники должны иметь в виду, что электронная почта компании предназначена для ведения бизнеса, тем самым ограничивая возможное попадание адресов электронной почты пользователей в Интернет. Сотрудники также должны быть внимательны при работе с вложениями из неизвестных источников. Организации также должны иметь формальную политику для использования социальных сетей, запрещающую или ограничивающую информацию, которую сотрудники могут размещать на своих личных страницах. Сайты социальных сетей представляют собой отличный инструмент для использования социального инжиниринга. Приведенные в порядок данные о компании, размещенные сотрудниками на сайтах социальных сетей, будут способствовать повышению уровня общей информационной безопасности компании.

1. Вести // <https://www.vedomosti.ru/finance/blogs/2018/09/25/781952-izbezhat-utechki-personalnoi/> [Электронный ресурс]. – Дата обращения 12.05.2018.
2. Социальная инженерия. Основные типы социальной инженерии // <https://efsol.ru/articles/social-engineering.html/> [Электронный ресурс]. – Дата обращения 12.05.2018.
3. Методы социальной инженерии или атаки на человеческий фактор // <https://www.a1qa.ru/blog/sotsialnaya-inzheneriya-ili-ataki-na-chelovecheskiy-faktor/> [Электронный ресурс]. – Дата обращения 12.05.2018.

Семенова А.Н., Аджиев Н.Д., Чочиева А.Н.

Сравнение современных методов детектирования объектов на изображениях

СКГА

(Россия, Черкесск)

doi: 10.18411/lj-11-2019-49

idsp: ljournal-11-2019-49

Аннотация

В современном мире системы видеонаблюдения становятся важной частью инфраструктуры городов и предприятий. Под системой видеонаблюдения понимается комплекс программных и аппаратных средств получения и анализа видео для помощи в принятии решения человеком. В настоящее время в большинстве случаев системы видеонаблюдения используются для видеофиксации событий с целью последующего анализа и разбора человеком-оператором, например, после возникновения какой-либо внештатной ситуации. Текущий уровень развития алгоритмов компьютерного зрения позволяет автоматизировать получение ответов на эти вопросы для ряда важных практических сценариев. Достижения в решении задач выделения автомобилей на дороге и их идентификации по номерному знаку позволили создать систему автоматической фиксации нарушений правил дорожного движения, принуждающую водителей им следовать. В последние несколько лет были разработаны эффективные алгоритмы выделения лиц людей в видео и идентификации человека по изображению лица.

Ключевые слова: алгоритмы, объекты, модель, методы, нейронные сети

Abstract

In today's world, video surveillance systems are becoming an important part of the infrastructure of cities and enterprises. A video surveillance system is understood as a complex of software and hardware for receiving and analyzing video to help a person make a decision. Currently, in most cases, video surveillance systems are used for video recording of events for the purpose of subsequent analysis and analysis by a human operator, for example, after the occurrence of an emergency situation. The current level of development of computer vision algorithms allows you to automate the receipt of answers to these questions for a number of important practical scenarios. Achievements in solving problems of distinguishing cars on the road and identifying them by license plate made it possible to create a system for automatically recording violations of traffic rules, forcing drivers to follow them. In the past few years, effective algorithms have been developed for highlighting people's faces in a video and for identifying a person by facial image.

Keywords: algorithms, objects, model, methods, neural networks

Способность "распознавать" считается основным свойством человеческих существ, как, впрочем, и других живых организмов. Распознавание образов - раздел кибернетики, разрабатывающий принципы и методы классификации, а также идентификации предметов, явлений, процессов, сигналов, ситуаций - всех тех объектов, которые могут быть описаны конечным набором некоторых признаков или свойств, характеризующих объект.

На протяжении десятилетий исследователи прокладывали путь в разработке передовых методов для повышения точности классификации. Традиционно, модели классификации могут хорошо работать только на небольших наборах данных, таких как CIFAR-10 и MNIST. Огромный скачок в развитии классификации изображений произошел, когда в 2009 году Feifei Li создал крупномасштабный набор данных изображений «ImageNet».

До появления методов глубокого обучения исследовательские работы прилагали много усилий для разработки масштабно-инвариантных функций (например, SIFT, HOG, GIST), представлений функций (например, Bag-of-Features), и классификаторы (например, SVM) для классификации изображений. Однако эти созданные вручную элементы работают с объектами на естественных изображениях со сложным фоном, различными цветами, текстурой, освещенностью и постоянно меняющимися позами и факторами обзора. На крупномасштабном конкурсе визуального распознавания ImageNet (ILSVRC) в 2012 году AlexNet выиграл первый приз со значительным отрывом от второго места, основанного на SIFT и векторах Фишера (FV). Это демонстрирует, что классификационная модель, основанная на глубоком CNN (сверточная нейронная сеть), работает намного надежнее, чем другие традиционные методы при наличии крупномасштабных вариаций. Это также представляет собой замечательную веху в современной истории нейронной сети после длительного периода стагнации.

Типичная глубокая модель CNN состоит из нескольких слоев свертки, за которыми следуют функции активации и пулы, и несколько полностью связанных слоев перед прогнозированием.

AlexNet имеет 8 слоев свертки, 3 пула и 3 полностью связанных слоя с общим количеством параметров 60 миллионов. Он успешно использует ReLU в качестве функции активации вместо сигмоида. Кроме того, увеличение данных и отсев широко используются сегодня в качестве эффективных стратегий обучения. Следовательно, AlexNet известен как фундаментальная работа современной глубокой CNN.

Вдохновленные AlexNet, VGGNet и GoogleNet сосредоточены на разработке более глубоких сетей для дальнейшего повышения точности. Они заняли второе место и победили в ILSVRC в 2014 году соответственно. Путем многократного объединения сверточных ядер 3×3 и максимального пула 2×2 VGGNet успешно создает сверточную нейронную сеть из 16–19 слоев. GoogleNet имеет 22 слоя, но его операции с плавающей запятой и количество параметров намного меньше, чем у AlexNet и VGGNet, так как они удаляют полностью связанные слои и оптимизируют операции разреженных матриц.

Хотя более глубокие сети предлагают лучшую точность, простое увеличение количества слоев не может постоянно улучшать точность из-за исчезновения/взрыва информации о градиенте во время обучения сети. ResNet, который делает еще один значительный прогресс в глубокой сетевой структуре, предлагает использовать быстрое соединение между остаточными блоками, чтобы в полной мере использовать информацию из предыдущих уровней и сохранять градиенты во время обратного распространения. Используя этот остаточный блок, ResNet успешно обучает очень глубокие сети до 152 уровней и стал победителем ILSVRC в 2015 году. Следуя идее ResNet, DenseNet устанавливает соединения между всеми предыдущими уровнями и текущим уровнем. Он объединяет и, следовательно, повторно использует элементы всех предыдущих слоев. DenseNet представляет большое преимущество в точности классификации на ImageNet, видно в Таблице 1. Соединение различных сетевых уровней показало многообещающее улучшение в обучении представлений более глубоких сетей.

Таблица 1

Обзор различных моделей CNN в задаче классификации ImageNet.

Модель	Год	Точность	Кол-во параметров	Кол-во FLOPs	Кол-во слоев
FE-Net [33]	2019	75.0%	5.9 М	563 М	–
ShiftNet-A [32]	2018	70.1%	4.1 М	1,400 М	–
ShuffleNet [31]	2018	73.7%	4.7 М	524 М	–
MobileNet [30]	2017	70.6%	4.2 М	569 М	28
SqueezeNet [29]	2016	57.5%	1.2 М	833 М	–
NASNet [27]	2018	82.7%	88.9 М	23,800 М	–
SENet [26]	2017	82.7%	145.8 М	42,300 М	–
DenseNet [25]	2017	79.2%	25.6 М	1,150 М	190
ResNet [24]	2015	78.6%	55 М	2,300 М	152
GoogleNet [23]	2014	69.8%	6.8 М	1,500 М	22
VGGNet [22]	2014	71.5%	138 М	15,300 М	16
AlexNet [21]	2012	57.2%	60 М	720 М	8

Вышеупомянутые сетевые модели кратко суммированы в Таблице 1. В дополнение к обычной проблеме классификации изображений с тысячами классов и сложных сцен, многокомпонентной классификации (например, распознавания лиц конкретных людей) представляют большой интерес в области компьютерного зрения.

Кроме того, большой успех глубокого обучения в области изображений стимулировал различные методы для изучения надежных представлений признаков для

классификации видео, где семантическое содержание, такое как действия человека или сложные события, автоматически классифицируется. Ранние работы часто рассматривают видеоклип как набор кадров. Классификация видео осуществляется путем агрегирования признаков CNN на уровне кадра путем усреднения или кодирования. Стандартные классификаторы, такие как SVM, используются для распознавания.

В отличие от методов классификации на уровне кадров, существует ряд других подходов, использующих сквозные модели CNN для изучения скрытых пространственно-временных моделей в видео. Например, типичные особенности C3D получены из глубокой трехмерной сверточной сети, обученной по крупномасштабному набору данных UCF101. Более того, двухпоточковый подход предлагается разделить проблему обучения представлению видео на пространственные и временные сигналы отдельно. В частности, пространственное CNN используется для моделирования информации о внешнем виде из кадров RGB, в то время как временное CNN используется для изучения информации о движении из плотного оптического потока среди соседних кадров.

Поскольку двухпоточковый подход отображает только движения в пределах короткого временного окна и не учитывает временной порядок различных кадров, несколько моделей периодических соединений для последовательных данных, включая модели рекуррентных нейронных сетей (RNN) и модели с кратковременной памятью (LSTM), используются для моделирования временной динамики для видео.

С появлением больших данных, продвинутых алгоритмов глубокого обучения и мощных аппаратных ускорителей современные системы компьютерного зрения претерпели значительные изменения. В будущем все более компактные и эффективные DNN необходимы для приложений реального времени и встроенных приложений. Кроме того, должны быть исследованы слабо контролируемые или неконтролируемые схемы DNN, чтобы воспринимать все категории объектов во всех сценах открытого мира. Кроме того, высокоэнерго-эффективные аппаратные движки необходимы для расширения существующих ускорителей в широком спектре сложных сценариев. По сравнению с линейными методами статистики, нейросети позволяют эффективно строить нелинейные зависимости, более точно описывающие наборы данных. Что же касается байесовского классификатора, строящего квадратичную разделяющую поверхность, нейронная сеть может построить поверхность более высокого порядка. Высокая нелинейность разделяющей поверхности наивного байесовского классификатора (он не использует ковариационные матрицы классов, как классический Байес, а анализирует локальные плотности вероятности) требует значительного суммарного числа примеров для возможности оценивания вероятностей при каждом сочетании интервалов значений переменных – тогда как нейросеть обучается на всей выборке данных, не фрагментируя её, что повышает адекватность настройки сети. В данной статье были изучены методы, применяемые для распознавания изображений, которые привели к выводу, что выбор метода в первую очередь зависит от исходной задачи, однако нейронные сети являются наиболее перспективным направлением в данный момент и покрывают наибольшее количество задач.

1. Л. П. Попова, И. О. Датъев. Обзор существующих методов распознавания образов. — М.: Сборник научных трудов, 2007. — 11 с.
2. Местецкий Л. В. Математические методы распознавания образов. Курс лекций.. — М.:, 2004. — 85 с.
3. Кузин Л.Т. Основы кибернетики: Основы кибернетических моделей. Т.2. - М.: Энергия, 1979