

Видеоаналитика и ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

О средствах аналитики на базе машинного и
глубокого обучения

Март 2021

Содержание

1	Краткая информация	3
2	Введение	4
3	Искусственный интеллект, машинное обучение и глубокое обучение	4
3.1	Машинное обучение	5
3.2	Глубокое обучение	6
3.3	Сравнение классического машинного обучения с глубоким обучением	7
4	Этапы машинного обучения	7
4.1	Сбор и аннотирование данных	7
4.2	Обучение	8
4.3	Тестирование	9
4.4	Развертывание	9
5	Встроенная аналитика	9
6	Аппаратное ускорение	10
7	Искусственный интеллект все еще находится на ранней стадии развития	10
8	Факторы, определяющие эффективность аналитики	11
8.1	Пригодность изображения	11
8.2	Расстояние обнаружения объектов	12
8.3	Настройка сигналов тревоги и записи	13
8.4	Обслуживание	13
9	Конфиденциальность и неприкосновенность личности	14
10	Приложение	16
10.1	Нейронные сети	16
10.2	Сверточные нейронные сети	17

1 Краткая информация

Видеоаналитика на базе искусственного интеллекта (ИИ) — одна из самых обсуждаемых тем в индустрии охранного видеонаблюдения. Некоторые приложения могут существенно ускорять анализ данных и автоматизировать повторяющиеся задачи. Однако на сегодняшний день системы ИИ не могут заменить человека, который принимает решения с учетом своих знаний и опыта. В такой ситуации успех принесет использование преимуществ систем ИИ для увеличения потенциала и возможностей человека.

Концепция искусственного интеллекта включает в себя алгоритмы машинного и глубокого обучения. Они автоматически создают математическую модель, используя значительные объемы выборочных (*обучающих*) данных, чтобы получить возможность вычислять результаты без специально созданной программы. Алгоритм ИИ разрабатывается посредством итеративного процесса, в котором цикл сбора обучающих данных, их маркировки, использования маркированных данных для обучения алгоритма и тестирования обученного алгоритма повторяется до тех пор, пока не будет достигнут желаемый уровень качества. После этого алгоритм готов к использованию в приложении обработки и анализа данных, которое можно приобрести и внедрить на объекте наблюдения. На этом обучение завершено, и приложение больше ничему новому не научится.

Типичной задачей видеоаналитики на базе искусственного интеллекта является обнаружение и классификация объектов (людей и транспортных средств) в видеопотоке. Алгоритм *машинного обучения* изучает комбинацию визуальных характеристик, позволяющую определить эти объекты. Алгоритм *глубокого обучения* более совершенен и может — если он был этому обучен — обнаруживать гораздо более сложные объекты. Однако он также требует больше усилий для разработки и обучения и гораздо больше вычислительных ресурсов при использовании окончательно доработанного приложения. Это значит, что следует четко определить потребности организации в видеонаблюдении и подумать, будет ли достаточно специализированного оптимизированного приложения на базе машинного обучения.

Разработка алгоритмов и расширение возможностей камер по обработке данных позволили использовать усовершенствованную видеоаналитику на основе ИИ непосредственно в камере, вместо того, чтобы выполнять все вычисления на сервере. Это означает возможность анализировать получаемые данные в реальном времени, ведь приложения могут сразу же использовать несжатый видеоматериал. Благодаря выделенным аппаратным ускорителям, таким как модуль обработки данных с технологией машинного обучения (MLPU) и модуль обработки данных с технологией глубокого обучения (DLPU), встроенная в камеры видеоаналитика требует меньших энергозатрат, чем использование центрального или графического процессора.

Перед установкой приложения для интеллектуального анализа видео необходимо внимательно изучить и строго соблюдать рекомендации производителя, основанные на известных предварительных условиях и ограничениях. Все системы видеонаблюдения уникальны, и эффективность приложения должна оцениваться на каждом конкретном объекте. Если качество окажется ниже ожидаемого, необходимо искать причину в комплексе, а не сосредотачиваться только на аналитическом приложении. Производительность видеоаналитики зависит от множества факторов, связанных с оснащением и конфигурацией камеры, качеством и динамическими характеристиками видеоизображения, а также с освещенностью в кадре. В большинстве случаев знание о влиянии этих факторов и их соответствующая оптимизация позволяют повысить эффективность видеоаналитики.

Поскольку искусственный интеллект все чаще применяется в видеонаблюдении, помимо преимуществ и новых сценариев его использования необходимо учитывать, когда и где будет применяться эта технология.

2 Введение

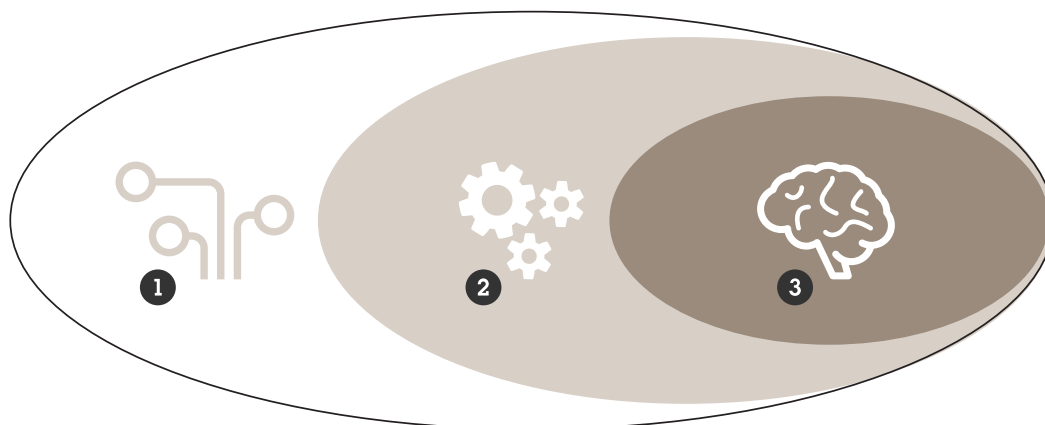
Искусственный интеллект (ИИ) разрабатывают и обсуждают с момента изобретения первых компьютеров. Хотя до полноценного искусственного интеллекта еще очень далеко, уже сейчас технологии с его элементами широко используются для выполнения четко определенных задач в таких приложениях и системах, как распознавание голоса, поисковые системы и виртуальные помощники. Кроме того, ИИ все чаще применяется в медицине, например в рентгеновской диагностике и для анализа результатов сканирования сетчатки.

Видеоаналитика на базе искусственного интеллекта — одна из самых обсуждаемых тем в индустрии охранного видеонаблюдения, и с ней связаны большие надежды. На рынке есть приложения, которые используют алгоритмы ИИ для значительного ускорения анализа данных и автоматизации повторяющихся задач. Но в более широком контексте видеонаблюдения искусственный интеллект сегодня и в ближайшем будущем должен рассматриваться лишь как один из нескольких элементов процесса создания точных решений.

В этом техническом обзоре представлена справочная информация об алгоритмах машинного и глубокого обучения, а также о том, как их можно разрабатывать и применять для видеоаналитики. В документе кратко описывается оборудование для ускорения ИИ, приводятся плюсы и минусы аналитических приложений, установленных в камере, по сравнению с видеоаналитикой на сервере, и рассматривается, как можно оптимизировать предварительные условия для повышения производительности видеоаналитики на основе ИИ с учетом широкого спектра факторов.

3 Искусственный интеллект, машинное обучение и глубокое обучение

Искусственный интеллект (ИИ) — это широкое понятие, связанное с машинами, которые могут решать сложные задачи, имитируя интеллект человека. Глубокое обучение и машинное обучение — это подмножества искусственного интеллекта.



- 1 *Искусственный интеллект*
- 2 *Машинное обучение*
- 3 *Глубокое обучение*

3.1 Машинное обучение

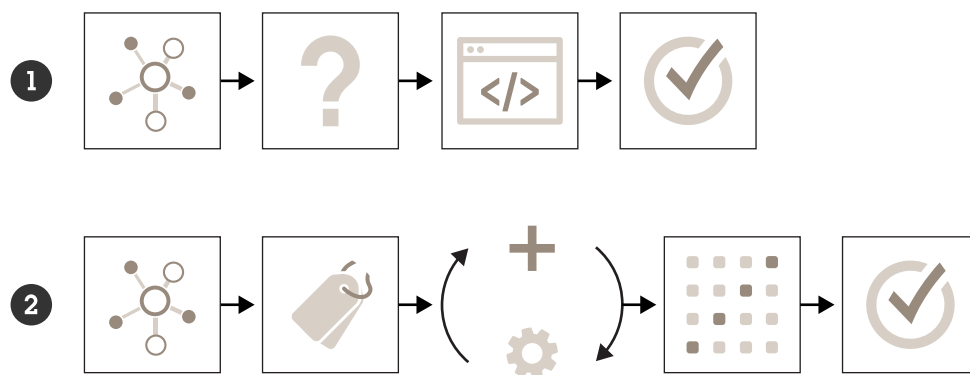
Машинное обучение – это подмножество ИИ, использующее статистические алгоритмы обучения для создания систем, которые могут автоматически учиться и совершенствоваться без явного программирования.

В этом разделе мы проводим различие между традиционным программированием и машинным обучением в контексте *компьютерного зрения* – дисциплины, позволяющей компьютерам понимать, что происходит в кадре, путем анализа фото- или видеоданных.

Традиционно программируемое компьютерное зрение основано на методах вычисления *признаков* изображения, например четких кромок и углов. Эти признаки должен задать вручную разработчик алгоритма, который знает, что именно нужно искать в визуальных данных. Затем разработчик объединяет эти признаки для создания алгоритма, который позволит сделать вывод о том, что обнаружено в кадре.

Алгоритмы машинного обучения автоматически создают математическую модель с использованием значительного количества выборочных (*обучающих*) данных, чтобы получить возможность принимать решения путем вычисления результатов без специального программирования. Признаки по-прежнему задаются вручную, но тому, как их объединять, алгоритм учится сам в процессе обработки больших объемов помеченных, или *аннотированных*, обучающих данных. Такая техника использования заданных вручную признаков в комбинациях, которым алгоритм обучается сам, в данном документе называется *классическим машинным обучением*.

Другими словами, в случае приложения машинного обучения нужно обучить компьютер использовать нужную программу. Данные собираются, а затем аннотируются людьми, иногда с помощью предварительного аннотирования, выполняемого серверными компьютерами. Результат передается в систему, и этот процесс продолжается до тех пор, пока приложение не узнает достаточно, чтобы обнаружить нужный объект, например конкретный тип автомобиля. Обученная модель становится программой. Обратите внимание: когда программа будет готова, система не научится ничему новому.



- 1 *Традиционное программирование:*
Выполняется сбор данных. Определяются критерии программы. Пишется программа (человеком). Готово.
- 2 *Машинное обучение:*
Выполняется сбор данных. Данные маркируются. Модель проходит итеративный процесс обучения. Доработанная обученная модель становится программой. Готово.

Преимущество ИИ перед традиционным программированием при создании программы компьютерного зрения заключается в способности обрабатывать обширные данные. Компьютер может просматривать тысячи изображений, постоянно сохраняя точность и высокую работоспособность, тогда как человек через некоторое время устанет и утратит концентрацию. Таким образом, используя искусственный интеллект, можно сделать приложение намного более точным. Однако чем сложнее приложение, тем труднее машине получить желаемый результат.

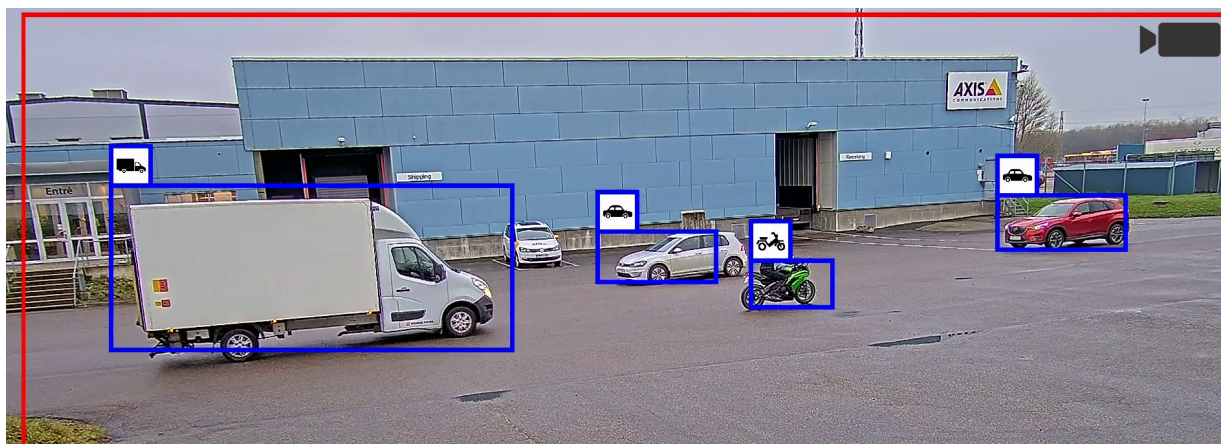
3.2 Глубокое обучение

Глубокое обучение представляет собой усовершенствованную версию машинного обучения, когда система обучается выделению признаков и способам объединения этих признаков в глубоких структурах правил для получения результата, и это обучение основано на входных данных. Алгоритм может автоматически определять, какие признаки следует искать в обучающих данных, а также может изучать очень глубокие структуры связанных комбинаций признаков.

Ядро алгоритмов, используемых в глубоком обучении, основано на том, как работают нейроны и как мозг использует их для формирования знаний более высокого уровня путем объединения выходных нейронных сигналов в глубокой иерархии, или *сети*, связанных правил. Мозг — это система, в которой комбинации также формируются нейронами, стирая различие между выделением и комбинированием признаков, делая эти процессы в некотором смысле одинаковыми. Ученые смоделировали эти структуры в так называемые *искусственные нейронные сети*, которые являются самым популярным типом алгоритмов глубокого обучения. См. краткий обзор нейронных сетей в приложении к этому документу.

Используя алгоритмы глубокого обучения, можно создавать сложные визуальные детекторы и автоматически обучать их обнаружению очень сложных объектов, устойчивых к масштабированию, вращению и прочим вариациям.

Причина такой гибкости заключается в том, что системы глубокого обучения могут учиться на гораздо большем объеме данных (к тому же характеризующихся более широким разнообразием), чем классические системы машинного обучения. В большинстве случаев они значительно превосходят алгоритмы компьютерного зрения, созданные вручную. Благодаря этому глубокое обучение идеально подходит для сложных задач, когда людям нелегко сформулировать комбинацию признаков (например, для классификации изображений, обработки лингвистической информации или обнаружения объектов).



Обнаружение объектов на основе глубокого обучения позволяет классифицировать сложные объекты. В этом примере аналитическое приложение может не только обнаруживать транспортные средства, но и классифицировать их по типам.

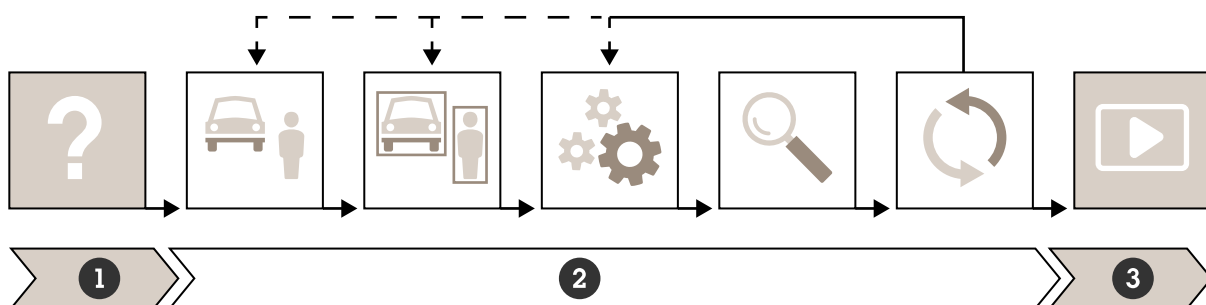
3.3 Сравнение классического машинного обучения с глубоким обучением

Оба вида обучения представляют собой схожие типы алгоритмов, но алгоритм глубокого обучения обычно использует гораздо больший набор выученных комбинаций признаков, чем классический алгоритм машинного обучения. Это означает, что аналитическое приложение на базе глубокого обучения может быть более гибким и — если его обучить — сможет выполнять гораздо более сложные задачи.

Впрочем, для отдельных приложений видеоаналитики может быть достаточно специально оптимизированного классического алгоритма машинного обучения. Для четко определенных задач этот алгоритм может выдавать такие же результаты, как и алгоритм глубокого обучения, при этом требуя меньше математических операций, что делает данное решение более экономичным и менее энергозатратным. Кроме того, для него необходимо гораздо меньше данных для обучения и, следовательно, ресурсов на разработку.

4 Этапы машинного обучения

Прежде чем развернуть готовое аналитическое приложение, алгоритм машинного обучения должен пройти серию этапов разработки и итераций, схематично изображенных ниже. В основе аналитического приложения лежит один или несколько алгоритмов, например детектор объектов. В случае приложений глубокого обучения ядром алгоритма является модель глубокого обучения.



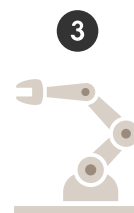
- 1 Подготовка: определение цели приложения.
- 2 Обучение: сбор обучающих данных; аннотирование данных; обучение модели; тестирование модели. Если качество не соответствует ожидаемому, предыдущие этапы повторяются в цикле итеративного улучшения.
- 3 Развертывание: установка и использование готового приложения.

4.1 Сбор и аннотирование данных

Для разработки аналитического приложения на основе ИИ необходимо собрать большие объемы данных. В охранном видеонаблюдении данные представляют собой отдельные снимки и видеоматериалы с изображением людей, транспортных средств или других объектов. Чтобы машина или компьютер могли распознать данные, требуется процесс аннотирования данных, то есть классификация и маркировка соответствующих объектов. Аннотирование данных — это в основном ручная и очень трудоемкая задача. Подготовленные данные должны охватывать достаточно большое количество разнообразных выборок, актуальных для контекста, в котором будет использоваться аналитическое приложение.

4.2 Обучение

Обучение представляет собой процесс, в ходе которого в модель загружаются аннотированные данные и используется схема обучения с итеративным изменением и совершенствованием модели до тех пор, пока не будет достигнуто желаемое качество. Другими словами, в ходе обучения модель оптимизируется для решения поставленной задачи. Обучение можно проводить одним из трех основных способов.



- 1 *Контролируемое обучение: модель учится делать точные прогнозы*
- 2 *Неконтролируемое обучение: модель учится идентифицировать кластеры*
- 3 *Обучение с подкреплением: модель учится на ошибках*

4.2.1 Контролируемое обучение

На сегодняшний день чаще всего в машинном обучении применяется контролируемое обучение, которое можно описать как обучение на примерах. Обучающие данные четко аннотированы, что означает, что входные данные уже соотнесены с желаемым выходным результатом.

Для контролируемого обучения обычно требуется очень большой объем аннотированных данных, а производительность обученного алгоритма напрямую зависит от качества обучающей выборки. Наиболее важным аспектом качества является использование набора данных, в котором представлены все потенциальные входные данные из реальной ситуации применения алгоритма. Что касается обнаружения объектов, разработчик обязан обучать алгоритм с использованием широкого спектра изображений, различных примеров объектов и условий, к которым относятся положение объектов, масштаб, освещение, фон и отвлекающие факторы. Готовое аналитическое приложение сможет давать точные прогнозы при обработке новых, не использовавшихся на этапе обучения, данных, только если обучающие данные являлись репрезентативными для планируемого сценария использования.

4.2.2 Неконтролируемое обучение

При неконтролируемом обучении используются алгоритмы для анализа и группировки немаркированных наборов данных. Этот метод обучения не очень популярен в индустрии видеонаблюдения, поскольку данная модель требует большого количества калибровок и проверок, при этом качество может быть непредсказуемым.

Наборы данных должны соответствовать аналитическому приложению, но не должны иметь четкую маркировку или разметку. При этом устраняется необходимость создавать аннотации вручную, но количество изображений или видеоматериалов, необходимых для обучения, должно быть увеличено на несколько порядков. На этапе обучения обучаемая модель выявляет в наборах данных общие признаки, что позволяет ей на этапе развертывания группировать данные в соответствии с шаблонами, а также обнаруживать аномалии, которые не вписываются ни в одну из изученных групп.

4.2.3 Обучение с подкреплением

Обучение с подкреплением используется, в частности, в робототехнике, автоматизации производства и планировании бизнес-стратегии, но из-за необходимости большого объема обратной связи этот метод в индустрии видеонаблюдения применяется нечасто. Обучение с подкреплением связано с выполнением подходящего действия, чтобы получить максимально возможное в конкретной ситуации *вознаграждение*, которое увеличивается по мере того, как модель делает правильный выбор. Алгоритм не использует пары данных/меток для обучения, а вместо этого оптимизируется путем тестирования своих решений во взаимодействии с окружающей средой. Цель алгоритма – обучиться таким действиям, которые помогут максимально увеличить ожидаемое вознаграждение.

4.3 Тестирование

После обучения модель необходимо тщательно протестировать. Этот этап обычно включает автоматизированную часть и тщательное тестирование в реальных условиях развертывания.

В ходе автоматизированной части приложение проверяется с помощью новых наборов данных, которые не использовались для обучения модели. Если результаты таких проверок не соответствуют ожидаемым, процесс начинается заново: собираются новые обучающие данные, создаются или уточняются аннотации, и модель снова подвергается обучению.

После достижения нужного уровня качества начинается испытание в условиях эксплуатации: приложение проверяется в ситуациях из реальной жизни. Объем и вариантность зависят от области применения приложения. Чем она меньше, тем меньше вариаций нужно тестировать. И наоборот, чем шире область применения, тем больше требуется тестов.

Результаты вновь сравниваются и оцениваются. Этот этап может снова привести к тому, что процесс начнется заново. Другой возможный результат может заключаться в определении предварительных условий, объясняющих сценарий, в котором приложение не рекомендуется использовать или рекомендуется использовать лишь частично.

4.4 Развертывание

Этап развертывания также называется этапом вывода или прогнозирования. *Получение вывода или прогнозирования* – это процесс выполнения обученной модели. Алгоритм использует то, чему он обучился на этапе обучения, для получения желаемого результата. В контексте аналитики для видеонаблюдения этап вывода – это приложение, используемое в системе для видеонаблюдения за объектами в реальных условиях.

Для достижения производительности в реальном времени при выполнении алгоритма на базе машинного обучения для входных аудио- или видеоданных обычно требуется специальное аппаратное ускорение.

5 Встроенная аналитика

Раньше высокопроизводительная видеоаналитика выполнялась на сервере, поскольку требовала большей мощности и охлаждения, чем могла предложить камера. Но благодаря разработке алгоритмов и увеличению вычислительной мощности периферийных устройств в последние годы появилась возможность использовать расширенную видеоаналитику на основе ИИ в самих камерах.

У встроенной видеоаналитики есть очевидные преимущества: аналитические приложения имеют почти мгновенный доступ к несжатому видеоматериалу, что позволяет использовать приложения в реальном времени, избегая дополнительных затрат и сложностей, связанных с перемещением данных в облако для вычислений. Встроенная аналитика также требует меньше денежных средств на оборудование и развертывание, поскольку системе наблюдения необходимо меньше серверных ресурсов.

Для некоторых приложений может оказаться полезным сочетание двух видов обработки данных: предварительной — на камере и дополнительной — на сервере. Такая гибридная система может способствовать экономичному масштабированию аналитических приложений за счет работы с видеопотоками нескольких камер.

6 Аппаратное ускорение

Зачастую конкретное аналитическое приложение можно запускать на нескольких типах платформ, но использование выделенного аппаратного ускорения позволяет существенно повысить производительность в случае нехватки электроэнергии. Аппаратные ускорители могут обеспечить высокую энергоэффективность аналитических приложений. При необходимости их можно использовать совместно с серверными и облачными вычислительными ресурсами.

- **Графический процессор.** Графические процессоры в основном разрабатывались для приложений обработки графики, но также используются для ускорения ИИ на серверных и облачных платформах. Хотя иногда их также используют во встроенных системах, графические процессоры не являются оптимальными для задач логического вывода с точки зрения энергоэффективности.
- **MLPU (модуль обработки данных с технологией машинного обучения).** Модуль MLPU может ускорить получение вывода в определенных классических алгоритмах машинного обучения для решения задач компьютерного зрения с очень высокой энергоэффективностью. Модуль предназначен для одновременного обнаружения в режиме реального времени ограниченного числа типов объектов (например, людей и транспортных средств).
- **DLPU (модуль обработки данных с технологией глубокого обучения).** Камеры со встроенным модулем DLPU могут ускорять получение общего вывода алгоритма глубокого обучения с высокой энергоэффективностью, что позволяет детальнее классифицировать объекты.

7 Искусственный интеллект все еще находится на ранней стадии развития

Представляется заманчивым провести сравнение между потенциалом системы на базе искусственного интеллекта и возможностями человека. Операторы систем видеонаблюдения способны сохранять бдительность и внимательность только в течение короткого времени, компьютер же может обрабатывать большие объемы данных неустанно и очень быстро. Однако было бы глубочайшим заблуждением думать, что искусственный интеллект может заменить человека. Успех заключается в использовании преимуществ систем ИИ для увеличения потенциала и возможностей оператора.

Машинное и глубокое обучение зачастую описывают как возможность системы автоматически учиться и совершенствоваться на основе полученного опыта. Но доступные на сегодняшний день системы ИИ *не* обучаются новым навыкам самостоятельно и *не* запоминают произошедшие события. Чтобы повысить эффективность системы, ее необходимо переобучить, используя более точные данные во время сеансов контролируемого обучения. Неконтролируемое обучение обычно требует

большого количества данных для создания кластеров и поэтому не используется в приложениях видеонаблюдения. Зато его применяют для анализа больших наборов данных при поиске аномалий, в частности в финансовых транзакциях. Большинство подходов, которые в индустрии видеонаблюдения рекламируются как «самообучение», основаны на анализе статистических данных, а не на фактическом переобучении моделей глубокого обучения.

Человеческий опыт по-прежнему превосходит большинство аналитических приложений на основе ИИ, используемых для целей видеонаблюдения. Особенно когда такие приложения должны выполнять очень общие задачи и когда понимание контекста имеет решающее значение. Приложение на основе машинного обучения может успешно обнаружить бегущего человека (если оно было специально обучено этому), но в отличие от человека, который способен поместить данные в контекст, приложение не понимает, почему человек бежит: чтобы успеть на автобус или чтобы его не догнал полицейский? Что бы ни обещали компании, применяющие ИИ в своих приложениях видеоаналитики, эти приложения не способны понимать, что происходит в кадре, с такой же проницательностью, как человек.

По этой же причине аналитические приложения на базе искусственного интеллекта могут выдавать ложную тревогу или, в случае реального происшествия, не подавать сигналы. Обычно такое происходит в сложной среде с интенсивным движением или когда человек несет крупный предмет, из-за которого приложение не может правильно классифицировать объект.

Аналитика на базе ИИ на современном этапе должна использоваться как вспомогательный инструмент, например чтобы определить, насколько важен инцидент, прежде чем оповестить оператора, чтобы тот решил, как реагировать. Таким образом, искусственный интеллект используется для обеспечения масштабируемости, а задача человека — оценивать потенциальные инциденты.

8 Факторы, определяющие эффективность аналитики

Чтобы знать, какое качество можно ожидать от аналитического приложения на базе ИИ, рекомендуется внимательно изучить и понять известные предварительные условия и ограничения, которые обычно указываются в документации к приложению.

Все системы видеонаблюдения уникальны, и эффективность приложения должна оцениваться на каждом конкретном объекте. Если качество не соответствует ожидаемому или прогнозируемому уровню, настоятельно рекомендуется искать причину не только в самом приложении. Необходимо выяснить причину в комплексе, поскольку эффективность аналитического приложения зависит от множества факторов, большинство из которых можно оптимизировать, если знать о том, как они влияют на качество. К таким факторам относятся оснащение камеры, качество и динамические характеристики видеоизображения, уровень освещения, а также конфигурация, положение и направление камеры.

8.1 Пригодность изображения

Часто говорят, что качество изображения зависит от высокого разрешения и высокой светочувствительности камеры. Несмотря на то, что важность этих параметров не подлежит сомнению, существуют и другие факторы, не менее важные при определении *пригодности* фото- или видеозаписи. Так, например, видеопоток непревзойденного качества с дорогостоящей камеры видеонаблюдения может оказаться абсолютно бесполезным из-за недостаточной подсветки в ночное время, перенаправления камеры в ненужную сторону или сбоя связи с системой.

Перед развертыванием аналитического приложения необходимо тщательно изучить место размещения камеры. Чтобы видеоаналитика работала надлежащим образом, камера должна быть расположена таким образом, чтобы зона наблюдения просматривалась полностью и беспрепятственно.

Пригодность изображения также может зависеть от сценария его использования. Качество видео, приемлемое для человеческого глаза, может быть недостаточно для приложения видеоаналитики. Более того, многие методы обработки изображений, обычно используемые для улучшения восприятия видео человеком, не рекомендуются при использовании видеоаналитики. К ним, в частности, относятся методы шумоподавления, технология широкого динамического диапазона и алгоритмы автоматического управления экспозицией.

Современные видеокамеры часто оснащены ИК-подсветкой, благодаря которой они могут работать в полной темноте. В этом есть положительный момент, ведь такая функция позволяет размещать камеры в труднодоступных местах и устраняет необходимость установки дополнительного освещения. Однако, если на объекте возможны сильные дожди или снегопад, настоятельно рекомендуется не полагаться на свет, исходящий от камеры или из источника, расположенного очень близко к ней. Капли дождя и снежинки могут отражать в камеру слишком много света и существенно затруднить аналитику. С другой стороны, при рассеянном свете больше шансов, что аналитика будет работать результативно даже в плохую погоду.

8.2 Расстояние обнаружения объектов

Определить максимальное расстояние обнаружения объектов аналитическим приложением на базе ИИ сложно: указанное в спецификации значение в метрах или футах не всегда соответствует истине. Нужно помнить, что на расстояние обнаружения существенно влияют качество изображения, характеристики места наблюдения, погодные условия и свойства объекта, в частности цвет и яркость. Очевидно, например, что яркий объект на темном фоне в солнечный день может быть обнаружен с гораздо большего расстояния, чем темный объект в дождливую погоду.

Расстояние обнаружения также зависит от скорости перемещения объектов. Для получения точных результатов приложению видеоаналитики необходимо «видеть» объект достаточно долго. Это время зависит от эффективности обработки (частоты кадров) платформы: чем она ниже, тем дольше объект должен находиться в кадре, чтобы его можно было обнаружить. Если выдержка камеры не соответствует скорости движения объекта, точность обнаружения может пострадать из-за размытости изображения.

Быстрые объекты могут оказаться необнаруженными, если они перемещаются вблизи камеры. Например, легко обнаружить человека, бегущего достаточно далеко от камеры, в то время как человек, бегущий с такой же скоростью очень близко к камере, может войти в сектор обзора и выйти из него так быстро, что сигнал тревоги не успеет сработать.

Для аналитических приложений, основанных на обнаружении движения, еще одну проблему представляют объекты, движущиеся непосредственно к камере или от нее. Особенно трудно обнаруживать медленно движущиеся объекты, которые вызывают лишь очень небольшие изменения в изображении по сравнению с движением через весь кадр.

Как правило, более высокое разрешение камеры не означает большее расстояние обнаружения. Возможности обработки, необходимые для выполнения алгоритма машинного обучения, пропорциональны размеру входных данных. Это означает, что для анализа полного разрешения камеры 4K требуется как минимум в четыре раза больше вычислительной мощности, чем для камеры с разрешением 1080p. Очень часто из-за ограничений в возможностях обработки камерой в

приложениях на основе ИИ используют более низкое разрешение, чем может предложить камера или видеопоток.

8.3 Настройка сигналов тревоги и записи

Из-за различных уровней применяемых фильтров аналитические приложения для обнаружения и классификации объектов генерируют очень мало ложных срабатываний. Однако такие приложения работают должным образом только при соблюдении всех указанных предварительных условий. В противном случае они могут пропустить важные события.

Если нет полной уверенности в том, что все условия будут выполняться во всех без исключения случаях, рекомендуется использовать консервативный подход и настроить систему таким образом, чтобы конкретная классификация объектов не являлась единственной причиной срабатывания сигнала тревоги. Такая настройка вызовет больше ложных срабатываний, но уменьшит риск пропустить важное событие. Когда сигналы тревоги или иницирующие их срабатывание данные поступают непосредственно в пункт охранной сигнализации, каждая ложная тревога оборачивается большими расходами. Совершенно очевидно, что необходима надежная классификация объектов, позволяющая отфильтровывать нежелательные сигналы тревоги. Однако механизм записи может и должен быть настроен так, чтобы полагаться не только на классификацию объектов. В случае пропущенного реального сигнала тревоги эта настройка позволяет оценить по записи причину пропуска, а затем внести изменения в монтаж и конфигурацию всей системы.

Если классификация объектов выполняется на сервере во время поиска инцидента, рекомендуется настроить систему на непрерывную запись и вообще не фильтровать исходную запись. Непрерывная запись занимает много места, но это в некоторой степени компенсируется современными алгоритмами сжатия, такими как Zipstream.

8.4 Обслуживание

Системы видеонаблюдения необходимо регулярно обслуживать. Рекомендуется не только просматривать картинку через интерфейс ПО для управления видео, но и проводить физический осмотр оборудования, чтобы своевременно обнаруживать и удалять все, что может мешать наблюдению или блокировать поле обзора. Это важно в отношении стандартных (только ведущих запись) систем, но еще более важно при использовании средств аналитики.

Если говорить о базовом видеообнаружении движения, то такая типичная помеха, как паутина, качающаяся на ветру, может увеличить количество сигналов тревоги, что приведет к более высокому потреблению дискового пространства. При использовании приложения для обнаружения и

классификации объектов паутина создаст зону исключения в области обнаружения. Ее нити будут скрывать объекты и значительно снизят вероятность их обнаружения и классификации.



Паутина может мешать обзору камеры видеонаблюдения.

Грязь на объективе или защитном колпаке камеры вряд ли создаст проблемы днем. Но в условиях низкой освещенности свет, падающий на грязный колпак сбоку (например, от фар автомобиля), может вызвать непредвиденное отражение, что снизит точность обнаружения.

Не менее важно проводить регулярное обслуживание зоны видеонаблюдения, ведь в течение срока службы камеры в секторе ее наблюдения может произойти многое. Выявить потенциальные проблемы позволит простое сравнение изображений местности «до и после». Как выглядела зона видеонаблюдения в момент установки камеры, и как она выглядит сегодня? Есть ли необходимость в корректировке зоны обнаружения? Нужно ли скорректировать поле обзора камеры или переместить камеру в другое место?

9 Конфиденциальность и неприкосновенность личности

Работа с системами безопасности и видеонаблюдения требует баланса между правом человека на конфиденциальность и неприкосновенность личности и стремлением повысить безопасность за счет предотвращения преступлений или проведения судебно-медицинской экспертизы. Каждый случай установки и использования приложения требует тщательного рассмотрения с точки зрения этики и местного законодательства. Также необходимо учитывать, например, требования по обеспечению кибербезопасности и предотвращению непреднамеренного доступа к видеоматериалам. В то же время встроенная аналитика и создание метаданных для статистических целей могут усилить защиту конфиденциальности, если для последующей обработки передаются только анонимные данные.

С ростом применения автоматизированной аналитики в системах видеонаблюдения необходимо учитывать некоторые новые аспекты. Поскольку использование аналитических приложений сопряжено с риском ложных обнаружений, важно, чтобы в процессе принятия решения участвовал опытный оператор или пользователь. Такой режим часто называют «оператор в контуре управления». Более того, важно понимать, что на решение человека может повлиять способ генерации и подачи сигнала тревоги. Без надлежащего обучения и понимания функционала аналитического приложения можно сделать неверные выводы.

Дополнительные проблемы могут возникнуть из-за способа разработки алгоритмов глубокого обучения — некоторые сценарии использования приложения требуют осторожности при применении данной технологии. Качество этих алгоритмов теснейшим образом связано с наборами данных, то есть с фото- и видеоматериалами, используемыми для обучения алгоритма. Тесты показывают, что если материалы не подвергать тщательному отбору, то некоторые системы ИИ при обнаружении объектов могут демонстрировать предвзятость по расовым и гендерным признакам. Это обстоятельство вызвало открытую дискуссию и привело как к законодательным ограничениям, так и к мерам по обеспечению того, чтобы такие аспекты учитывались во время разработки систем.

Учитывая, что искусственный интеллект все чаще применяется для охранного видеонаблюдения, помимо преимуществ и новых сценариев его использования необходимо учитывать, когда и где будет применяться эта технология.

10 Приложение

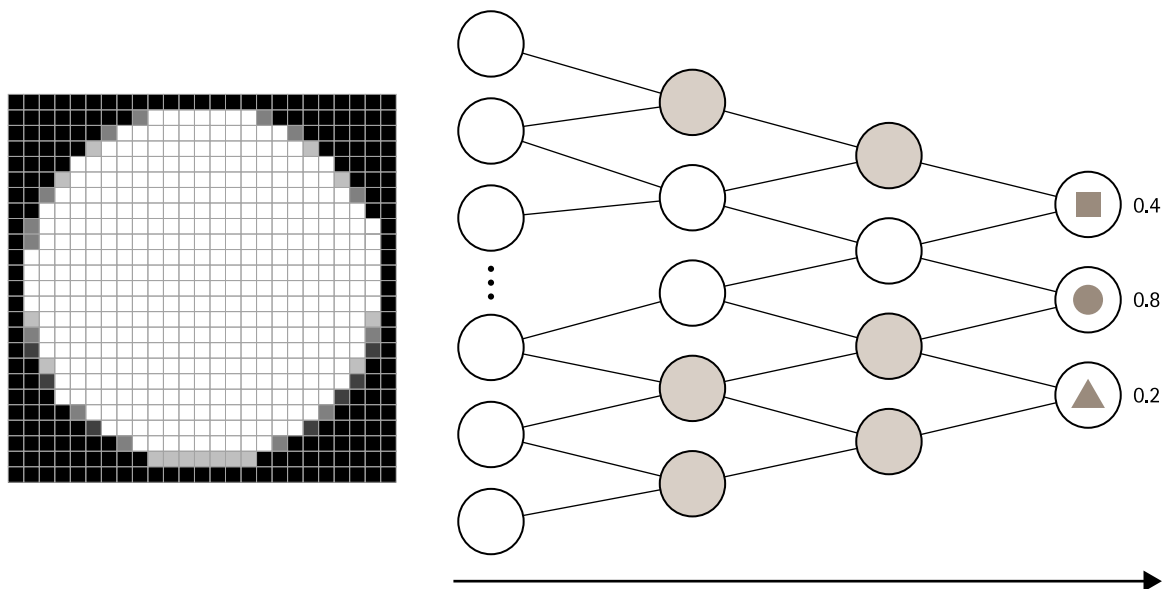
В приложении представлена справочная информация об искусственных нейронных сетях, составляющих основу глубокого обучения.

10.1 Нейронные сети

Нейронные сети — это семейство алгоритмов, которые используются для распознавания взаимосвязей в наборах данных посредством процесса, в некоторой степени похожего на работу человеческого мозга. Нейронная сеть состоит из иерархии множества слоев так называемых узлов, или нейронов, связанных между собой, и информация передается через сеть по соединениям от входного слоя к выходному.

Принцип работы нейронных сетей заключается в том, что выборка входных данных может быть уменьшена до конечного набора признаков, создавая достоверное представление входных данных. Затем эти признаки можно объединить для классификации входных данных, например описывая содержимое изображения.

На рисунке ниже показан пример, в котором нейронная сеть используется для определения того, к какому классу принадлежит входное изображение. Каждый пиксель изображения представлен одним входным узлом. Все входные узлы соединены с узлами первого слоя и производят выходные значения, которые передаются как входные значения на второй слой, где схема повторяется. В каждом слое в процессе также участвуют функции весов, значения смещения и функции активации.



Пример входного изображения (слева) и нейронной сети (справа). По достижении выходного слоя сеть определяет вероятности для каждой возможной категории (квадрат, круг или треугольник). Категория с наибольшим значением вероятности является наиболее вероятной формой входного изображения.

Этот процесс называется *прямым распространением*. В случае несовпадения результата прямого распространения параметры сети незначительно изменяются посредством *обратного распространения*. Во время этого итеративного процесса обучения производительность сети постепенно повышается.

Как правило, после развертывания нейронная сеть «не помнит» о предыдущих прямых проходах. Это означает, что сеть не улучшается со временем и может обнаруживать только те типы объектов или решать те типы задач, для которых она была обучена.

10.2 Сверточные нейронные сети

Сверточные нейронные сети (СНС) — это подвид искусственных нейронных сетей, которые особенно хорошо подходят для задач компьютерного зрения и лежат в основе быстрого прогресса глубокого обучения. В случае компьютерного зрения сеть обучается автоматически искать отличительные признаки изображения, например края, углы и цвета, по сути идентифицируя формы объектов.

При этом основной используемой операцией является математическая операция, называемая *сверткой*. Это очень эффективная операция, поскольку выход данных каждого узла зависит только от ограниченного окружения входных данных, созданного предыдущим слоем, а не от использования всего объема входных данных. Другими словами, в СНС каждый узел соединен не с каждым узлом на предыдущем уровне, а только с небольшим подмножеством. Свертки дополняются другими операциями, которые уменьшают размер данных, сохраняя при этом наиболее полезную информацию. Как и в стандартной искусственной нейронной сети, данные становятся тем более абстрактными, чем глубже они проникают в сеть.

На этапе обучения СНС обучается лучшему способу применять слои, а именно тому, как свертки должны комбинировать признаки из предыдущего слоя, чтобы выходные данные сети максимально совпадали с аннотациями обучающих данных. Во время получения вывода обученная сверточная нейронная сеть последовательно применяет слои сверток, которые были результатом обучения.

О компании Axis Communications

Компания Axis вносит весомый вклад в формирование более разумного и безопасного мира, разрабатывая решения, которые повышают безопасность и эффективность бизнеса. Занимая в отрасли технологий сетевого видео ведущие позиции, компания Axis предоставляет решения для видеонаблюдения, контроля доступа, сетевых домофонов и звукового сопровождения. Эффективность наших решений повышается благодаря приложениям интеллектуальной аналитики и высококачественному обучению.

Около 4000 специалистов компании Axis трудятся более чем в 50 странах мира, вместе с нашими партнерами по технологиям и по системной интеграции разрабатывая и внедряя решения задач, стоящих перед клиентами по всему миру. Компания Axis была основана в 1984 году. Штаб-квартира компании находится в городе Лунд, Швеция