



(19) **KG** (11) **432** (13) **C2** (46) **31.07.2025**

(51) **G06F 21/62** (2024.01)

МИНИСТЕРСТВО НАУКИ, ВЫСШЕГО ОБРАЗОВАНИЯ И ИННОВАЦИЙ КЫРГЫЗСКОЙ РЕСПУБЛИКИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ

к патенту Кыргызской Республики

(21) 20240043.1

(22) 07.08.2024

(46) 31.07.2025. Бюл. № 7

(71) (73) Общество с ограниченной ответственностью «Түлкү Мурун» (KG)

(72) Бородачев Евгений Владимирович (KG)

(56) Патент RU 2 710 942 C1, кл. G06N 3/063, G06N 3/08, G06K 9/62, 14.01.2020

(54) **Компьютерно-реализованный способ использования системы распознавания лиц**

(57) Данное изобретение в целом относится к области вычислительной техники, но не в порядке ограничения, а в частности относится к системам и способам распознавания лиц. Компьютерно-реализованный способ использования системы распознавания лиц для идентификации человека, выполняемый, по меньшей мере, одним вычислительным устройством, в котором обучают базовую сверточную нейронную сеть (СНС) на предварительно выбранном наборе изображений; модифицируют СНС посредством обеспечения, по меньшей мере, одного скрытого слоя с регуляризацией методом дропаута; обеспечивают поверх, по меньшей мере, одного скрытого слоя независимые полносвязные слои для

распознавания атрибутов лиц, причем каждый из этих независимых слоев соответствует ответственному одному из атрибутов лиц и имеет соответствующую нелинейность; извлекают набор мультимодальных атрибутов человека посредством слоев базовой СНС первого представления первого изображения человека; применяют выходные данные классификатора атрибутов и выходные данные измеренного расстояния, сгенерированные с использованием второго представления второго изображения человека, к ранее обученной модели машинного обучения для проверки слияния; генерируют выходные данные системы распознавания лиц с использованием ранее обученной модели машинного обучения проверки слияния; и контролируют доступ к защищенному объекту с использованием выходных данных системы распознавания лиц. Технический результат заключается в повышении эффективности кластеризации лиц и распознавания атрибутов лиц путем обучения представлений лиц с предварительным обучением сверточной нейронной сети (СНС) для задачи идентификации лиц из предварительно собранной базы данных.

1 н. п. ф., 10 з. п. ф., 6 фиг.

(19) **KG** (11) **432** (13) **C2** (46) **31.07.2025**

3

Данное изобретение относится к области вычислительной техники, но не в порядке ограничения, а в частности относится способам распознавания лиц.

Технологии классификации изображений (например, технологии классификации лиц) имеют множество вариантов использования. Среди примеров использования можно выделить разрешение авторизованным лицам (и запрет неавторизованным лицам) входить в безопасное физическое пространство, аутентификация пользователя электронного устройства или идентификация объекта (например, стула, стола и т. д.). Одной из проблем некоторых реализаций технологии классификации изображений является высокая частота неправильной идентификации обученной нейронной сетью классификации изображений.

Верификация лица является важным аспектом практически любой современной системы распознавания лиц. Верификация лица может вычислять сходство один-к-одному, чтобы определить, являются ли два изображения одним и тем же предметом, где информация о распознанном объекте может быть получена через конвейер классификации лиц.

Верификация может не выполняться на уровне пикселей между зондом и галереей из-за того, что существует слишком много вариаций и помех в необработанных изображениях лица. Вместо этого из изображений лиц извлекаются высокоуровневые черты лица либо обычными методами, такими как HOG, SIFT и другие, либо более продвинутыми и основанными на данных нейросетевыми подходами, такими как Dlib, Arcface и другие. Затем может быть проведена проверка векторов черт лица с использованием метрик сходства, таких как евклидово расстояние или косинусное сходство, известных из уровня техники.

Хотя высокоуровневые черты лица могут упростить верификацию лица, авторы настоящего изобретения признали, что все еще существуют ограничения, связанные с векторами отпечатков лиц, которые иногда могут приводить к сбоям верификации. Ограничения в основном связаны с аспектами, описанными ниже.

4

Искусственная нейронная сеть, используемая для извлечения черт лица, может сильно зависеть от качества сгенерированного отпечатка лица, которое определяет эффективность верификации. Между тем, на производительность нейронной сети может влиять множество факторов, таких как набор обучающих данных, структура сети и функция потерь.

Что касается обучающего набора данных, большинство общедоступных наборов данных изображений лиц имеют множество проблем, связанных с демографической погрешностью. Например, этнические группы среди обучающих идентичностей часто несбалансированы, при этом большинство идентичностей относятся к одной этнической группе. Например, в таких наборах изображений лиц, как VGGFace, MS-celeb-1M и Youtube Face, преобладают лица европейской этнической принадлежности, в то время как в CASIA - восточноазиатской. Дисбаланс этнических групп может привести к тому, что работа нейронных сетей будет благоприятствовать одной этнической группе, что приведет к негативному влиянию на эффективность верификации для других этнических групп.

В дополнение к этнической принадлежности возрастные различия в обучающем наборе данных также могут приводить к смещению в сторону определенных возрастных групп в зависимости от того, как проводился процесс сбора данных. Такая возрастная погрешность может привести к погрешности в производительности обученных нейронных сетей.

В дополнение к этнической и возрастной принадлежности проблемой также является гендерный дисбаланс. Гендерный дисбаланс в большинстве случаев можно контролировать при составлении списка для обучения идентификаторов на основе публичной информации, но он может быть не идеальным, если при сборе данных не проводилась тщательная оценка.

Наконец, еще одной причиной неудачной верификации может быть большой разброс положений лиц. Сравнение двух профильных изображений лица для целей верификации представляет собой более сложную задачу даже с точки зрения человека.

5

Из уровня техники известно решение по распознаванию лиц (RU 2710942 C1 <https://patents.google.com/patent/RU2710942C1/ru>), в котором обучают базовую сверточную нейронную сеть (СНС) на предварительно выбранном наборе изображений; модифицируют СНС посредством обеспечения по меньшей мере одного скрытого слоя с регуляризацией методом дропаута; обеспечивают по верху, по меньшей мере, одного скрытого слоя независимые полносвязные слои для распознавания атрибутов лиц, причем каждый из этих независимых слоев соответствует ответственному одному из атрибутов лиц и имеет соответствующую нелинейность (см. RU 2710942 C1, <https://patents.google.com/patent/RU2710942C1/ru> п. 1 формулы).

Из статьи Changxing Ding et al. "Robust Face Recognition via Multimodal Deep Face Representation" arxiv.org (<https://arxiv.org/abs/1509.00244v1>) известно распознавание лиц на основе мультимодальных алгоритмов. В частности выполнение мультимодального выделения признаков с использованием набора сверточных нейронных сетей (CNN) и объединение (слияние) на уровне признаков набора CNN-признаков с использованием трехслойного стекового автоэнкодера (SAE) (см. вышеуказанную статью <https://arxiv.org/abs/1509.00244v1>, разделы II.B, III.A и III.B).

Также известна статья Мультимодальная биометрическая система распознавания личности cyberleninka.ru (<https://cyberleninka.ru/article/n/multimodalnaya-biometricheskaya-sistema-raspoznavaniya-lichnosti-po-litsu-i-raduzhnoy-obolochke-glaza?ysclid=lx01lil1dt419792618>), раскрывающая выделение различных признаков лица независимо друг от друга и слияние результатов распознавания.

Недостатки данных решений очевидны для любого специалиста в уровне техники за счет отсутствия эффективности.

На изображениях, которые необязательно выполнены в масштабе, посредством подобных численных величин можно описывать схожие компоненты в различных видах. Подобные численные величины с различными буквенными суффиксами могут представлять различные экземпляры аналогичных компонентов.

На фиг. 1 представлена концептуальная схема, иллюстрирующая пример системы

6

распознавания лиц, использующей мультимодальные атрибуты лица.

На фиг. 2 показан пример системы распознавания лиц, использующей один или несколько вспомогательных атрибутов, применяющихся в сочетании с визуальной информацией для повышения эффективности идентификации лица.

На фиг. 3 представлена блок-схема одного из примеров реализованного способа использования одного или нескольких вспомогательных атрибутов, применяемых в сочетании с визуальной информацией для повышения эффективности идентификации лица в системе распознавания лиц, как было описано выше в отношении фиг. 2.

На фиг. 4 представлена блок-схема одного из примеров реализованного способа 100 использования мультимодальных атрибутов лица в системе распознавания лиц для повышения эффективности идентификации лица в системе распознавания лиц, как было описано выше в отношении фиг. 1.

На фиг. 5 показан пример модуля машинного обучения 200 в соответствии с некоторыми примерами реализации настоящего изобретения.

На фиг. 6 показана блок-схема примера устройства, для распознавания которой может быть реализован один или несколько способов (например, методологий), рассмотренных в настоящем документе.

Технический результат заключается в повышении эффективности кластеризации лиц и распознавания атрибутов лиц путем обучения представлений лиц с предварительным обучением сверточной нейронной сети (СНС) для задачи идентификации лиц из предварительно собранной базы данных.

В настоящем изобретении рассмотрено использование мультимодальных атрибутов лица в системах распознавания лиц. Кроме того, использование одного или нескольких вспомогательных атрибутов, таких как темпоральный атрибут, может быть использовано в сочетании с визуальной информацией для повышения эффективности идентификации лица в системе распознавания лиц. В некоторых примерах реализации мультимодальных атрибутов лица в системах распознавания лиц может сочетаться с использованием одного или нескольких вспомогательных атрибутов,

7

таких как темпоральный атрибут. Каждый из этих методов может повысить эффективность верификации системы распознавания лиц.

В некоторых вариантах реализации настоящее описание относится к реализуемому вычислительным средством способу использования системы распознавания лиц для идентификации человека, причем способ включает в себя: извлечение атрибута человека путем применения первого представления первого изображения человека к ранее обученной модели машинного обучения классификатора атрибутов для создания выходных данных классификатора атрибутов; применение выходных данных классификатора атрибутов и выходных данных измерения расстояния, созданных с использованием второго изображения второго изображения человека, к ранее обученной модели машинного обучения для верификации слиянием; генерирование выходных данных системы распознавания лиц с использованием ранее обученной модели машинного обучения для верификации слиянием; и контроль доступа к защищенному активу с использованием выходных данных системы распознавания лиц.

В некоторых аспектах настоящее раскрытие относится к реализуемому компьютером методу использования системы распознавания лиц для идентификации человека по его изображению, метод включает в себя: составление отпечатка лица; применение отпечатка лица к конвейеру классификации для идентификации похожего изображения; генерирование выходных данных конвейера классификации на основе идентифицированного изображения; применение выходных данных конвейера классификации к модели совместной классификации; применение вспомогательного атрибута к модели совместной классификации; генерирование выходных данных совместной классификации с использованием модели совместной классификации; и управление доступом к защищенному активу с использованием выходных данных совместной классификации.

Задача мультимодальной аутентификации состоит из нескольких этапов:

- a. Получить результаты из отдельных систем.
- b. Произвести нормализацию этих результатов.

8

с. Осуществить процедуру слияния нормированных составляющих.

Таким образом, получается результат, который имеет такой же вид, как если бы производилась унимодальная аутентификация. При этом он содержит в себе информацию от всех составных частей мультимодальной модели.

Первый этап не является сложным в реализации, так как аутентификацию по отдельным системам можно производить одновременно. Например, можно просить пользователя одновременно зафиксировать палец на сканере отпечатков и произвести сканирование сетчатки глаза. Дальнейшие шаги же выполняются внутри общей системы.

Данный подход позволяет значительно увеличить уровень безопасности систем аутентификации. В нем используется уже не одна характеристика человека, а несколько, что усложняет задачу подделки биометрических данных. Также существенным является то, что такие системы не только с большей точностью выявляют злоумышленников, но и имеют меньшее количество отказов для зарегистрированных пользователей (то есть более высокий True Positive Rate). Данное качество позволяет в значительной степени повысить качество решений в сфере аутентификации.

Распознавание лиц происходит в три этапа: обнаружение, анализ и распознавание.

Обнаружение - это процесс поиска лица на изображении. С помощью компьютерного зрения распознавание лиц позволяет обнаружить и идентифицировать отдельные лица на изображении, содержащем лица одного или многих людей. Оно может распознавать данные лица как в профиль, так и в анфас.

Машинное зрение позволяет компьютерам идентифицировать на изображениях людей, места и предметы с точностью, которая сравнима с человеческими способностями или даже превышает их, и при этом с гораздо более высокой скоростью и эффективностью. Используя сложную технологию искусственного интеллекта, компьютерное зрение автоматизирует извлечение, анализ, классификацию и понимание полезной информации из данных изображения. Данные изображения принимают различные формы, например, следующие:

- одиночные изображения;
- последовательности видео;
- вид с нескольких камер;
- трехмерные данные.

Затем система распознавания лиц анализирует изображение лица. Она отображает и считывает геометрию лица и мимику, а затем определяет ориентиры на лице, которые являются ключевыми для отличия лица от других объектов (атрибуты). Технология распознавания лиц обычно определяет следующее:

- расстояние между глазами;
- расстояние от лба до подбородка;
- расстояние между носом и ртом;
- глубина глазниц;
- форма скул;
- контур губ, ушей и подбородка.

Затем система преобразует данные распознавания лица в строку цифр или точек, называемую отпечатком лица (данный набор данных может формироваться в векторе). Каждый человек имеет уникальный отпечаток лица, подобный отпечатку пальца. Информация, используемая при распознавании лиц, также может быть использована в обратном направлении для цифровой реконструкции лица человека.

Распознавание лиц позволяет идентифицировать человека путем сравнения лиц на двух или более изображениях и оценки вероятности их совпадения. Например, так можно проверить, совпадает ли лицо, изображенное на селфи, снятом камерой мобильного телефона, с лицом на изображении выданного правительством удостоверения личности, такого как водительские права или паспорт, а также узнать, совпадает ли лицо, изображенное на селфи, с лицом в коллекции лиц, снятых ранее.

В настоящем изобретении описаны различные способы преодоления проблем, известных из уровня техники. Авторы настоящего изобретения осознали необходимость создания нового представления признаков на основе мультимодальных атрибутов лица. Новое представление признаков может являться сочетанием одной или нескольких черт лица, таких как возраст, пол, эмоции и этническая принадлежность, как будет более подробно описано ниже.

В частности, при наличии на пользовательском изображении человеческого лица

может быть определена и маркирована позиция лица.

Операции по извлечению атрибутов могут быть реализованы посредством признаков Хаара в совокупности с алгоритмом детектирования лица *adaboost*. Признаки Хаара могут быть разделены на три категории, в число которых входят граничный признак, линейный признак, а также центральный признак и диагональный признак. Алгоритм детектирования лица *adaboost* может включать в себя выбор признаков Хаара и расчет признаков. В том числе расчет признаков может быть реализован посредством способа интегрального изображения. Согласно способу интегрального изображения, три категории признаков Хаара могут быть скомбинированы для формирования шаблона признаков, содержащего белый прямоугольник и черный прямоугольник, и величина признака шаблона может быть получена вычитанием суммарного значения пикселей черного прямоугольника из суммарного значения пикселей белого прямоугольника.

Согласно основному принципу интегрального изображения, суммарное значение пикселей каждой прямоугольной области на изображении, которая может быть сформирована от начальной точки до каждой точки на изображении, может быть сохранено в памяти в качестве элемента массива. При необходимости расчета суммарного значения пикселей некоторой области может быть непосредственно использован элемент массива без пересчета суммарного значения пикселей этой области, что позволяет ускорить расчеты. При наличии различных размеров одной области интегральное изображение позволяет рассчитать различные признаки области, затрачивая одинаковое время, что значительно повышает скорость детектирования.

Алгоритм *adaboost* может считаться обычным методом детектирования лица и далее может повторно не упоминаться.

Помимо описания применения мультимодальных атрибутов лица в системах распознавания лиц, в данном изобретении описывается использование одного или нескольких вспомогательных атрибутов, таких как темпоральный атрибут, может быть использовано в сочетании с визуальной информацией для повышения эффективности идентификации лица в системе распознавания лиц. Каждый из

11

этих способов может повысить эффективность верификации системы распознавания лиц.

Современная система распознавания лиц в неограниченной среде решает две задачи:

д. 1) классификация лиц (1-ко-многим), которая заключается в попытке сопоставить запрошенное изображение лица с наиболее похожим лицом в существующей базе данных хранилища, собранной в результате регистрации пользователей;

е. 2) верификация лиц (1-к-1), которая заключается в попытке определить, являются ли запрошенное изображение и наиболее похожее изображение из хранилища одним и тем же лицом или нет.

Оба процесса в значительной степени зависят от отпечатков лиц, полученных с помощью конвейера извлечения признаков лица, который может представлять собой как разработанный вручную традиционный экстрактор признаков на изображении, так и искусственные нейронные сети.

Как следует из результатов данного изобретения, эффективность верификации лица в некоторой степени зависит от эффективности классификации лиц. Для повышения эффективности сопоставления можно использовать методы машинного обучения, такие как линейный классификатор, например SVM или логистическая регрессия, или деревья решений, например XGboost или LightGBM. Идея состоит в том, чтобы обучить модель классификации на отпечатках лиц, извлеченных из базы данных зарегистрированных лиц, чтобы получить лучшее представление о целевой группе. В некоторых примерах реализации для каждого человека регистрируется небольшое количество изображений лица с высоким качеством и контролируемой вариативностью.

На фиг. 1 представлена примерная схема реализации, иллюстрирующая работу системы распознавания лиц 10, использующую мультимодальные атрибуты лица. Чтобы идентифицировать человека 12 или лицо сеть распознавания лиц 14 может генерировать отпечаток лица 16 из первого представления первого изображения человека, которым может быть первый шаблон, первый вектор или первое изображение, вне зависимости от вида представления. Отпечаток лица 16 может

12

включать числовые векторы, которые могут представлять каждое обнаруженное лицо в представлении изображения. Чтобы получить этот вектор, сверточная нейронная сеть предварительно обучается с учителем идентифицировать лица из большого набора данных, например, таких как CASIA-WebFace, VGGFace/VGGFace2 или MS-Celeb-1M.

В некоторых примерах реализации для генерации отпечатка лица 16 может быть использована модель машинного обучения или нейронная сеть. В качестве примера нейронная сеть ArcFace может использоваться для генерации отпечатков лиц. Например, предварительно обученные веса нейронной сети могут быть преобразованы из git-репозитория deepinsight/insightface с открытым исходным кодом, который может быть обучен с архитектурой сетью LResNet50E-IR и функцией потерь ArcFace.

На предварительном этапе базовую искусственную сверточную нейронную сеть (далее - СНС) обучают на имеющемся очень большом наборе изображений для получения отпечатка лица. Базовая СНС предпочтительно представляет собой СНС с низкой вычислительной сложностью и затратам памяти, например, MobileNet v1/v2.

На следующем этапе модифицируют СНС, обеспечивая, по меньшей мере, один скрытый слой регуляризации методом дропаута поверх слоев базовой СНС. Исключение или дропаут (от англ. dropout) - метод регуляризации искусственных нейронных сетей, который предназначен для уменьшения переобучения сети за счет предотвращения сложных коадаптаций отдельных нейронов на тренировочных данных во время обучения. После этого, поверх упомянутого по меньшей мере одного скрытого слоя обеспечиваются независимые полносвязные слои для распознавания атрибутов лиц. Каждый из этих независимых слоев соответствует конкретному атрибуту лица и имеет конкретную нелинейность. Далее осуществляют обучение этих независимых полносвязных слоев. При обучении поочередно используют подмножество (батч) обучающих данных, относящийся только к одному из независимых слоев, обучаемому в данное время. Здесь следует отметить, что каждый батч обучающих данных может иметь соответствующую метку, указывающую конкретный атрибут лица, к которо-

13

му данный батч относится. Например, отдельный батч это одна раса, отдельный батч это одна эмоция и так далее. Затем на вход модифицированной и обученной СНС подаются одно или несколько входных изображений. Далее слои базовой СНС извлекают из, по меньшей мере, частей введенных изображений индивидуальные характерные признаки, подходящие для идентификации лиц.

Перед извлечением признаков часто во входных изображениях детектируют области, ассоциированные с лицами, например, с помощью метода MTCNN (сеть для классификации), и эти области затем используются в качестве упомянутых частей входных изображений. Затем скрытый слой СНС предоставляет ввод для независимых полносвязных слоев, используя извлеченные индивидуальные характерные признаки. И наконец, атрибуты лиц соответственно распознаются независимыми полносвязными слоями на основе ввода от скрытого слоя.

Предложенная СНС имеет следующие преимущества. Прежде всего, она очень эффективна, так как позволяет использовать СНС с высокой вычислительной скоростью и низкими затратами памяти (например, базовую сеть MobileNet) в качестве базовой СНС и одновременно решать все вышеупомянутые задачи распознавания возраста, пола, этнической принадлежности, эмоций и идентификации личности, без необходимости применения нескольких разных СНС. Во-вторых, в отличие от общедоступных наборов данных, обычно используемых для задач распознавания атрибутов лиц, которые имеют довольно малый объем и низкое качество, в предложенной модели используется потенциал очень больших качественных наборов данных для идентификации лиц для изучения высококачественных представлений лиц. Кроме того, скрытый слой между индивидуальными характерными признаками и выходами дополнительно объединяет знания, необходимые для классификации атрибутов лиц. В результате, эта модель позволяет повысить точность распознавания атрибутов лиц по сравнению с моделями, основывающимися на обучении только на специфических наборах данных.

Используя различные способы настоящего изобретения, реализуемый вычислительным устройством способ может постро-

14

ить новое представление признаков на основе извлеченных мультимодальных атрибутов лица. Новое представление признаков может быть использовано для обучения бинарного классификатора для верификации лица. Новое представление признаков может представлять собой комбинацию нескольких черт лица, таких как возраст, пол, эмоции и этническую принадлежность, как было описано выше.

Например, реализуемый вычислительным устройством способ может извлечь один или несколько атрибутов человека 12 путем применения первого представления первого изображения человека к ранее обученной модели машинного обучения классификатора атрибутов. Наиболее распространены два способа представления изображений: растровый и векторный. Существует еще фрактальный способ представления графических изображений. К атрибутам могут относиться, например, возраст, пол, этническая принадлежность и/или положение головы человека. Например, отпечаток лица может сгенерировать вектор, который применяется к одной или нескольким ранее обученным моделям машинного обучения классификатора атрибутов, а затем каждая модель может извлечь свой конкретный соответствующий атрибут. В другом примере, одна или несколько ранее обученных моделей машинного обучения классификатора атрибутов могут получать необработанные изображения в качестве входных данных, генерировать признак, а затем извлекать соответствующий атрибут.

В качестве неограничивающего конкретного примера, способ может использовать вычисленные оценки возраста в 5 категориях, таких как 15-20, 25-32, 38-43, 48-53 и 60+; вычисленные оценки пола в 2 категориях: мужчина, женщина; вычисленные оценки этнической принадлежности в 5 категориях: азиатско-индийская, азиатско-восточная, афроамериканская, европеоидная и латиноамериканская; и оценка углов наклона, движения и крена головы для положения головы человека.

Обученная ранее модель (модели) машинного обучения классификатора признаков, как описано выше, может являться частью механизма верификации 18. Модель машинного обучения классификатора возраста 20 может извлекать возраст человека 12 из

15

представления изображения человека. Аналогично, модель машинного обучения 22 классификатора пола может извлекать пол человека 12. Модель машинного обучения 24 классификатора этнической принадлежности может извлекать этническую принадлежность человека 12. Модель машинного обучения 26 классификатора положения головы может извлекать положение головы, например углы наклона, движения и крена, человека 12.

Кроме того, реализуемый вычислительным устройством способ может получить измерение расстояния между различными признаками, используя второе представление второго изображения человека, которым может быть второй шаблон, второй вектор или второе изображение. В некоторых примерах первый шаблон может быть таким же, как и второй шаблон. В некоторых примерах первый шаблон может отличаться от второго шаблона. В некоторых примерах первое изображение может быть таким же, как и второе изображение. В некоторых примерах первое изображение может отличаться от второго, например, быть снятым под другим углом.

Классификатор измерения расстояния 28 модуля верификации 18 может сравнивать два вектора из двух отпечатков, которыми могут быть шаблоны. Например, классификатор измерения расстояния 28 может сравнивать вектор из отпечатка лица, представленного считывающему устройству, например, с вектором из отпечатка лица человека 12, которое было получено на месте происшествия, например, с помощью камеры. Эти два вектора представлены нижним входом 30 в классификатор измерения расстояния 28.

В некоторых вариантах реализации классификатор измерения расстояния 28 может в качестве альтернативы сравнивать вектор из отпечатка лица, представленного считывающему устройству или камере, например, с вектором из отпечатка лица человека, найденного при поиске изображений в базе данных. Например, конвейер классификации 32, которым может быть предварительно обученная модель машинного обучения или алгоритм, может выполнять поиск наиболее похожих шаблонов в базе данных шаблонов. Выходные данные конвейера классификации 32 могут включать наиболее похожие шаблоны. Затем данный способ может выполнить проверку изображений, используя, например,

16

поиск по центроиду 34. Кластер, один из основных терминов кластерного анализа, - группа из элементов со схожими свойствами. Слово «кластер» образовано от английского «cluster», что означает пучок, группа, скопление, гроздь, куст. Не менее важное понятие центроид - центр кластера.

Классификатор измерения расстояния 28 механизма верификации 18 может сравнивать вектор отпечатка лица, представленного, например, считывающему устройству или камере, с вектором отпечатка лица наиболее похожего изображения, полученного при использовании конвейера классификации 32 и поиска 34. Эти два вектора представлены верхним входом 36 в классификатор измерения расстояния 28. В некоторых примерах классификатор измерения расстояния 28 может использовать оба входа 30, 36.

Выходные данные 38 классификатора измерения расстояния 28 могут представлять собой математическое значение, полученное в результате сравнения двух векторов. В некоторых примерах классификатор измерения расстояния 28 может быть ранее обученной моделью машинного обучения. В других примерах классификатор измерения расстояния 28 может представлять собой алгоритм.

Каждая из ранее обученных моделей машинного обучения классификатора атрибутов 20-26 может обеспечивать генерирование соответствующих выходных данных классификатора атрибутов 40-46. Используя способы настоящего изобретения, один или несколько выходов классификатора атрибутов 40-46 могут быть применены совместно с результатом измерения расстояния 38 к ранее обученной модели машинного обучения для верификации слиянием 48. В некоторых примерах, ранее обученная модель машинного обучения для верификации слиянием 48 может являться бинарным классификатором XGBoost, построенным на основе деревьев решений.

Путем слияния или объединения выходных данных измерения расстояния 38 с выходными данными одного или более классификаторов атрибутов 40-46, предварительно обученная модель машинного обучения для верификации слиянием 48 может повысить эффективность верификации системы распознавания лиц принимая таким образом решение.

17

Примерный вариант реализации может обеспечивать генерирование выходных данных системы распознавания лиц 50 с использованием предварительно обученной модели машинного обучения верификации слиянием 48, которая может, к примеру, контролировать доступ к защищенному объекту, такому как помещение, здание или компьютерная система. Модуль верификации 18 может представлять собой программу для ЭВМ. Модуль верификации 18, ранее обученные модели машинного обучения и любые алгоритмы, описанные в настоящем изобретении, могут быть реализованы с помощью системы 300 в соответствии с фиг. 6. Например, один или несколько модулей верификации 18, моделей машинного обучения 20-26 и модели машинного обучения 48 могут быть реализованы с помощью инструкций 324, выполняемых процессором 302. Модуль верификации и модели машинного обучения могут находиться на одной или разных машинах.

Авторы изобретения создали исчерпывающий список комбинаций для всех новых категорий признаков, причем цифра "1" означает, что признак (возраст, пол, этническая принадлежность и угол рысканья) используется, а "0" означает, что признак не используется. Для каждой комбинации обучался классификатор XGBoost на тестовом наборе данных VGGFace2, из которого половина пользователей была отнесена к подлинным, а другая половина - к ложным.

Как упоминалось выше, в дополнение к использованию мультимодальных признаков лица в системах распознавания лиц, в настоящем изобретении описано использование одного или нескольких вспомогательных признаков, таких как темпоральный атрибут, которые могут быть использованы в сочетании с визуальной информацией для повышения эффективности идентификации лица в системе распознавания лиц. Использование вспомогательных признаков показано и описано ниже применительно к фиг. 2.

Несколько факторов могут ограничивать производительность визуального классификатора. Например, нейронная сеть, генерирующая отпечатки лиц, несовершенна. В настоящее время проводятся исследования, направленные на разработку современного способа извлечения черт лица, поскольку он играет важную роль во всем процессе распо-

18

знавания лиц. Несмотря на то, что производительность нейронных сетей достигла очень высокого уровня, до сих пор не найдено решения, которое позволило бы добиться идеального распознавания лиц на крупномасштабных и специальных наборах данных. Кроме того, в условиях неограниченной обработки изображений вариативность запрашиваемых изображений также влияет на эффективность распознавания. Некоторые из наиболее распространенных нюансов были описаны выше, включая угол наклона лица, возрастные различия, этническую принадлежность или качество изображения.

В свете вышеизложенного авторы настоящего изобретения осознали необходимость создания новой архитектуры, объединяющей визуальные и вспомогательные атрибуты, такие как темпоральная информация, для повышения эффективности идентификации лица, показанной и описанной ниже на фиг. 2.

На фиг. 2 представлена концептуальная схема, иллюстрирующая пример системы распознавания лиц 60 посредством одного или нескольких вспомогательных атрибутов, которые могут быть использованы в сочетании с визуальной информацией для повышения эффективности идентификации лица. На фиг. 2 концептуально изображен реализуемый компьютером метод использования системы распознавания лиц для идентификации человека по его изображению. Как описано ниже, когда визуальный классификатор предсказывает несколько кандидатов с близким уровнем достоверности, вспомогательный атрибут (атрибуты), такой как временная информация, может быть использован для повторного ранжирования прогнозов для улучшения идентификации. Способы, описанные на фиг. 2, могут повысить точность классификации лиц (1-ко-многим).

Для идентификации человека 62 (ID_a) сеть распознавания лиц 64 может генерировать или выполнять составление отпечатков лиц из первого представления первого изображения человека, которым может быть первый шаблон, первый вектор или первое изображение. Отпечатки лиц могут содержать числовые векторы, которые могут представлять каждое обнаруженное лицо в представлении изображения.

Отпечатки лиц, например, векторы, сгенерированные сетью распознавания лиц, могут быть затем применены к конвейеру классификации 66. Например, конвейер классификации 66, такой как предварительно обученная модель машинного обучения или алгоритм, может выполнять поиск наиболее похожих шаблонов в базе данных шаблонов. Выходные данные конвейера классификации 66 могут содержать наиболее похожие шаблоны. В примере, показанном на фиг. 2, наиболее похожим шаблоном, найденным конвейером классификации, является шаблон ID_b, показанный на позиции 68. Конвейер классификации 66 на фиг. 2 определил, что он имеет более высокую достоверность прогноза для шаблона ID_b, чем для шаблона ID_a для человека 62 (ID_a).

Конвейер классификации 66 может генерировать выходные данные конвейера классификации 70 на основе идентифицированного изображения. В некоторых примерах выходные данные конвейера классификации 70 могут включать измерение расстояния. В других примерах выходные данные конвейера классификации 70 могут включать оценку или вероятность совпадения для каждого зарегистрированного пользователя в системе.

Например, конвейер классификации 66 может включать классификатор измерения расстояния, такой как описан выше в отношении фиг. 1, который может сравнивать два вектора из двух отпечатков, например, из шаблонов. В некоторых примерах классификатор измерения расстояния может сравнивать вектор из отпечатка лица, представленного считывающему устройству или камере, например, с вектором из отпечатка лица наиболее похожего изображения, найденного при использовании конвейера классификации и поиска центраида.

Выходные данные классификатора измерения расстояния, которые также могут быть выходными данными конвейера классификации 66, могут представлять собой математическое значение, полученное в результате сравнения двух векторов. В некоторых примерах классификатор измерения расстояния может быть ранее обученной моделью машинного обучения. В других примерах классификатор измерения расстояния может представлять собой алгоритм.

При использовании различных способов настоящего изобретения вспомогательный атрибут 72 может применяться в сочетании с визуальной информацией, полученной через конвейер классификации 66, для повышения эффективности идентификации лица системой распознавания лиц 60. Вспомогательный атрибут 72 может среди прочего включать темпоральную информацию, рост человека 62 и/или социальную принадлежность человека 62.

Темпоральная информация может включать, например, временной диапазон момента, когда человек 62 обычно регистрируется системой распознавания лиц у охраняемого входа, например, у двери или на входе в здание. Модель машинного обучения может со временем определить этот временной диапазон для человека 62. В некоторых примерах модель машинного обучения может узнать конкретный охраняемый вход, например, дверь или вход в здание, которым пользуется человек 62, если имеется несколько охраняемых входов. Затем, используя различные методы настоящего изобретения, темпоральная информация или другой вспомогательный атрибут может быть применен вместе с выходными данными конвейера классификации 70 к модели совместной классификации 74.

В качестве неограничивающего конкретного примера для целей иллюстрации, модель машинного обучения системы распознавания лиц 60 ранее определила, что человек 62 регулярно прибывает к охраняемому входу в здание между 6:00 и 6:15 утра. Эта темпоральная информация или другие вспомогательные атрибуты могут быть применены к модели совместной классификации 74 с выходными данными конвейера классификации 70 для повышения эффективности идентификации лица системой распознавания лиц 60. Конвейер классификации 66 первоначально определил более высокую достоверность прогноза в шаблоне ID_b, чем в шаблоне ID_a для человека 62 (ID_a), как показано на 68 на фиг. 2. Однако человек, связанный с шаблоном ID_b, регулярно прибывает к охраняемому входу между 9:00 и 9:15 утра, а текущее время - 6:10 утра. Таким образом, человек 62, скорее всего, связан с шаблоном ID_a, а не с шаблоном ID_b.

21

Используя темпоральную информацию, например, система распознавания лиц 60 и, в частности, модель совместной классификации 74 могут осуществлять генерирование выходных данных 76 с использованием модели совместной классификации 74, которая устанавливает более высокую достоверность прогноза для шаблона ID_a , чем для шаблона ID_b для человека 62 (ID_a), как показано на 78 на рис. 2. Затем, способ может контролировать доступ к защищенному активу, используя выходные данные совместной классификации 76.

Темпоральная информация является одним из возможных вспомогательных атрибутов. В других примерах может быть использован рост человека. В других примерах может использоваться одежда человека. В других примерах может быть использована походка человека. Модели машинного обучения могут быть обучены для классификации роста, одежды и походки людей.

В некоторых примерах может использоваться социальная среда человека 62. Социальная среда может включать информацию о сети людей, с которыми общается 62. Например, социальная среда человека 62 может включать людей, с которыми человек 62 уходит с работы, приходит на работу, ходит на обед и т. п. Модель машинного обучения может изучать личности этих людей с течением времени, чтобы определить социальную среду человека 62. Когда человек 62 прибывает на охраняемый вход, например, ранее обученная модель машинного обучения может определить достоверность прогноза для одного или более лиц с человеком 62, сравнить эту достоверность прогноза с социальной средой человека 62 и применить выходные данные конвейера социальной среды и конвейера классификации 70 к модели совместной классификации 74 для повышения производительности идентификации лица системы распознавания лиц 60.

В некоторых примерах система распознавания лиц 60 может генерировать выходные данные 70 совместной классификации путем генерирования суммарной вероятности с использованием модели совместной классификации 74. Например, суммарная вероятность, основанная на двух входах в модель совместной классификации 74, может быть определена с помощью теории Байеса:

22

Приведенные выше уравнения показывают, что вероятность совпадения лица в запросе и каждого лица в хранилище регистрации с совместным использованием визуальной и темпоральной информации ($Pr(query=galleryn|v,t)$) (уравнение 1), пропорциональна умножению вероятности совпадения от визуального классификатора, обученного на визуальных признаках ($Pr(query=galleryn|v)$) и вероятности прогнозируемого лица, прибывающего на проходную или вход с одной временной меткой ($Pr(tlquery=galleryn)$) (уравнение 2).

Таким образом, сочетание визуальной информации и вспомогательного атрибута может быть использовано для повышения эффективности идентификации лица системой распознавания лиц.

На фиг. 3 представлена блок-схема примера способа 80 использования одного или более вспомогательных атрибутов, которые могут быть применены в сочетании с визуальной информацией для повышения эффективности идентификации лица системой распознавания лиц, как было описано выше в отношении фиг. 2.

В блоке 82 реализуемый компьютером метод 80 может составить отпечаток лица. Например, сеть распознавания лиц, такая как сеть распознавания лиц 60 на фиг. 2, может осуществлять генерирование или составление отпечатка лица из первого представления первого изображения человека, такого как первый шаблон, первый вектор или первое изображение.

В блоке 84 реализуемый компьютером метод 80 может применить отпечаток лица к конвейеру классификации для идентификации похожего изображения. Например, отпечаток лица, содержащий векторы, сгенерированные сетью распознавания лиц, могут быть применены к конвейеру классификации, такому как конвейер классификации 66 на фиг. 2.

В блоке 86 реализуемый компьютером метод 80 может генерировать выходные данные классификационного конвейера на основе идентифицированного изображения. Например, классификационный конвейер 66 на фиг. 2, такой как предварительно обученная модель машинного обучения или алгоритм, может выполнить поиск наиболее похожих шаблонов в базе данных шаблонов и

23

выдать наиболее похожий шаблон. В некоторых примерах выходные данные конвейера классификации 70 на фиг. 2 могут представлять собой измерение расстояния.

В блоке 88 реализуемый компьютером метод 80 может применить выходные данные конвейера классификации к совместной модели классификации. Например, выходные данные конвейера классификации, такие как выходные данные конвейера классификации 70 на фиг. 2, могут быть применены к совместной модели классификации, такой как совместная модель классификации 74 на фиг. 2, которая может вычислять совместную вероятность на основе теории Байеса.

В блоке 90 способ 80 может применить вспомогательный атрибут к совместной модели классификации. Например, темпоральная информация или другой вспомогательный атрибут может быть применен к модели совместной классификации, такой как модель совместной 74 на фиг. 2.

В блоке 92 способ 80 может генерировать выходные данные совместной классификации с использованием модели совместной классификации.

В блоке 94 способ 80 может контролировать доступ к защищенному активу, используя выходные данные совместной классификации. Например, выходные данные совместной классификации 76 на фиг. 2 могут быть использованы для управления доступом к помещению, зданию или компьютерной системе, например, путем подачи управляющего сигнала на систему управления, связанную с защищенным активом.

На фиг. 4 представлена блок-схема примера способа 100 использования мультимодальных атрибутов лица в системе распознавания лиц для повышения эффективности идентификации лица в системе распознавания лиц, как было описано выше в отношении фиг. 1.

В блоке 102 способ 100 может извлекать атрибут лица, такой как возраст, пол, этническая принадлежность и положение головы, путем применения первого представления первого изображения лица, такого как первый шаблон, первый вектор или первое изображение, к ранее обученной модели машинного обучения классификатора атрибутов для генерирования выходных данных классификатора атрибутов. Например, один или

24

несколько атрибутов, таких как возраст, пол, этническая принадлежность и/или положение головы человека, могут быть извлечены с помощью соответствующих ранее обученных моделей машинного обучения классификатора атрибутов 20-26 на фиг. 1.

В некоторых примерах отпечаток лица может генерировать вектор, который применяется к одной или нескольким ранее обученным моделям машинного обучения классификатора признаков, а затем каждая модель может извлекать свой конкретный соответствующий атрибут. В других примерах одна или несколько ранее обученных моделей машинного обучения классификатора атрибутов могут получать необработанные изображения в качестве входных данных, генерировать признак, а затем извлекать соответствующий атрибут.

В блоке 104 способ 100 может применить выходные данные классификатора атрибутов и выходные данные измерения расстояния, созданные с использованием второго представления второго изображения человека, такого как второй шаблон, второй вектор или второе изображение, к ранее обученной модели машинного обучения для верификации слиянием. Например, один или несколько выходных данных классификатора атрибутов 40-46 на фиг. 1 и выходные данные 38 классификатора измерения расстояния 28 на фиг. 1 могут быть применены к ранее обученной модели машинного обучения для верификации слиянием 48 на фиг. 1.

В блоке 106 способ 100 может генерировать выходные данные системы распознавания лиц с использованием ранее обученной модели машинного обучения для верификации слиянием. Например, ранее обученная модель машинного обучения для верификации слиянием 48 на фиг. 1 может генерировать выходные данные системы распознавания лиц 50.

В некоторых примерах использование мультимодальных атрибутов лица в системах распознавания лиц может сочетаться с использованием одного или нескольких вспомогательных атрибутов, таких как темпоральный атрибут. Каждый из этих методов может повысить эффективность верификации системы распознавания лиц. То есть в некоторых примерах способы на фиг. 2 могут быть дополнительно объединены со способами на

25

фиг. 1. Такие дополнительные способы показаны в блоках 108 и 110 на фиг. 4.

В блоке 108 способ 100 может дополнительно применять выходные данные системы распознавания лиц и вспомогательный атрибут к совместной модели классификации. Например, выходные данные системы распознавания лиц 50 на фиг. 1 могут быть применены со вспомогательным атрибутом, таким как темпоральная информация, к совместной классификационной модели 74 на фиг. 2.

В блоке 110 способ 100 может дополнительно генерировать выходные данные совместной классификации с использованием модели совместной классификации, такой как модель совместной классификации 74 на фиг. 2.

В блоке 112 способ 100 может контролировать доступ к защищенному активу, используя выходные данные совместной классификации или выходные данные системы распознавания лиц 50, если вспомогательный атрибут не используется. Например, выходные данные 74 совместной классификации могут быть использованы для управления доступом в помещение, здание или компьютерную систему, например, путем подачи управляющего сигнала на систему управления, связанную с защищенным активом.

На фиг. 5 показан пример модуля машинного обучения 200 в соответствии с некоторыми примерами настоящего изобретения. Модуль машинного обучения 200 может быть полностью или частично реализован одним или несколькими вычислительными устройствами. В некоторых примерах модуль обучения 202 может быть реализован другим устройством, чем модуль прогнозирования 204. В этих примерах модель 214 может быть создана на первой машине, а затем отправлена на вторую машину.

Модуль машинного обучения 200 использует модуль обучения 202 и модуль прогнозирования 204. Модуль обучения 202 может реализовать способ обучения схемы обработки, такой как процессор 302 на фиг. 6, с использованием машинного обучения в системе распознавания лиц для идентификации человека. Модуль обучения 202 вводит обучающие данные 206 в селекторный модуль 208.

Обучающие данные 206 могут включать, например, изображения или векторы

26

признаков, сгенерированные сетью СНС извлечения признаков. Используя обучающие данные 206, алгоритм машинного обучения 212 может получить изображения или векторы признаков с желаемой меткой (например, возраст, пол и др.) для каждого образца и пропустить их через алгоритм машинного обучения, такой как SVM, XBoost.

В некоторых примерах обучающие данные 206 могут быть помечены. В других примерах обучающие данные могут быть не помечены, и модель может быть обучена с использованием данных обратной связи, например, с помощью метода обучения с подкреплением.

Селекторный модуль 208 выбирает вектор обучения 210 из обучающих данных 206. Выбранные данные могут заполнить обучающий вектор 210 и включают набор обучающих данных, которые определены как прогнозирующие классификацию материала. Информация, выбранная для включения в обучающий вектор 210, может представлять собой все обучающие данные 206 или, в некоторых примерах, может быть подмножеством всех обучающих данных 206. Обучающий вектор 210 может быть использован (вместе с любыми применимыми метками) алгоритмом машинного обучения 212 для создания модели 214 (обученной модели машинного обучения). В некоторых примерах могут использоваться другие структуры данных, отличные от векторов. Алгоритм машинного обучения 212 может обучать один или несколько слоев модели.

Примерные слои могут включать сверточные слои, выпадающие слои, слои с объединенной / восходящей выборкой, слои SoftMax и им подобные. Примерные модели могут представлять собой нейронную сеть, где каждый слой состоит из множества нейронов, которые принимают множество входных сигналов, взвешивают входные сигналы, вводят взвешенные входные сигналы в функцию активации для получения выходных данных, которые затем могут быть отправлены на другой слой. Примерные функции активации могут включать блок линейной ректификации (ReLU) и ему подобные. Слои модели могут быть полностью или частично соединены.

В модуле прогнозирования 204 данные 216 могут быть введены в селекторный мо-

27

дуль 218. Данные 216 могут включать набор данных акустической визуализации, например, матрицу рассеивания. Селекторный модуль 218 может работать так же, как и селекторный модуль 208 модуля обучения 202, или иным образом. В некоторых примерах селекторные модули 208 и 218 являются одними и теми же модулями или различными экземплярами одного и того же модуля. Селекторный модуль 218 генерирует вектор 220, который вводится в модель 214 для создания изображения материала, отображающего вероятность дефекта на пиксель или вексель, в результате чего формируется изображение 222.

Например, весовые коэффициенты и/или структура сети, изученные обучающим модулем 202, могут быть реализованы на векторе 220 путем применения вектора 220 к первому слою модели 214 для получения входных данных для второго слоя модели 214, и так далее до получения выходных данных классификации материала. Как отмечалось ранее, возможно использование других структур данных, помимо вектора (например, матрицы).

Модуль обучения 202 может работать в режиме офлайн для обучения модели 214. Модуль прогнозирования 204, однако, может быть разработан для работы в режиме онлайн. Следует отметить, что модель 214 может периодически обновляться посредством дополнительного обучения и/или обратной связи с пользователем. Например, дополнительные обучающие данные 206 могут собираться по мере предоставления пользователями обратной связи о возрасте, поле, этнической принадлежности, положении головы и других параметрах, и/или вспомогательных атрибутов, таких как темпоральная информация. Обратная связь, а также данные 216, соответствующие этой обратной связи, могут быть использованы для уточнения модели модулем обучения 202.

Алгоритм машинного обучения 212 можно выбрать из множества различных потенциальных контролируемых или неконтролируемых алгоритмов машинного обучения. Среди примеров алгоритмов обучения можно выделить искусственные нейронные сети, сверточные нейронные сети, байесовские сети, обучение на основе экземпляров, методы опорных векторов, деревья решений (например, Iterative Dichotomiser 3, C4.5,

28

Classification and Regression Tree (CART), Chi-squared Automatic Interaction Detector (CHAID) и т. п.), случайные леса, линейные классификаторы, квадратичные классификаторы, метод k-ближайших соседей, линейная регрессия, логистическая регрессия, региональные CNN, полные CNN (для семантической сегментации), алгоритм маскирования R-CNN для сегментации экземпляров и скрытые марковские модели. Среди примеров алгоритмов обучения без учителя можно отметить алгоритмы максимизации ожидания, векторное квантование и метод информационного узкого места.

Таким образом, модуль машинного обучения 200 на фиг. 5 может помочь в реализации компьютерного метода использования системы распознавания лиц для идентификации человека в соответствии с настоящим изобретением.

Способы, показанные и описанные в настоящем документе, могут быть реализованы с использованием части или всей машины 300, как рассмотрено ниже в отношении фиг. 6.

Ниже будут представлены результаты экспериментов по распознаванию атрибутов лиц на основе потока кадров при идентификации лица.

В данном подразделе предложенная модель сравнивается со следующими известными СНС для предсказания возраста/пола:

1. Age_net/gender_net, обученная на наборе данных Adience.

2. Сеть Deep expectation (DEX) VGG16, обученная на довольно большом наборе данных IMDB-Wiki.

Кроме того, анализируются два специальных случая модели на основе MobileNet. Во-первых, модель сжимается с использованием стандартных преобразований графа при квантовании в Tensorflow. Во-вторых, все слои в этой модели дообучают на предсказание возраста и пола. Хотя такое дообучение, по-видимому, снижает точность идентификации лица по индивидуальным характерным признакам на выходе базовой сети MobileNet, она вызывает повышение валидационной точности на 1 % и 2 % для классификации по полу и возрасту, соответственно.

Эксперименты выполнялись на ноутбуке MacBook 2016 Pro (ЦП: 4xCore i7 2,2 ГГц, ОЗУ: 16 ГБ) и двух мобильных телефонах, в

частности: 1) Honor 6C Pro (ЦП: MT6750 4×1 ГГц и 4×2,5 ГГц, ОЗУ: 3 ГБ); и 3) Samsung S9+ (процессор: 4×2,7 ГГц, Mongoose M3 и 4×1,8 ГГц, Cortex-A55, ОЗУ: 6 ГБ).

Как и ожидалось, сети MobileNets в несколько раз быстрее, чем более глубокие сверточные сети, и им требуется меньше памяти для хранения их весов. Хотя квантование уменьшает размер модели в 4 раза, оно не уменьшает время обработки изображений. И наконец, хотя время вычислений для ноутбука значительно меньше по сравнению с выводом в мобильных телефонах, все их более современные модели стали более пригодными для распознавания изображений в автономном режиме. Фактически, предлагаемая модель требует всего 60 мс для извлечения индивидуальных характерных признаков и распознавания всех атрибутов лиц (возраста, пола и т. д.), что позволяет осуществлять на этом устройстве комплексную аналитику лиц.

В следующих экспериментах сравнивается точность моделей в распознавании пола и предсказании возраста. Были использованы следующие наборы видеоданных:

- Eurecom Kinect, который содержит 9 фотографий для каждого из 52 человек (14 женщин и 38 мужчин).

- база данных Indian Movie Face (IMFDB) с 332 видеофайлами, содержащими 63 мужчин и 33 женщин. Предусмотрено всего четыре возрастные категории: "Ребенок" (0-15 лет), "Молодой" (16-35), "Среднего возраста" (36-60) и "Старый" (60+).

- Acted Facial Expressions in the Wild (AFEW) из конкурса аудиовизуального распознавания эмоций EmotiW 2018 (Emotions recognition in the wild). Содержит 1165 видеофайлов. Лицевые области детектируются с помощью метода MTCNN.

- IARPA Janus Benchmark A (IJB-A) с общим количеством кадров более 13000 на 1165 видеофайлах. В этом наборе данных имеется только информация о поле.

При распознавании пола на основе видео сначала классифицируется пол в каждом видеокадре. После этого используются две простые стратегии агрегации, а именно: простое голосование и правило умножения.

Во-первых, предложенные модели намного более точны, чем известные СНС. Этот факт можно объяснить предварительным обучением базовой MobileNet на задаче

идентификации лиц с очень большим набором данных, что облегчает извлечение высококачественных представлений лиц. Во-вторых, использование правила умножения обычно приводит к уменьшению частоты ошибок на 1-2 % по сравнению с простым голосованием. В-третьих, дообученная версия модели обеспечивает наименьшую частоту ошибок только для набора данных Kinect и менее точна на 1-3 % в других случаях. И наконец, хотя сжатие СНС позволяет значительно уменьшить размер модели, оно снижает коэффициент распознавания на величину до 7 %.

Для наборов данных Kinect и AFEW (с известным возрастом) считается распознавание возраста корректным, если разность между реальным и предсказанным возрастом не превышает 5 лет. Агрегация предсказаний возраста по отдельным видеокадрам осуществляется посредством: 1) простого голосования, 2) максимизации произведения апостериорных вероятностей возраста (1) и 3) усреднения ожидаемого значения (3) с выбором $L=3$ лучших предсказаний в каждом кадре.

Можно отметить, что и здесь эти модели более точны практически во всех случаях. Модели DEX сопоставимы с сетями СНС только для набора данных AFEW. Наименьшие коэффициенты ошибок получены для расчета ожидаемого значения предсказаний возраста. Например, этот расчет на 2 % и 8 % точнее, чем простое голосование для данных Kinect и AFEW. Этот эффект особенно заметен для изображений IMFDB, в которых ожидаемое значение приводит к повышению коэффициента распознавания на величину до 45 %.

На фиг. 6 показана блок-схема одного из примеров вычислительного устройства, на которой может быть реализована любая из одной или нескольких методик (например, методологий), рассмотренных в настоящем документе. В альтернативных вариантах осуществления изобретения машина 300 может работать как автономное устройство или быть подключенной к другим машинам (например, являться сетевой). При сетевом развертывании машина 300 может работать в качестве серверной машины, клиентской машины или обеих в сетевых средах "сервер-клиент". В одном из примеров машина 300 может дей-

ствовать как одноранговая машина в одноранговой (P2P) (или другой распределенной) сетевой среде. Машина 300 представляет собой персональный компьютер (ПК), планшетный ПК, телевизионную приставку (STB), персональный цифровой помощник (PDA), мобильный телефон, смартфон, веб-устройство, сетевой маршрутизатор, коммутатор или мост, серверный компьютер, базу данных, оборудование конференц-зала или любую машину, способную выполнять инструкции (последовательные или иные), которые определяют действия, подлежащие выполнению этой машиной. В различных вариантах осуществления изобретения машина 300 может выполнять один или несколько процессов, описанных выше. Кроме того, несмотря на то, что проиллюстрирована только одна машина, термин "машина" должен также включать любую совокупность машин, которые по отдельности или совместно выполняют набор (или несколько наборов) инструкций для выполнения любой одной или нескольких методик, обсуждаемых в настоящем документе, например, облачные вычисления, программное обеспечение как услуга (SaaS), другие конфигурации компьютерных кластеров.

Как описано в настоящем документе, приведенные примеры могут содержать или работать на основе логических или ряде из нескольких компонентов, модулей или механизмов (в дальнейшем совокупно именуемые "модулями"). Модули представляют собой материальные объекты (например, аппаратные средства), способные выполнять заданные операции и сконфигурированные или расположенные определенным образом. В одном из примеров схемы расположены (например, внутри или по отношению к внешним объектам, таким как другие схемы) определенным образом в виде модуля. В одном из примеров вся или часть одной или нескольких компьютерных систем (например, автономная, клиентская или серверная компьютерная система) или один или несколько аппаратных процессоров сконфигурированы с использованием микропрограммы или программного обеспечения (например, инструкций, части приложения или приложения) в качестве модуля, который работает для выполнения заданных операций. В одном из примеров программное обеспечение может

находиться на нетранзитном или другом машиночитаемом носителе. В одном из примеров программное обеспечение, выполняемое аппаратным обеспечением, лежащим в основе модуля, побуждает аппаратное обеспечение выполнять заданные операции.

Соответственно, под термином "модуль" понимается материальный объект, будь то объект, который физически создан, специально сконфигурирован (например, жестко подключен) или временно (например, транзитно) сконфигурирован (например, запрограммирован) для работы определенным образом или для выполнения части или всех операций, описанных в настоящем документе. При рассмотрении примеров, в которых модули временно сконфигурированы, каждый из модулей не обязательно должен быть инстанцирован в любой момент времени. Например, если в состав модулей входит аппаратный процессор общего назначения, настраиваемый с помощью программного обеспечения, аппаратный процессор общего назначения настраивается как соответствующие различные модули в разные моменты времени. Программное обеспечение может соответствующим образом сконфигурировать аппаратный процессор, например, так, чтобы он представлял собой определенный модуль в один момент времени и другой модуль в другой момент времени.

Машина (например, компьютерная система) 300 может включать аппаратный процессор 302 (например, центральный процессор (CPU), графический процессор (GPU), аппаратное процессорное ядро или любое их сочетание), основное запоминающее устройство 304 и статическое запоминающее устройство 306, некоторые или все из которых могут взаимодействовать друг с другом по каналу связи 308 (например, шине). Машина 300 может также содержать блок отображения 310, устройство буквенно-цифрового ввода 312 (например, клавиатуру) и устройство навигации пользовательского интерфейса (UI) 314 (например, мышь). В одном из примеров блок отображения 310, устройство ввода 312 и устройство навигации пользовательского интерфейса 314 представляют собой сенсорный дисплей. Машина 300 может дополнительно включать запоминающее устройство (например, накопитель) 316,

устройство генерирования сигналов 318 (например, громкоговоритель), устройство сетевого интерфейса 320 и один или несколько датчиков 321, таких как датчик глобальной системы позиционирования (GPS), компас, акселерометр или другой датчик. Машина 300 может включать контроллер вывода 328, такой как последовательный (например, универсальная последовательная шина (USB), параллельный или другое проводное или беспроводное соединение (например, инфракрасный порт (ИК), связь ближнего поля (NFC) и т. д.) для подключения или управления одним или несколькими периферийными устройствами (например, принтером, устройством чтения карт и другими).

Запоминающее устройство 316 может включать машиночитаемый носитель 322, на котором хранится один или несколько наборов структур данных или инструкций 324 (например, программное обеспечение), реализующих или использующих любую одну или несколько методик или функций, описанных в настоящем документе. Инструкции 324 могут также полностью или, по меньшей мере, частично находиться в основном запоминающем устройстве 304, в статическом запоминающем устройстве 306 или в аппаратном процессоре 302 во время их выполнения машиной 300. В одном из примеров, один или любая комбинация аппаратного процессора 302, основного запоминающего устройства 304, статического запоминающего устройства 306 или накопителя 316 могут представлять собой машиночитаемый носитель.

Хотя машиночитаемый носитель 322 показан как единый носитель, термин "машиночитаемый носитель" может включать единый носитель или несколько носителей (например, централизованную или распределенную базу данных и/или связанные кэши и серверы), сконфигурированных для хранения одной или нескольких инструкций 324.

Термин "машиночитаемый носитель" может включать любой носитель, который способен хранить, кодировать или переносить инструкции для выполнения машиной 300, которые побуждают машину 300 выполнять любой один или несколько способов настоящего изобретения, или который способен хранить, кодировать или переносить структуры данных, используемые или связанные с такими инструкциями. Без всяких ограниче-

ний примерами машиночитаемого носителя могут служить твердотельные запоминающие устройства, оптические и магнитные носители. К конкретным примерам машиночитаемых носителей можно отнести: энергонезависимые запоминающие устройства, такие как полупроводниковые запоминающие устройства (например, электрически программируемые запоминающие устройства с возможностью чтения (EPROM), электрически стираемые запоминающие устройства с возможностью чтения (EEPROM)) и устройства флэш-памяти; магнитные диски, такие как внутренние жесткие диски и съемные диски; магнитооптические диски; оперативные запоминающие устройства (RAM); твердотельные накопители (SSD); диски CD-ROM и DVD-ROM. В некоторых примерах машиночитаемые носители могут включать нетранзиторные машиночитаемые носители. В некоторых примерах машиночитаемый носитель может включать машиночитаемый носитель, который не является транзитным распространяющимся сигналом.

Инструкции 324 могут далее передаваться или приниматься по сети связи 326 с использованием среды передачи через устройство сетевого интерфейса 320. Машина 300 может взаимодействовать с одной или несколькими другими машинами, используя любой из множества протоколов передачи данных (например, ретрансляция кадров, интернет-протокол (IP), протокол управления передачей данных (TCP), протокол пользовательских дейтаграмм (UDP), протокол передачи гипертекста (HTTP) и другие).

Среди прочих к сетям связи можно отнести локальную сеть (LAN), глобальную сеть (WAN), сеть пакетной передачи данных (например, Интернет), сети мобильной телефонной связи (например, сотовые сети), стационарные телефонные сети (POTS), беспроводные сети передачи данных (например, семейство стандартных сетей

Института инженеров по электротехнике и электронике (IEEE) 802.11, известных как Wi-Fi®, семейство стандартов IEEE 802.16, известное как WiMax®, семейство стандартов IEEE 802.15.4, семейство стандартов Long Term Evolution (LTE), семейство стандартов Universal Mobile Telecommunications System (UMTS), одноранговые (P2P) сети и многие другие. В одном из примеров

35

устройство сетевого интерфейса 320 может включать один или несколько физических разъемов (например, Ethernet, коаксиальный или телефонный разъем) или одну или несколько антенн для подключения к сети связи 326. В одном из примеров устройство сетевого интерфейса 320 может включать в себя множество антенн для беспроводной связи с использованием, по меньшей мере, одного из методов однократного ввода-вывода (SIMO), многократного ввода-вывода (MIMO) или многократного ввода - однократного вывода (MISO). В некоторых примерах устройство сетевого интерфейса 320 может осуществлять беспроводную связь с использованием техники Multiple User MIMO.

Приведенные в настоящем документе примеры могут содержать или оперировать логическими или несколькими компонентами, модулями или механизмами. Модули представляют собой материальные объекты (например, аппаратные средства), способные выполнять заданные операции и сконфигурированные или расположенные определенным образом. В одном из примеров схемы расположены (например, внутри или по отношению к внешним объектам, таким как другие схемы) определенным образом в виде модуля. В одном из примеров вся или часть одной или нескольких компьютерных систем (например, автономная, клиентская или серверная компьютерная система) или один или несколько аппаратных процессоров сконфигурированы с использованием микропрограммы или программного обеспечения (например, инструкций, части приложения или приложения) в качестве модуля, который работает для выполнения заданных операций. В одном из примеров программное обеспечение может находиться на машиночитаемом носителе. В одном из примеров программное обеспечение, выполняемое аппаратным обеспечением, лежащим в основе модуля, побуждает аппаратное обеспечение выполнять заданные операции.

Соответственно, под термином "модуль" понимается материальный объект, будь то объект, который физически создан, специально сконфигурирован (например, жестко

36

подключен) или временно (например, транзиторно) сконфигурирован (например, запрограммирован) для работы определенным образом или для выполнения части или всех операций, описанных в настоящем документе. При рассмотрении примеров, в которых модули временно сконфигурированы, каждый из модулей не обязательно должен быть инстанцирован в любой момент времени. Например, если в состав модулей входит аппаратный процессор общего назначения, настраиваемый с помощью программного обеспечения, аппаратный процессор общего назначения настраивается как соответствующие различные модули в разные моменты времени. Программное обеспечение может соответствующим образом сконфигурировать аппаратный процессор, например, так, чтобы он представлял собой определенный модуль в один момент времени и другой модуль в другой момент времени.

Различные варианты реализации изобретения полностью или частично реализованы в программном обеспечении и/или микропрограммном обеспечении. Данное программное обеспечение и/или микропрограммное обеспечение может иметь форму инструкций, содержащихся на нетранзитивном считываемом компьютером носителе. Названные инструкции могут быть считаны и выполнены одним или несколькими процессорами для обеспечения выполнения операций, описанных в настоящем документе. Инструкции могут быть без каких-либо ограничений представлены в любой подходящей форме, такой как исходный код, скомпилированный код, интерпретируемый код, исполняемый код, статический код, динамический код и др. Такой компьютерно-читаемый носитель может содержать любой осязаемый нетранзитивный носитель для хранения информации в форме, считываемый одним или несколькими компьютерами, например, без каких-либо ограничений - постоянное запоминающее устройство (ROM); память с произвольным доступом (RAM); носители на магнитных дисках; оптические носители; флэш-память и другим.

Формула изобретения

1. Компьютерно-реализованный способ идентификации человека с использованием системы распознавания лиц, выполняемый, по меньшей мере, одним вычислительным устройством, включающий этапы, на кото-рых:

- обучают базовую сверточную нейронную сеть (СНС) на заранее выбранном наборе изображений;
- модифицируют указанную СНС посредством обеспечения, по меньшей мере, одного скрытого слоя с регуляризацией методом дропаута;
- формируют над, по меньшей мере, одним скрытым слоем независимые полно-связные слои для распознавания атрибутов лица, при этом каждый из указанных независи-мых слоев соответствует одному из атрибу-тов лица и содержит соответствующую нели-нейную функцию активации;
- извлекают набор мультимодальных атрибутов лица на основе базовой СНС, при-меняемой к первому представлению первого изображения человека;
- применяют выходные данные класси-фикатора атрибутов и выходные данные из-меренного расстояния, сгенерированные на основе второго представления второго изоб-ражения человека, к предварительно обучен-ной модели машинного обучения для выпол-нения проверки слияния;
- генерируют выходные данные систе-мы распознавания лиц на основе указанной модели проверки слияния;
- осуществляют контроль доступа к охраняемому объекту на основе выходных данных системы распознавания лиц.

2. Способ по п. 1, о т л и ч а ю щ и й с я тем, что атрибут включает, по меньшей мере, один из следующих параметров: возраст, пол, этническая принадлежность, угол поворота головы.

3. Способ по п. 1, о т л и ч а ю щ и й с я тем, что первое представление изображения включает, по меньшей мере, один вектор при-знаков.

4. Способ по п. 1, о т л и ч а ю щ и й с я тем, что выходные данные измеренного рас-стояния формируют путем применения пер-вого измерения расстояния к классификатору расстояний, при этом способ дополнительно включает:

- выполнение встраивания лица (face embedding);
- применение указанного встраивания к классификационному конвейеру для иденти-фикации похожего изображения;
- формирование второго измерения рас-стояния на основе указанного идентифициро-ванного изображения;
- применение второго измерения рас-стояния к классификатору расстояний для получения выходных данных измеренного расстояния.

5. Способ по п. 1, дополнительно вклю-чающий:

- применение выходных данных систе-мы распознавания лиц к совместной модели классификации;
- применение вспомогательного атрибу-та к совместной модели классификации;
- формирование выходных данных сов-местной классификации.

6. Способ по п. 5, о т л и ч а ю щ и й с я тем, что вспомогательный атрибут включает временной признак.

7. Способ по п. 5, о т л и ч а ю щ и й с я тем, что вспомогательный атрибут включает социальный статус человека.

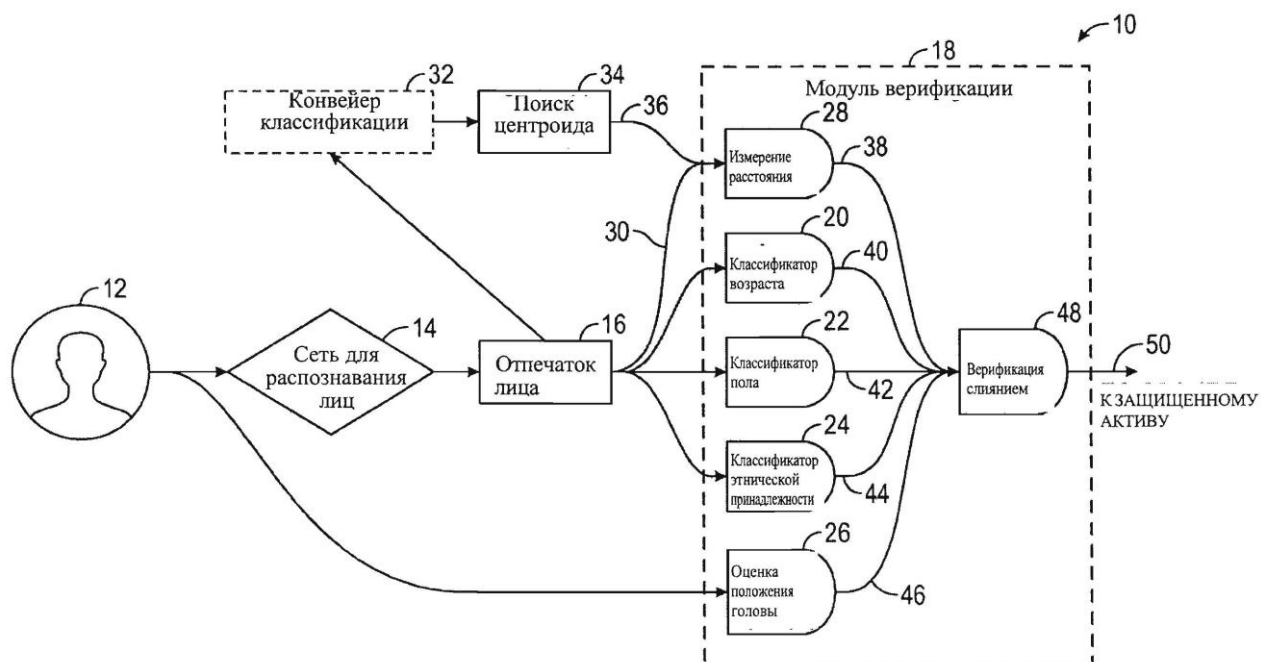
8. Способ по п. 5, о т л и ч а ю щ и й с я тем, что вспомогательный атрибут включает рост человека.

9. Способ по п. 1, о т л и ч а ю щ и й с я тем, что первое представление изображения идентично второму представлению изобра-жения.

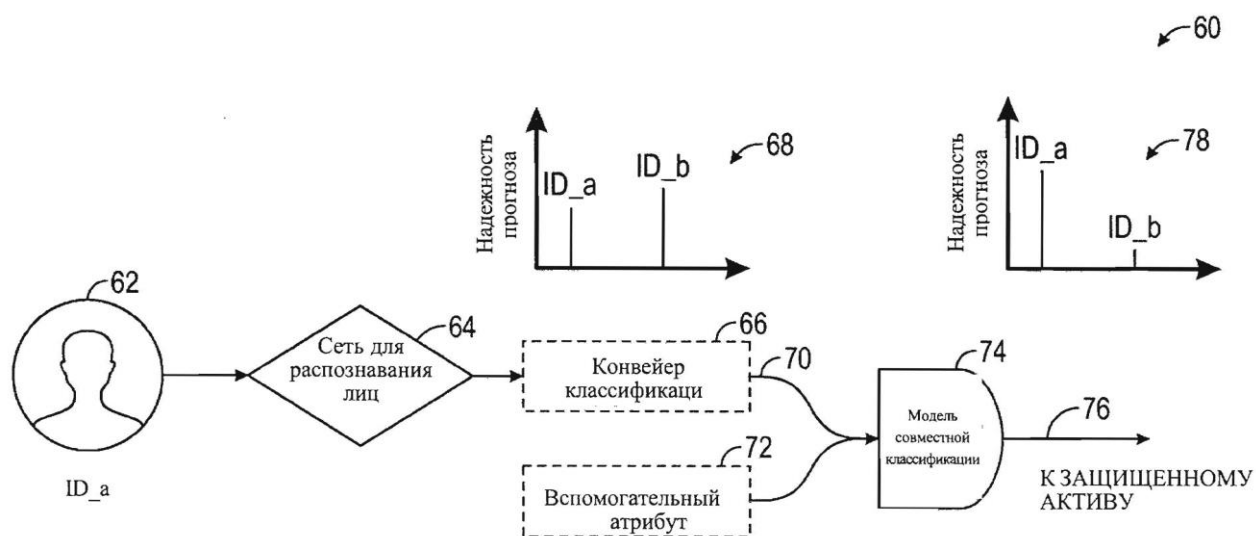
10. Способ по п. 1, о т л и ч а ю щ и й с я тем, что первое изображение идентично второму изображению.

11. Способ по п. 1, о т л и ч а ю щ и й с я тем, что контроль доступа к охраняемому объекту включает управление доступом к входу в здание.

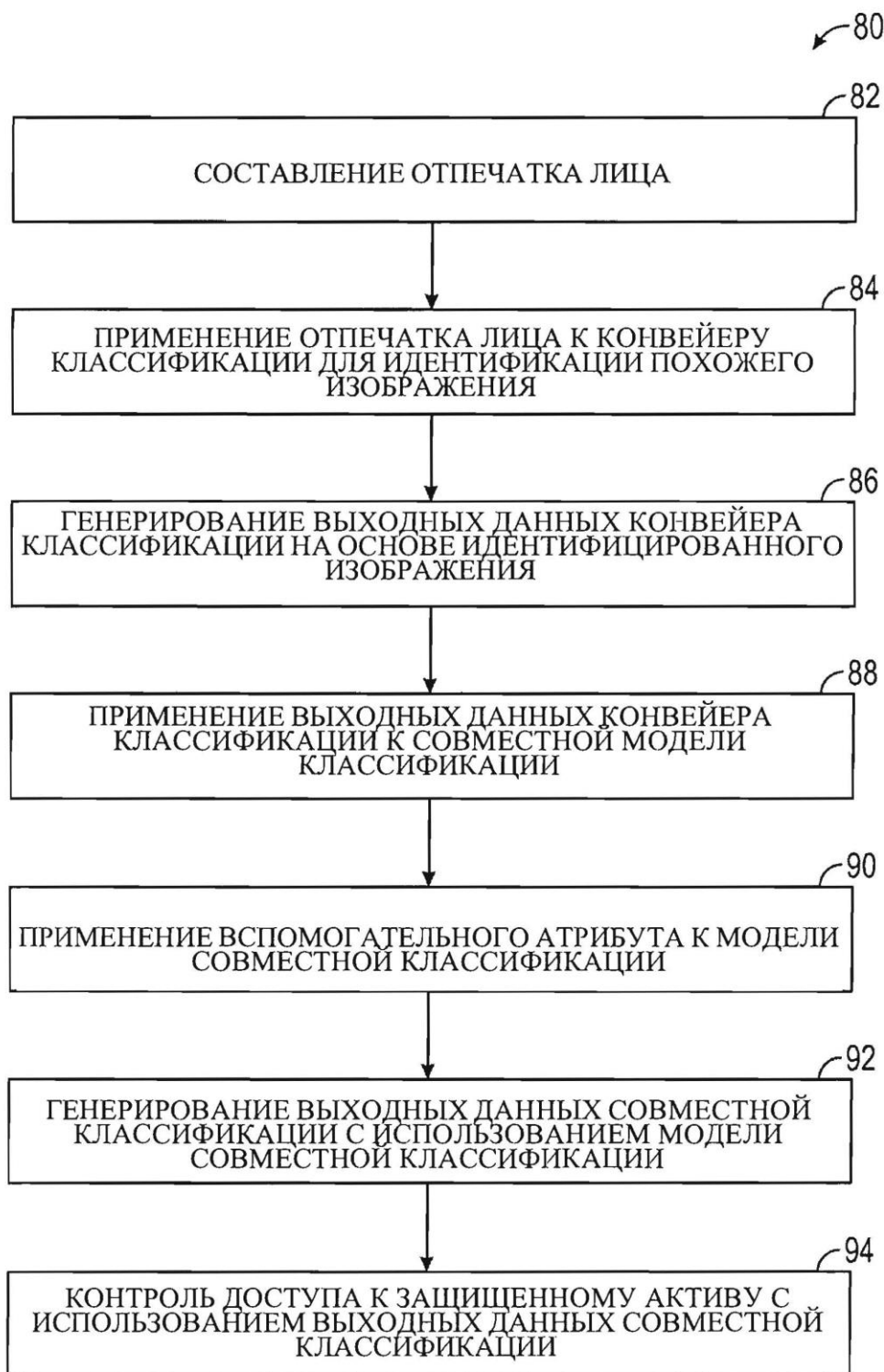
Компьютерно-реализованный способ использования системы распознавания лиц



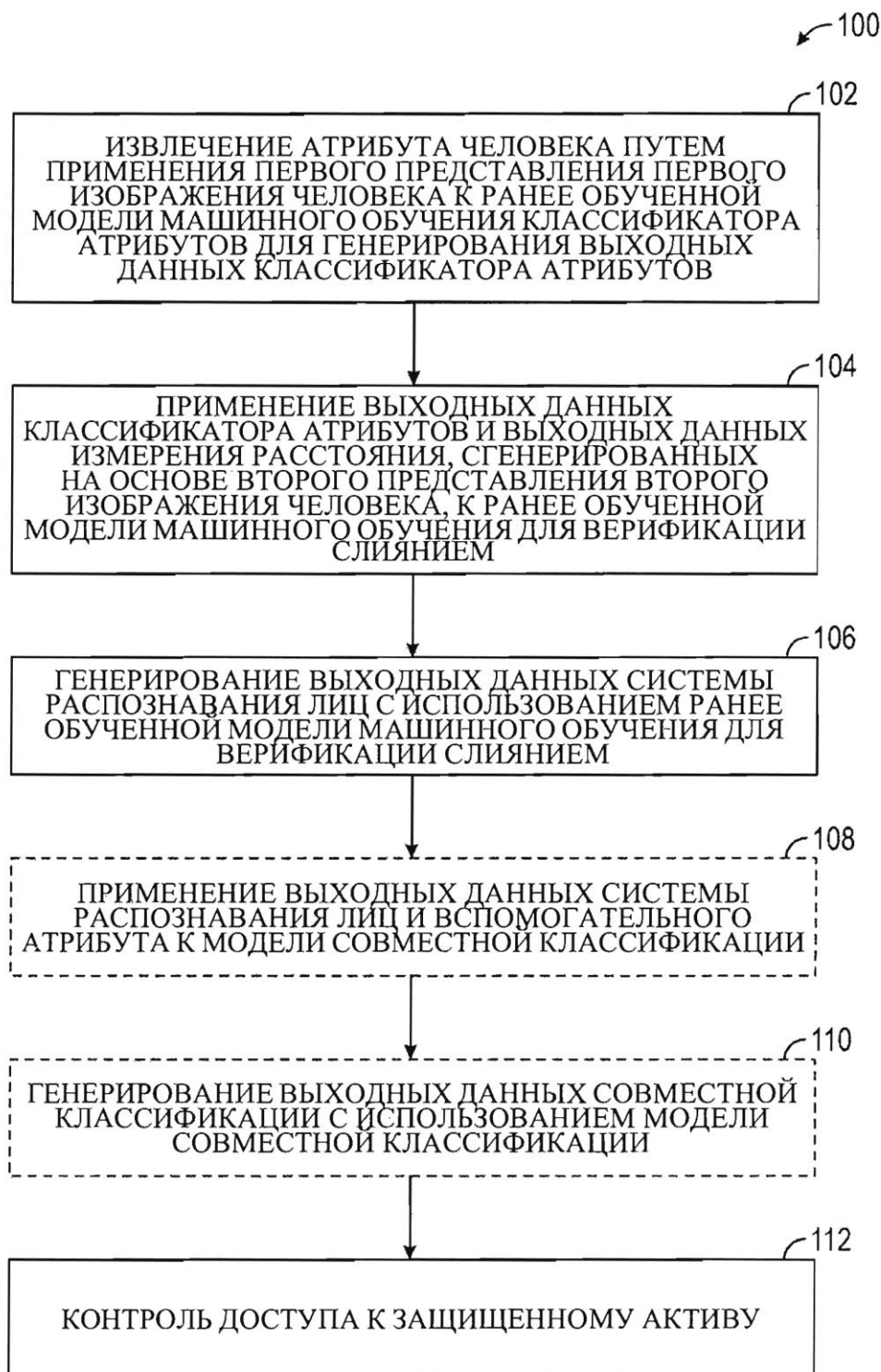
Фиг. 1



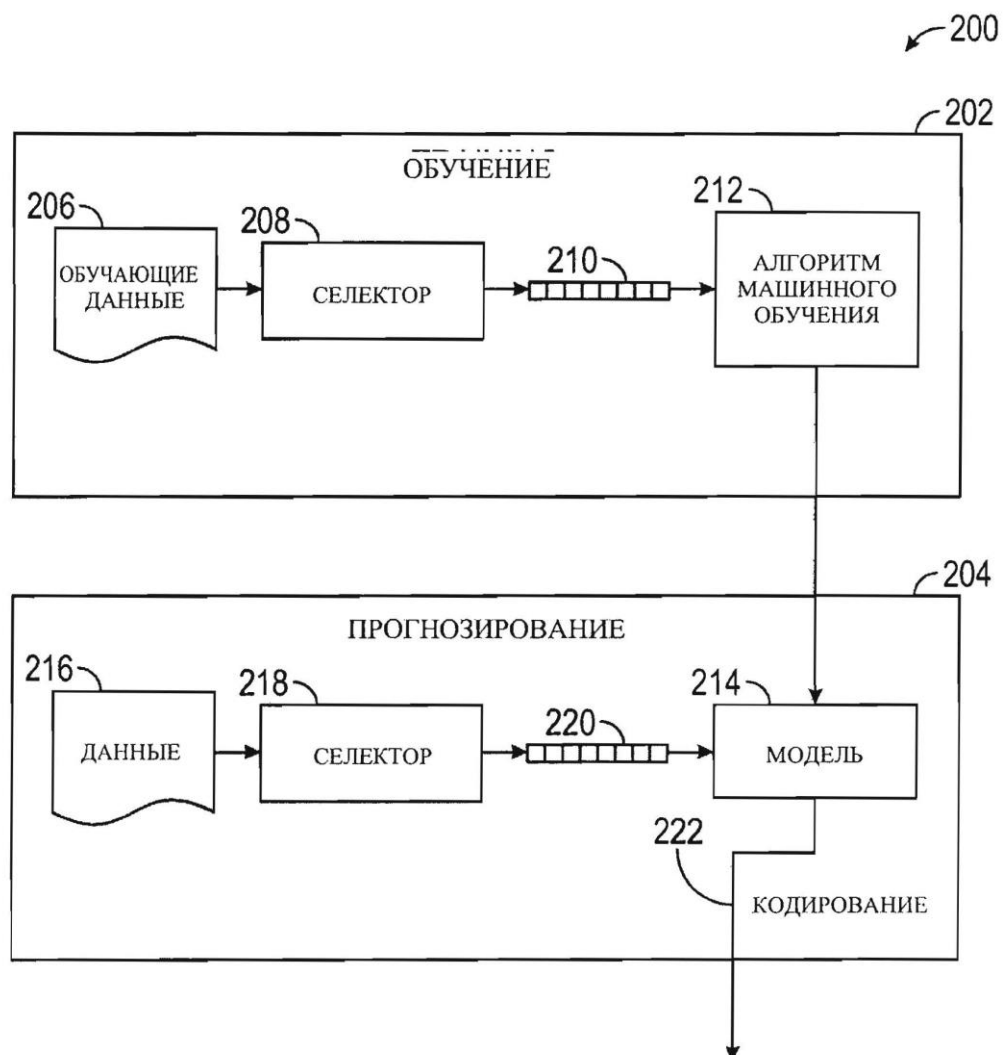
Фиг. 2



Фиг. 3

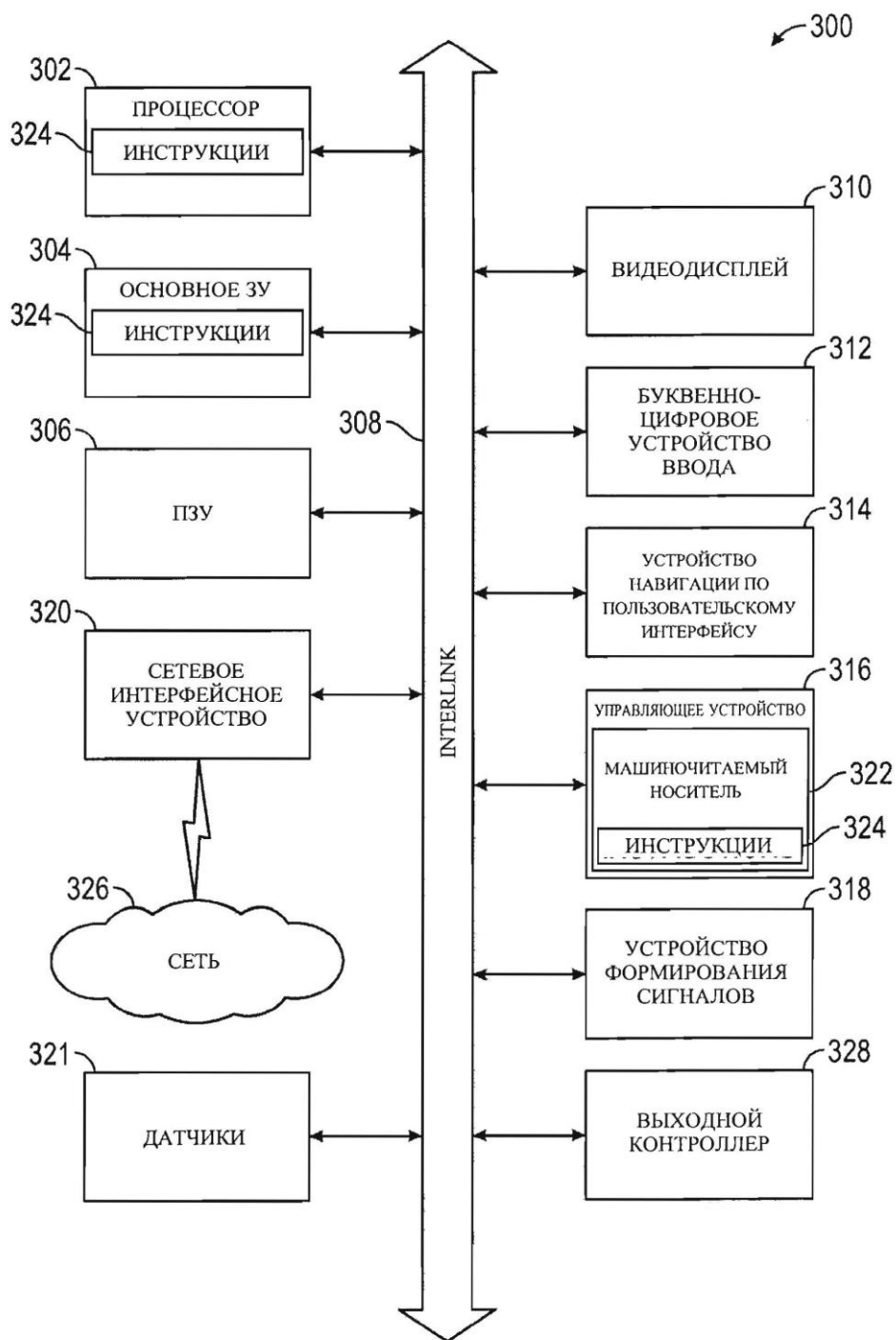


Фиг. 4



Фиг. 5

Компьютерно-реализованный способ использования системы распознавания лиц



Фиг. 6

Выпущено отделом подготовки официальных изданий