

ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

Челябинский физико-математический журнал. 2022. Т. 7, вып. 4. С. 490–504.

УДК 004.056+004.94+004.89

DOI: 10.47475/2500-0101-2022-17408

КОМБИНИРОВАННЫЙ 2D/3D-ПОДХОД ДЛЯ ПОВЫШЕНИЯ НАДЁЖНОСТИ СИСТЕМ РАСПОЗНАВАНИЯ ЛИЦ С ПРИМЕНЕНИЕМ ГЛУБОКОГО ОБУЧЕНИЯ

К. А. Дорофеев^{1,a}, А. Н. Ручай^{1,2,b}

¹ *Челябинский государственный университет, Челябинск, Россия*

² *Южно-Уральский государственный университет*

(национальный исследовательский университет), Челябинск, Россия

^a *Kostuan1989@mail.ru*, ^b *ran@csu.ru*

Широко известны системы распознавания лиц, использующие цветные двумерные изображения. Однако их основной проблемой является неустойчивость к различным условиям освещения, эмоциональным и мимическим выражениям, перекрытиям и углам поворота при считывании кадра. В статье предложен новый подход, комбинирующий обработку двумерных данных в виде цветного изображения и трёхмерных данных в виде облака точек, или карта глубины. Особое внимание было уделено вопросам архитектуры нейронных сетей, качеству и точности распознавания. Было предложено несколько комбинированных методов распознавания лиц с применением машинного обучения и глубокого обучения. Проведён сравнительный анализ результатов экспериментов в терминах точности распознавания на открытых базах данных лиц. Был выбран самый надёжный комбинированный метод распознавания лиц. Основная цель — построение надёжной биометрической системы распознавания лиц, устойчивой к сложным внешним факторам, таким, как мимические экспрессии, изменения масштаба, освещения, частичное перекрытие посторонними предметами, большие углы поворота.

Ключевые слова: *информационная безопасность, биометрия, распознавание лиц, нейронные сети, машинное обучение, глубокое обучение, идентификация, аутентификация, мультибиометрия.*

Введение

Системы обнаружения и распознавания лиц являются неотъемлемой частью нашей жизни с момента их появления в начале 1960-х годов. Наиболее важная область применения распознавания лиц — системы безопасности [1]. Также распознавание лиц применяется в медицине, при проектировании умных городов, умных магазинов без касс [2; 3]. Интерес к надёжным системам распознавания лиц обусловлен высокими финансовыми потерями от несанкционированного доступа, от действий киберпреступников по всему миру.

В настоящее время в связи с развитием технологий машинного обучения и нейронных сетей стали доступны новые надёжные алгоритмы и подходы к реализации систем распознавания лиц. Достигнуты серьёзные успехи в решении задач распознавания человека по лицу благодаря применению нейронных сетей глубокого обучения с тремя и более слоями [4–9], которые объединяют в себе как выбор и расчёт признаков, так и классификацию. Эксперты отмечают, что точность методов глубокого обучения при использовании очень большой обучающей выборки и при сложной архитектуре сети в контролируемых условиях зачастую превышает точность распознавания человеком [10–12].

Основными проблемами при использовании двумерного подхода являются различные условия освещения, эмоциональные и мимические выражения, перекрытия части лица, углы поворота и расположение головы при считывании кадра. Системы распознавания лиц в реальном мире должны быть нечувствительны к таким изменениям, чтобы считаться эффективными. Информации о человеческом лице на обычном двумерном изображении может быть недостаточно для точного распознавания. Именно поэтому исследователи и разработчики начали использовать трёхмерные данные [13; 14]. Объединение информации о карте глубины и текстуре лица приводит к увеличению точности распознавания. Информация о глубине не зависит от изменения освещения и условия захвата изображения. Важным фактором также явилось появление экономичных и широкодоступных RGB-D датчиков, таких как Microsoft Kinect, Intel RealSense и прочих, что дополнительным образом подстегнуло отрасль к развитию.

Однако стоит отметить, что развитие трёхмерного распознавания лиц происходит медленнее с использованием камер RGB-D, чем двумерного распознавания лиц, по разным причинам: отсутствие достаточно больших наборов данных о лицах в 3D-варианте, доступных для использования в открытом виде; отсутствие эффективного способа передачи и обработки трёхмерных данных нейронной сетью; высокая стоимость как аппаратной, так и программной части подобного рода систем, сложность в обслуживании.

Разные исследователи пытаются решить сопутствующие проблемы [15–17], например, применяя техники дополнения обучающего набора данных (data augmentation) [18–20]. Известны результаты преобразования двумерных баз данных трёхмерного распознавания лиц с помощью применения модулей оценки глубины (depth estimation) [21].

Кроме того, во многих исследованиях используют нейронные сети, обученные на двумерных изображениях лиц, и применяют методы переноса обучения (transfer learning) и тонкой настройки (fine tuning) для работы с трёхмерными данными [18; 22; 23].

Исследователи предложили следующие варианты передачи трёхмерных данных на вход нейронной сети: передача данных глубины, азимута, углов вектора нормали в качестве третьего канала входных данных [19]; нормализация данных глубины к тому же диапазону значений, как у RGB-кадра [24]; в работах [25; 26] предложено использовать облака точек вместо данных глубины; проецирование облаков точек в карту диапазонов [18]; использование многовидовых изображений [27]; вокселизация в объёмное представление [28].

1. Предложенный комбинированный 2D/3D-подход

1.1. База данных и используемые инструменты

Для тестов использовалась база данных Vosphorus [29], поскольку база предназначена для исследований задач обработки человеческих лиц в двумерных и трёхмерных вариантах, включая распознавание выражений, обнаружение лицевых движений, распознавание лиц в сложных, неблагоприятных условиях, деформируемое моделирование и трёхмерную реконструкцию лица. В базе содержится 4666 сканов 105 людей в различных позах, с разной мимикой, эмоциями, перекрытием части лица.

Разбиение кадров базы производилось следующим образом: в качестве обучающей выборки использовались кадры без перекрытия лица, с ровным фронтальным расположением (без поворотов), с нейтральными эмоциональным и мимическим выражениями (имена таких файлов в базе подходят под шаблон *NN*); в качестве тестовой выборки использовались все остальные кадры того же субъекта (повороты лица, изменения мимики, эмоций, перекрытия). При таком разбиении размер обучающей выборки — 299 кадров, размер тестовой выборки — 4367 кадров.

1.2. Модель VggFace2 для распознавания двумерного изображения



Рис. 1. Пример цветного изображения из базы Vosphorus

В качестве нейронной сети для обработки цветного двумерного изображения была выбрана модель VggFace2 [30]. Авторы проекта подготовили большой набор данных размером 3.31 миллиона изображений для 9131 субъекта, а также выложили в открытый доступ предобученные модели сетей VggFace2. Модели обучались на наборе данных MS-Celeb-1M [31], а затем производилась тонкая настройка (fine tuning) на наборе данных VggFace2. Основой для архитектуры моделей послужила ResNet-50. На вход нейронной сети подаётся цветное двумерное изображение лица человека разрешением 224×224 пикселя (рис. 1).

Для адаптации модели к базе данных Vosphorus была произведена тонкая настройка (fine tuning), состоящая из шагов:

- 1) срезание у предобученной модели всех слоёв после "avgpool";
- 2) добавление полносвязного слоя с функцией активации softmax, с количеством выходов, соответствующим количеству классов (в базе Vosphorus это значение равно 105);
- 3) дообучение нейронной сети на базе Vosphorus.

1.3. Модель IrisFaceRgbD для распознавания трёхмерных данных

В качестве нейронной сети для работы с картой глубины была выбрана модель IrisFaceRgbD [18]. Зачастую приходится иметь дело не с картами глубины, а с облаками точек, поэтому авторы предложили использовать ортогональную проекцию облака точек и подавать на вход нейронной сети двумерные серые изображения. Авторы проекта взяли за основу нейронную сеть VggFace первой версии, провели тонкую настройку и выложили предобученную модель в открытый доступ. На вход

модели подаётся двумерное изображение в виде плоской ортогональной проекции облака точек с разрешением 224×224 пикселя (рис. 2).

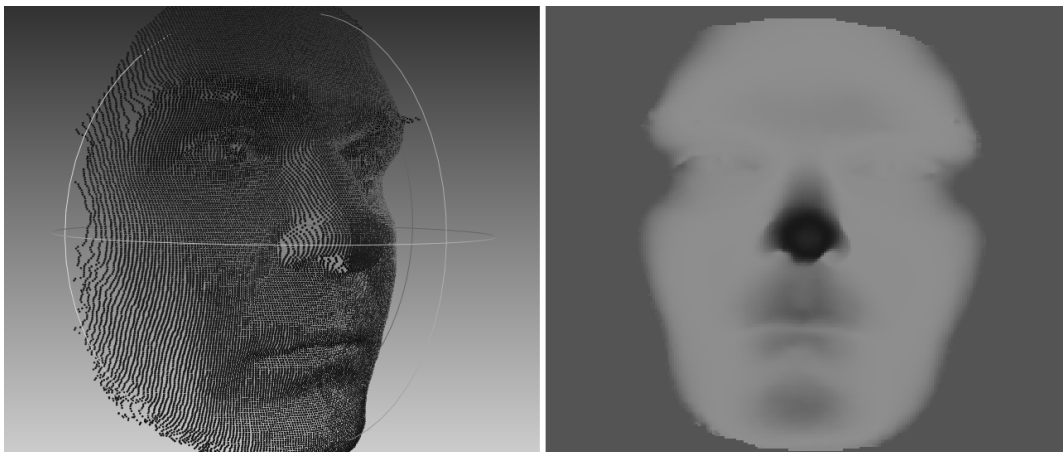


Рис. 2. Пример облака точек и его плоской ортогональной проекции из базы Vosphorus

Для адаптации модели к базе данных Vosphorus была произведена тонкая настройка (fine tuning), состоящая из шагов:

- 1) срезание у предобученной модели всех слоёв после "block5pool";
- 2) добавление полносвязного слоя, содержащего 2048 нейронов с функцией активации relu;
- 3) добавление слоя регуляризации (dropout) для уменьшения переобучения сети;
- 4) добавление полносвязного слоя с функцией активации softmax, с количеством выходов, соответствующим количеству классов (в базе Vosphorus это значение равно 105);
- 5) дообучение нейронной сети на базе Vosphorus.

1.4. Комбинированный метод принятия решения на основе выходов нескольких нейронных сетей

Самым простым решением при построении комбинированного 2D/3D-подхода может быть метод принятия решения на основе выходов нескольких нейронных сетей (DMM — decision making method). В таком случае при идентификации человека по лицу подаётся цветное двумерное изображение в нейронную сеть VggFace2, а плоская проекция облака точек подаётся в нейронную сеть IrisFaceRgb, в итоге получается два выхода, которые можно анализировать по некоторому правилу (рис. 3).

Таких правил можно предложить достаточно много, например, принимать решение о положительной идентификации субъекта при выполнении хотя бы одного из следующих условий:

- 1) выходы двух нейронных сетей совпадают, когда обе сети уверены, что это один и тот же субъект;
- 2) какая-то из нейронных сетей имеет выход с большой уверенностью, например, больше некоторого порогового значения.

При применении такого подхода точность на тестовой выборке составляет 0.94, что не сильно отличается от использования только одной нейронной сети VggFace2. Однако тяжело подобрать наиболее точное и эффективное правило вручную, поэтому имеет смысл использовать машинное обучение.

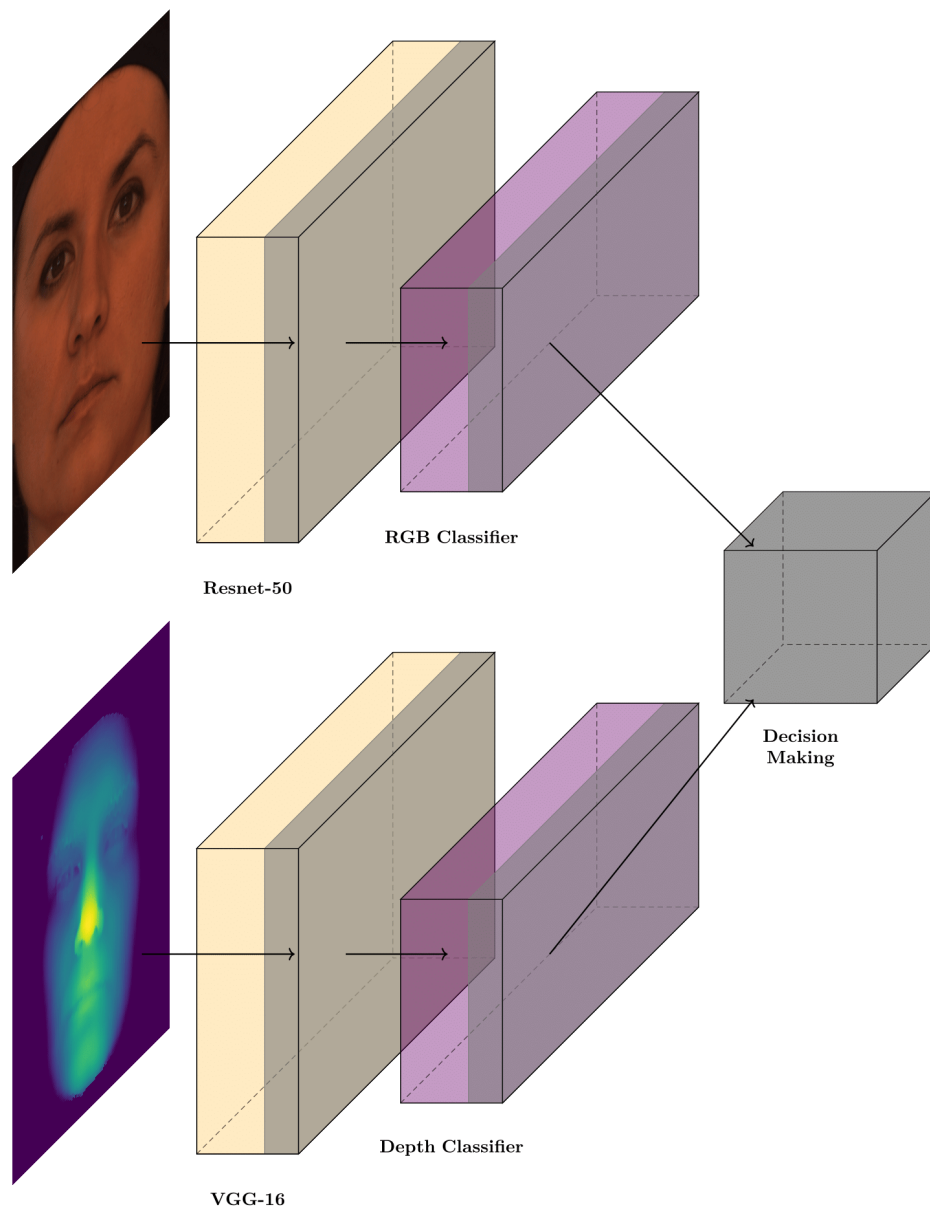


Рис. 3. Метод принятия решения на основе двух нейронных сетей DMM

1.5. Комбинированный метод с применением машинного обучения

Выходы нейронных сетей содержат векторы значений в виде вероятности классификации субъекта. Эти векторы можно использовать для задачи классификации с помощью машинного обучения. Был предложен комбинированный метод с применением машинного обучения (MLD, machine learning decision), когда цветное двумерное изображение подаётся в нейронную сеть VggFace2, и плоская проекция облака точек подаётся в нейронную сеть IrisFaceRgb, выходы нейронных сетей в виде векторов вероятностей объединяются в общий вектор признаков для классификации с помощью машинного обучения.

Разбиение на обучающую и тестовую выборку при применении машинного обучения производилось схожим образом (см. раздел 1.1). Для машинного обучения использовалась библиотека `sklearn`, для анализа данных было применено большинство из основных методов библиотеки: метод опорных векторов (SVC), метод линейных опорных векторов (LinearSVC), метод k -ближайших соседей (KNeighborsClassifier), пассивно-агрессивный классификатор (PassiveAggressiveClassifier), классификатор гауссова процесса (GaussianProcessClassifier), классификатор случайного леса (RandomForestClassifier), классификатор повышения ансамбля

(AdaBoostClassifier), дерево решений (DecisionTreeClassifier), логистическая регрессия (LogisticRegression), гауссовский наивный байесовский метод (GaussianNB), линейный дискриминантный анализ (LinearDiscriminantAnalysis), многоуровневый перцептрон (MLPClassifier), классификатор дополнительных деревьев (ExtraTreesClassifier), голосование среди соседей в заданном радиусе (RadiusNeighborsClassifier), классификатор опорных векторов Nu (NuSVC).

Стоит отметить, что многие методы содержат достаточно большое количество параметров, которые тяжело перебирать вручную. Помогает разрешить эту проблему библиотека `autosklearn`. `Autosklearn` позволяет автоматизировать процесс поиска оптимальных гиперпараметров модели, избавить разработчика от рутинных задач. В результате работы библиотеки было найдено лучшее решение — классификатор `LinearSVC` со следующим набором параметров: `C=6.3959`, `dual=False`, `intercept_scaling=1.0`, `random state=1`, `tol=0.0001642`.

1.6. Комбинированный метод с применением глубокого обучения

В качестве альтернативы применения машинного обучения можно построить нейронную сеть, на вход которой подавать выходы нейронных сетей `VggFace2` и `IrisFaceRgbd` после слоёв классификации (DNND, deep neural network decision) (рис. 4).

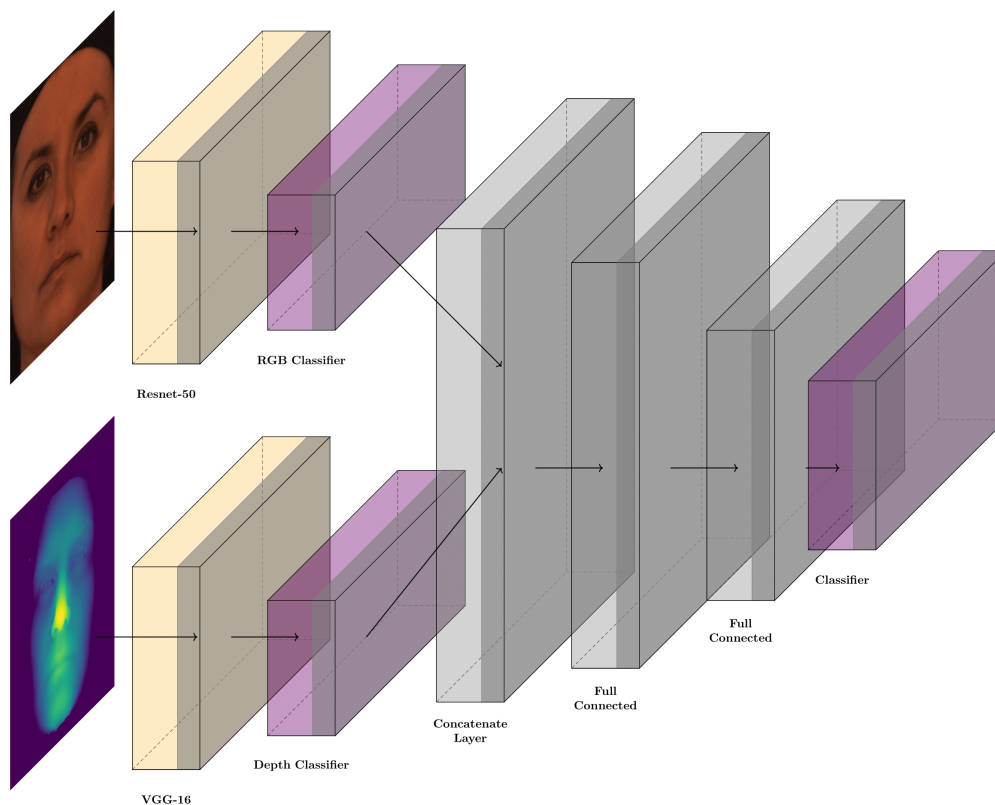


Рис. 4. Схема нейронной сети для принятия решения DNND

Лучших результатов удалось добиться при следующей архитектуре нейронной сети (табл. 1). Общее число обучающихся параметров составляет 136 миллионов. По аналогии с применением машинного обучения необходимо перебирать большое количество параметров нейронной сети, тестировать разные архитектуры. Для этих целей можно применять библиотеку `autokeras`, однако этот инструмент сильно требователен к вычислительным ресурсам, как к характеристикам оборудования, так и ко времени работы.

Наиболее точным и удобным в применении оказался предложенный комбинированный метод на основе глубокой нейронной сети с множественными входами и смешанными данными (DNNMIMD, deep neural network with multiple inputs and mixed data) (см. рис. 5). При таком подходе входной вектор формируется с помощью конкатенации данных цветного двумерного изображения и карты глубины. Модель нейронной сети представляет собой параллельную комбинацию (в этом случае модель не является последовательной) слоёв VggFace2 (включительно до слоя "avgpool") и IrisFaceRgb (включительно до слоя "block5pool"). Отличие DNNMIMD-метода от нейронной сети для принятия решения DNND заключается в том, что после указанных слоёв не добавляются слои классификации, а формируется новый конкатенированный слой признаков двух сетей. После чего добавляются полносвязный слой с 4096 нейронами и с функцией активации relu, слой регуляризации (dropout), ещё один полносвязный слой с 2048 нейронами и с функцией активации relu, слой регуляризации (dropout) для уменьшения переобучения сети, и последний слой классификации с функцией активации softmax. Общее число параметров такой сети составляет 94 миллиона, обучающихся параметров — 55 миллионов.

Таблица 1

Архитектура нейронной сети для принятия решения на основе метода DNND

Model: "sequential"		
Layer	Output shape	Params
Fc1 (Dense)	None, 16256	3430016
Dropout	None, 16256	0
Fc2 (Dense)	None, 8128	132136896
Dropout	None, 8128	0
Classifier	105	853545

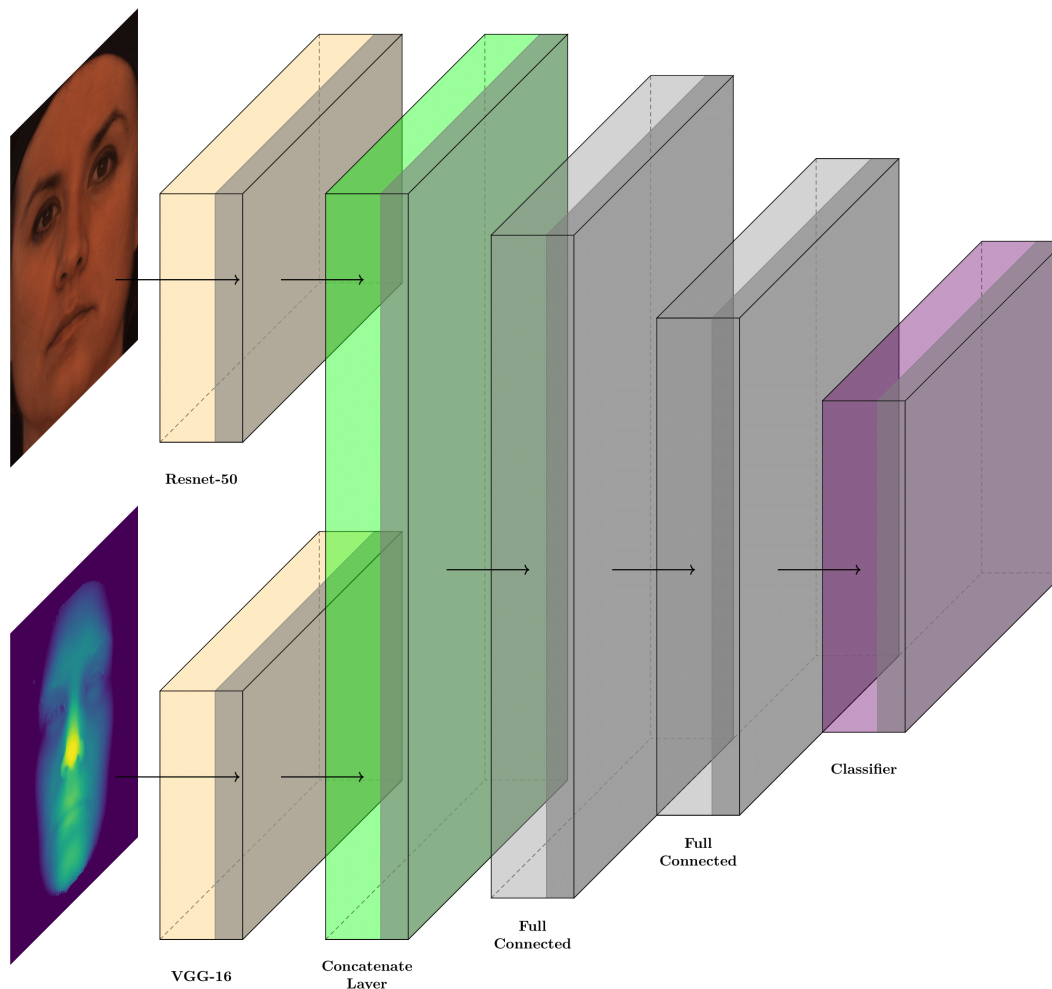


Рис. 5. Архитектура глубокой нейронной сети с множественными входами и смешанными данными DNNMIMD

2. Результаты экспериментов

Для тестов использовалась база данных Vosphorus [29], поскольку данные этой базы наиболее точно подходят под основные задачи исследования. Для написания кода использовался язык python, фреймворк keras и бесплатная интерактивная облачная среда для работы с кодом google colab.

История обучения модели VggFace2 представлена на рис. 6. Результаты обучения и тестирования VggFace2 в терминах точности показаны в табл. 2. Хорошим показателем является отсутствие переобучения. Разница точности на обучающей и тестовой выборке составляет всего 3%.

Таблица 2
Точность обучения и тестирования модели VggFace2

Обучение		
	Accuracy	Loss
Эпоха 1/2	0.3813	5.9956
Эпоха 2/2	0.9666	0.1618
Тестирование		
	Accuracy	Loss
	0.9380	0.2111
Вес модели		
	VggFace2Resnet-50	VggFace2TransferLearning
	90.4 mb	92.9 mb

Также можно констатировать, что нейронная сеть VggFace2 вполне позволяет обрабатывать цветные двумерные изображения из базы Vosphorus с точностью 0.9380 на тестовой выборке, однако применение комбинированного метода позволяет улучшить точность.

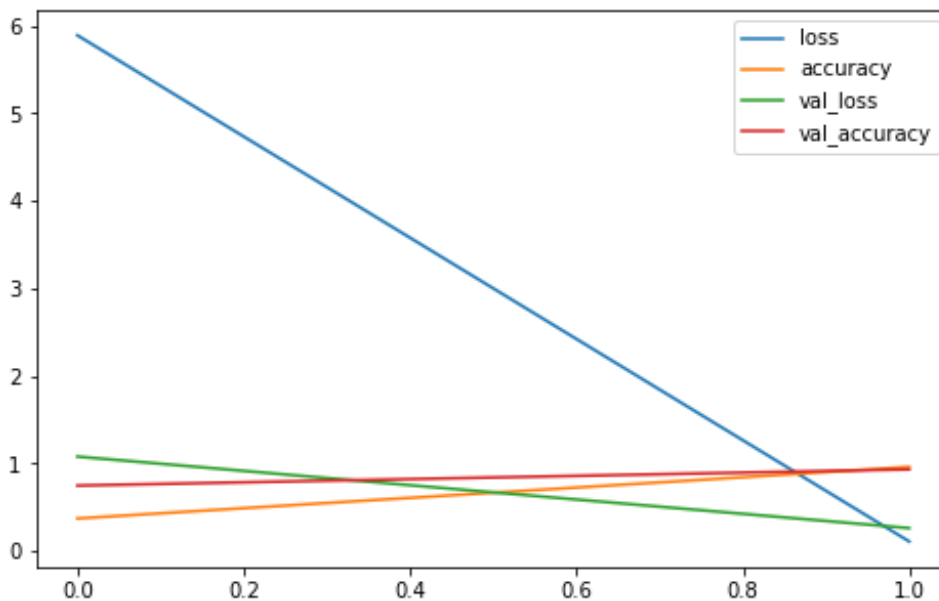


Рис. 6. История обучения модели VggFace2

История обучения модели IrisFaceRgbD представлена на рис. 7. Результаты обучения и тестирования IrisFaceRgbD в терминах точности представлены в табл. 3. Присутствует более заметная разница в точности на обучающей и тестовой выборках. Это можно объяснить тем, что в сложных условиях (например, большие углы поворота) данных только в карте глубины недостаточно, чтобы нейронная сеть могла правильно идентифицировать объект.

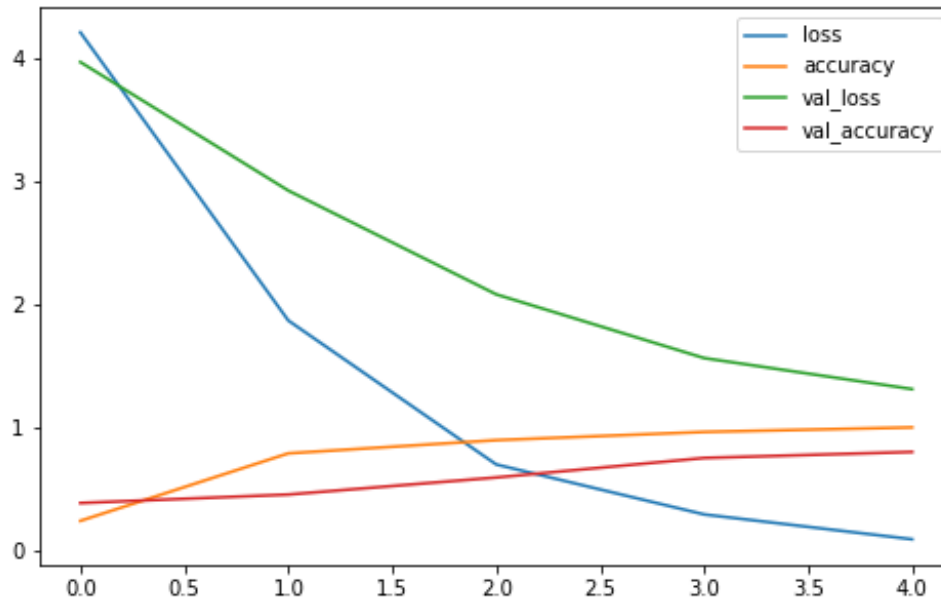


Рис. 7. История обучения модели IrisFaceRgbD

Результаты применения машинного обучения для принятия решения MLD приведены в табл. 4. Лучших результатов удалось достичь с помощью KNeighborsClassifier. Известно, что метод k -ближайших соседей хорошо подходит для мультимодальных задач, хорошо обрабатывает данные изображений. В компьютерном зрении считается прекрасным вариантом для классификации множества разнообразных изображений на основе сходства.

Таблица 3
Точность обучения и тестирования модели IrisFaceRgbD

Обучение		
	Accuracy	Loss
Эпоха 1/5	0.2174	4.2521
Эпоха 2/5	0.7659	1.9365
Эпоха 3/5	0.8863	0.7219
Эпоха 4/5	0.9202	0.3010
Эпоха 5/5	0.9632	0.0905
Тестирование		
	Accuracy	Loss
	0.8544	0.6265
Вес модели		
	IrisFaceRgbD	IrisFaceRgbDTransferLearning
	521 mb	660 mb

Таблица 4

Точность машинного обучения при принятии решения MLD

Метод	Accuracy
SVC	0.9439
LinearSVC	0.9476
KNeighborsClassifier	0.9480
PassiveAggressiveClassifier	0.9478
GaussianProcessClassifier	0.9370
RandomForestClassifier	0.7096
AdaBoostClassifier	0.1504
DecisionTreeClassifier	0.4774
LogisticRegression	0.8823
GaussianNB	0.6387
LinearDiscriminantAnalysis	0.8273
MLPClassifier	0.0767
ExtraTreesClassifier	0.6883

Таблица 5

Результаты обучения и тестирования нейронной сети для принятия решения DNND

Обучение		
	Accuracy	Loss
Эпоха 1/4	0.3645	4.0153
Эпоха 2/4	0.7763	0.9670
Эпоха 3/4	0.9231	0.5288
Эпоха 4/4	0.9805	0.0294
Тестирование		
	Accuracy	Loss
	0.9505	0.3752
Вес модели		
	DNND	
	1520 mb	

Результаты обучения и тестирования нейронной сети DNND представлены в табл. 5. Можно отметить отсутствие переобучения и хорошие результаты на тестовой выборке.

Результаты обучения и тестирования глубокой нейронной сети с множественными входами DNNMIMD представлены в табл. 6. Метод DNNMIMD оказался лучшим в терминах точности на тестовой выборке, однако его результаты можно ещё улучшить, если найти более оптимальные архитектуры глубокой нейронной сети.

Таблица 6

Результаты обучения и тестирования комбинированной глубокой нейронной сети с множественными входами и смешанными данными DNNMIMD

Обучение		
	Accuracy	Loss
Эпоха 1/3	0.4515	5.1402
Эпоха 2/3	0.8595	0.6280
Эпоха 3/3	0.9799	0.1051
Тестирование		
	Accuracy	Loss
	0.9601	0.2731
Вес модели		
	DNNMIMD	
	785 mb	

Таблица 7

Итоговое сравнение комбинированных методов на базе Bosphorus

Метод	Точность
VggFace2	0.9380
IrisFaceRgb	0.8544
DMM	0.9400
MLD	0.9485
DNND	0.9505
DNNMIMD	0.9601

Итоговое сравнение результатов применения комбинированных методов в терминах точности на тестовой выборке приведены в табл. 7. Комбинированный метод на основе глубокой сети DNNMIMD имеет самую большую точность 0.9601 на тестовой выборке. Такой результат является вполне объяснимым, поскольку именно концепция глубокого обучения является наиболее гибкой, масштабируемой и настраиваемой. Существуют разные задачи, которые требуют мультимодальных входов, объединяющих данные, поступающие из разных источников. Для обработки каждого вида данных используются разные типы нейронных слоёв или даже целых нейронных сетей. Построение надёжной системы распознавания лиц вполне относится к такого рода задачам, поскольку для её решения можно комбинировать разнородные данные в виде двумерных, трёхмерных или других и повышать таким образом надёжность системы.

Заключение

Был предложен новый подход, комбинирующий обработку двумерных в виде цветного изображения и трёхмерных данных в виде облака точек или карта глубины, для повышения надёжности систем распознавания лиц. Было реализовано несколько комбинированных методов с использованием глубокого обучения. Была доказана эффективность применения комбинированного подхода в сравнении с двумерным распознаванием лиц на основе модели VggFace2 и трёхмерным распознаванием лиц на основе модели IrisFaceRgb. В результате самым надёжным методом является комбинированная глубокая нейронная сеть DNNMIMD с точностью 0.9601 на тестовой выборке.

Среди основных направлений и задач для будущих исследований можно указать следующие: поиск новых архитектур глубоких нейронных сетей; повышение точности распознавания в сложных условиях; подготовка и формирование больших наборов данных трёхмерных лиц; поиск эффективных способов передачи трёхмерных данных в нейронную сеть; решение проблем с высокой стоимостью комбинированных систем, со сложностью разворачивания и обслуживания.

Список литературы

1. **Callaway S., Cheng J., Contratti A., Fu D., Gelivi H., Wachulec J., Purohit S.** Comparative analysis of image processing algorithms for airport security // 2020 IEEE MIT Undergraduate Research Technology Conference (URTC), 2020.
2. **Bansal M., Sharma D.** Facial recognition system for security resolutions in smart city // International Journal of Advanced Research in Engineering and Technology. 2020. Vol. 11, no. 10. P. 146–151.
3. **Praveen G., Dakala J.** Face recognition: Challenges and issues in smart city/environments // International Conference on Communication Systems and Networks. 2020. P. 791–793.
4. **Taigman Y., Yang M., Ranzato M., Wolf L.** Deepface: Closing the gap to human-level performance in face verification // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. 2014. P. 1701–1708.
5. **He R., Zhang X., Ren S., Sun J.** Delving deep into rectifiers: Surpassing human-level performance on imagenet classification // Proceedings of IEEE International Conference on Computer Vision. 2015. P. 1026–1034.
6. **Sun Y., Wang X., Tang X.** Deeply learned face representations are sparse, selective, and robust // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. 2015. P. 2892–2900.
7. **Goodfellow I., Bengio Y., Courville A.** Deep Learning. Cambridge, MA: MIT Press, 2016.
8. **Guo Y., Liu Y., Oerlemans A., Lao S., Wu S., Lew M. S.** Deep learning for visual understanding: A review // Neurocomputing. 2016. Vol. 187. P. 27–48.
9. **Rawat W., Wang Z.** Deep convolutional neural networks for image classification: A comprehensive review // Neural Computation. 2017.
10. **Schroff F., Kalenichenko D., Philbin J.** Facenet: A unified embedding for face recognition and clustering // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. 2015. P. 815–823.
11. **Sun Y., Liang D., Wang X., Tang X.** Deepid3: Face recognition with very deep neural networks // arXiv 1502.00873. 2015.
12. **Zhou E., Cao Z., Yin Q.** Naive-deep face recognition: Touching the limit of LFW benchmark or not? // arXiv 1501.04690. 2015.
13. **Li B. Y., Mian A., Liu W., Krishna A.** Using kinect for face recognition under varying poses, expressions, illumination and disguise // IEEE Workshop on Applications of Computer Vision (WACV). 2013.
14. **Abbad A., Abbad K., Tairi H.** 3D Face recognition: Multi-scale strategy based on geometric and local descriptors // Computers & Electrical Engineering. 2018. Vol. 70. P. 525–537.
15. **Dorofeev K., Ruchay A., Kober A., Kober V.** 3D face recognition using depth filtering and deep convolutional neural network // Applications of Digital Image Processing. 2019. Vol. XLII. P. 11137.
16. **Ruchay A., Dorofeev K., Kalschikov V.** A switching morphological algorithm for depth map recovery // Analysis of Images, Social Networks and Texts: 8th International Conference. 2019.
17. **Dorofeev K., Ruchay A.** Design of autonomous mobile systems for face recognition based on a DCNN with compression and pruning // Applications of Digital Image Processing. 2020. Vol. XLIII. P. 11510.
18. **Kim D., Hernandez M., Choi J., Medioni G.** Deep 3d face identification // arXiv 1703.10714. 2017.
19. **Gilani S. Z., Mian A.** Learning from millions of 3d scans for large-scale 3d face recognition // arXiv 1711.05942. 2018.

20. **Cai Y., Lei Y., Yang M., You Z., Shan S.** A fast and robust 3d face recognition approach based on deeply learned face representation // *Neurocomputing*. 2019. Vol. 363. P. 375–397.
21. **Chui M., Cheng H., Wang C., Lai S.** High-Accuracy RGB-D Face Recognition via Segmentation-Aware Face Depth Estimation and Mask-Guided Attention Network. Department of Computer Science, National Tsing Hua University, Taiwan, Microsoft AI R&D Center, 2018.
22. **Lee Y.-C., Chen J., Tseng C. W., Lai S.-H.** Accurate and robust face recognition from RGB-D images with a deep learning approach // *Proceedings of the British Machine Vision Conference (BMVC)*. 2016.
23. **Xiong X., Wen X., Huang C.** Improving RGB-D face recognition via transfer learning from a pretrained 2D network // *International Symposium on Benchmarking, Measuring and Optimization*. 2019.
24. **Jiang L., Zhang J., Deng B.** Robust RGB-D face recognition using attribute-aware loss // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2019.
25. **Charles R., Hao S., Kaichun M., Leonidas J.** PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. Stanford : Stanford University, 2017.
26. **Zhang Z., Da F., Yu Y.** Data-free point cloud network for 3d face recognition // *arXiv*, 1911.04731. 2019.
27. **Feng Y., Zhang Z., Zhao X., Ji R., Gao Y.** Group-view convolutional neural networks for 3d shape recognition // *Conference on Computer Vision and Pattern Recognition (CVPR)*. 2018.
28. **Jackson A. S., Bulat A., Argyriou V., Tzimiropoulos G.** Large pose 3D face reconstruction from a single image via direct volumetric CNN regression // *International Conference on Computer Vision (ICCV)*. 2017.
29. **Savran A., Alyuz N., Dibeklioglu H., Celiktutan O., Gokberk B., Sankur B., Akarun L.** Bosphorus database for 3D face analysis // *Workshop on Biometrics and Identity Management*. 2008.
30. **Cao Q., Shen L., Xie W., Parkhi O., Zisserman A.** VGGFace2: A dataset for recognising faces across pose and age // *IEEE Conference on Automatic Face and Gesture Recognition (F&G)*. 2018.
31. **Guo Y., Zhang L., Hu Y., He X., Gao J.** MS-Celeb-1M: A dataset and benchmark for large scale face recognition // *European Conference on Computer Vision (ECCV)*. 2016.

Поступила в редакцию 24.08.2022.

После переработки 24.09.2022.

Сведения об авторах

Дорофеев Константин Андреевич, старший преподаватель кафедры компьютерной безопасности и прикладной алгебры; Челябинский государственный университет, Челябинск, Россия; e-mail: Kostuan1989@mail.ru.

Ручай Алексей Николаевич, кандидат физико-математических наук, доцент, заведующий кафедрой компьютерной безопасности и прикладной алгебры; Челябинский государственный университет, Челябинск, Россия; доцент кафедры защиты информации, Южно-Уральский государственный университет (национальный исследовательский университет), Челябинск, Россия; e-mail: ran@csu.ru.

COMBINED 2D/3D APPROACH FOR IMPROVING THE ACCURACY OF FACE RECOGNITION SYSTEMS USING DEEP LEARNING

K.A. Dorofeev^{1,a}, A.N. Ruchay^{1,2,b}

¹*Chelyabinsk State University, Chelyabinsk, Russia*

²*South Ural State University (National Research University), Chelyabinsk, Russia*

^a*Kostuan1989@mail.ru*, ^b*ran@csu.ru*

Face recognition systems using color images are widely known. However, their main problem is the instability to various lighting conditions, emotional and facial expressions, overlaps and rotation angles. The article proposed a new approach that combines the processing of two-dimensional data on a color image and three-dimensional data on a point cloud or depth map. The main attention was paid to the architecture of neural networks, the quality and accuracy of recognition. Some combined methods for the face recognition using machine learning and deep learning have been proposed. A comparative analysis of the results of experiments in terms of the recognition accuracy on open face databases was carried out. The best combined method was chosen. The main goal is to build a reliable, accurate combined biometric face recognition system that is resistant to complex external factors, such as facial expressions, scale changes, lighting, partial overlap with foreign objects, large rotation angles.

Keywords: *information security, biometrics, face recognition, neural networks, machine learning, deep learning, identification, authentication, multi-biometrics.*

References

1. Callaway S., Cheng J., Contratti A., Fu D., Gelivi H., Wachulec J., Purohit S. Comparative analysis of image processing algorithms for airport security. *IEEE MIT Undergraduate Research Technology Conference (URTC)*, 2020.
2. Bansal M., Sharma D. Facial recognition system for security resolutions in smart city. *International Journal of Advanced Research in Engineering and Technology*, 2020, vol. 11, no. 10, pp. 146–151.
3. Praveen G., Dakala J. Face recognition: Challenges and issues in smart city/environments. *International Conference on Communication Systems and Networks*, 2020, pp. 791–793.
4. Taigman Y., Yang M., Ranzato M., Wolf L. Deepface: Closing the gap to human-level performance in face verification. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1701–1708.
5. He R., Zhang X., Ren S., Sun J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *Proceedings of IEEE International Conference on Computer Vision*, 2015, pp. 1026–1034.
6. Sun Y., Wang X., Tang X. Deeply learned face representations are sparse, selective, and robust. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 2892–2900.
7. Goodfellow I., Bengio Y., Courville A. *Deep Learning*. Cambridge, MA, MIT Press, 2016.
8. Guo Y., Liu Y., Oerlemans A., Lao S., Wu S., Lew M.S. *Deep learning for visual understanding: A review*. *Neurocomputing*, 187, 2016, pp. 27–48.

9. **Rawat W., Wang Z.** Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation*, 2017.
10. **Schroff F., Kalenichenko D., Philbin J.** Facenet: A unified embedding for face recognition and clustering. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 815–823.
11. **Sun Y., Liang D., Wang X., Tang X.** Deepid3: Face recognition with very deep neural networks. *arXiv* 1502.00873, 2015.
12. **Zhou E., Cao Z., Yin Q.** Naive-deep face recognition: Touching the limit of LFW benchmark or not? *arXiv* 1501.04690, 2015.
13. **Li B.Y., Mian A., Liu W., Krishna A.** Using kinect for face recognition under varying poses, expressions, illumination and disguise. *IEEE Workshop on Applications of Computer Vision (WACV)*, 2013.
14. **Abbad A., Abbad K., Tairi H.** 3D face recognition: Multi-scale strategy based on geometric and local descriptors. *Computers & Electrical Engineering*, 2018, vol. 70, pp. 525–537.
15. **Dorofeev K., Ruchay A., Kober A., Kober V.** 3D face recognition using depth filtering and deep convolutional neural network. *Applications of Digital Image Processing*, 2019, vol. XLII, p.11137.
16. **Ruchay A., Dorofeev K., Kalschikov V.** A switching morphological algorithm for depth map recovery. *Analysis of Images, Social Networks and Texts*, 8th International Conference, AIST, 2019.
17. **Dorofeev K., Ruchay A.** Design of autonomous mobile systems for face recognition based on a DCNN with compression and pruning. *Applications of Digital Image Processing*, 2020, vol. XLIII, p. 11510.
18. **Kim D., Hernandez M., Choi J., Medioni G.** Deep 3D face identification. *arXiv* 1703.10714, 2017.
19. **Gilani S.Z., Mian A.** Learning from millions of 3d scans for large-scale 3d face recognition. *arXiv* 1711.05942, 2018.
20. **Cai Y., Lei Y., Yang M., You Z., Shan S.** A fast and robust 3D face recognition approach based on deeply learned face representation. *Neurocomputing*, 2019, vol. 363, pp. 375–397.
21. **Chui M., Cheng H., Wang C., Lai S.** *High-Accuracy RGB-D Face Recognition via Segmentation-Aware Face Depth Estimation and Mask-Guided Attention Network*. Department of Computer Science, National Tsing Hua University, Taiwan, Microsoft AI R&D Center, 2018.
22. **Lee Y.-C., Chen J., Tseng C.W., Lai S.-H.** Accurate and robust face recognition from RGB-D images with a deep learning approach. *Proceedings of the British Machine Vision Conference (BMVC)*, 2016.
23. **Xiong X., Wen X., Huang C.** Improving RGB-D face recognition via transfer learning from a pretrained 2D network. *International Symposium on Benchmarking, Measuring and Optimization*, 2019.
24. **Jiang L., Zhang J., Deng B.** Robust RGB-D face recognition using attribute-aware loss. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.
25. **Charles R., Hao S., Kaichun M., Leonidas J.** *PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation*. Stanford, Stanford University, 2017.
26. **Zhang Z., Da F., Yu Y.** Data-free point cloud network for 3d face recognition. *arXiv* 1911.04731, 2019.
27. **Feng Y., Zhang Z., Zhao X., Ji R., Gao Y.** Group-view convolutional neural networks for 3D shape recognition. *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.
28. **Jackson A.S., Bulat A., Argyriou V., Tzimiropoulos G.** Large pose 3D face reconstruction from a single image via direct volumetric CNN regression. *International Conference on Computer Vision (ICCV)*, 2017.

29. **Savran A., Alyuz N., Dibeklioglu H., Celiktutan O., Gokberk B., Sankur B., Akarun L.** Bosphorus database for 3D face analysis. *Workshop on Biometrics and Identity Management*, 2008.
30. **Cao Q., Shen L., Xie W., Parkhi O., Zisserman A.** VGGFace2: A dataset for recognising faces across pose and age. *IEEE Conference on Automatic Face and Gesture Recognition (FG)*, 2018.
31. **Guo Y., Zhang L., Hu Y., He X., Gao J.** MS-Celeb-1M: A Dataset and Benchmark for Large Scale Face Recognition. *European Conference on Computer Vision (ECCV)*, 2016.

Article received 24.08.2022.

Corrections received 24.09.2022.