

Использование современных алгоритмов машинного обучения для задачи распознавания эмоций¹

Д. А. Астахов, А. В. Катаев

Волгоградский государственный технический университет
400005, Волгоград, пр. Ленина, 28

e-mail: astakhov.d.a@gmail.com, alexander.kataev@gmail.com

Аннотация. В работе сравниваются результаты работы алгоритмом машинного обучения SVM, ANN на наборе данных из нормализованных ключевых точек и дистанций между ними, а также метода CNN на монохромных изображениях. Проводится описание процесса подготовки базы для обучения и экспериментального тестирования используемых методов.

Ключевые слова: система распознавания эмоций, ключевые точки, машинное обучение, распознавание лица, эмоции.

1. Введение

Изменение выражения лица — одно из наиболее важных и естественных средств общения, передачи эмоций и намерений для человека. Автоматизированная классификация эмоционального состояния лица человека является интересной и сложной задачей, которая встречается во многих областях, например человеко-машинное взаимодействие, анимация лица, определение эмоционального состояния. Благодаря широкому кругу приложений, автоматический анализ выражения лица привлек большое внимание исследователей в последние годы.

В частности, в работах под руководством Заболеевой-Зотовой А. В. [1–3] рассматривается возможность построения системы определения эмоциональных реакций человека по мимике, голосу и невербальным движениям. В работах Алфимцева А. Н. и Девяткова В. В. рассматривается задача распознавания динамических жестов человека [4]. В докторской диссертации Дементенко В. В. разрабатывается автоматизированная система предугадывания засыпания водителя по частоте моргания и саккадам [5]. В работе Цайфенг Шан рассматривается алгоритм детектирования эмоций на основе локальных бинарных паттернов [6]. В последнее время для решения классификации эмоций стали часто применяться сверточные нейронные сети, это можно заметить по таким работам, как [7–11].

¹Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 18-07-01308а.

Большинство систем автоматического распознавания выражения лица, найденных в литературе, осуществляют классификации напрямую по основным эмоциям. Существует множество методик кодирования эмоций, среди них наиболее известной является «Система кодирования лицевых движений» (FACS), разработанная Полом Экманом и Уолессом Фризенем [12]. Система FACS является самым распространенным стандартом классификации выражений лица человека и активно используется как в областях психологии, так и в областях компьютерной анимации.

Тем не менее главным ограничением системы FACS является отсутствие учета временной и детальной пространственной информации [13]. В то время как «ярко» выраженные эмоциональные состояния человека могут быть распознаны с высокой точностью, но определение скрытых или промежуточных выражений требует сравнения во времени изменения формы и внешнего вида лица [14].

В связи с этим целью данного исследования является реализация и сравнение наиболее популярных методов машинного обучения для задач классификации: метод опорных векторов (SVM), искусственная нейронная сеть (ANN), сверточная нейронная сеть. Обучение методов ANN и SVM производится на двух наборах данных: нормализованные ключевые точки и дистанции между ключевыми точками. Идея подхода к обучению методов SVM и ANN на различных наборах данных взята из статьи [8]. Следующим исследуемым алгоритмом распознавания эмоций является сверточная нейронная сеть, обучаемая на монохромных изображениях, ее обучение производится на относительной небольшой выборке данных, которая состоит из последовательно изменяющихся изображений с течением времени от слабо до «ярко» проявляемой эмоции.

2. Определение эмоционального состояния человека

Известно, что выражение эмоций может быть весьма разнообразным и различаться как в зависимости от индивидуальных особенностей, так и от ситуативного контекста. Культурный контекст играет важную роль, если речь идет об эмоциях, не относящихся к категории базовых, поскольку принятые в определенной общности способы выражения сложных эмоциональных состояний различаются. Кроме того, проявления, соответствующие «смешанным» эмоциям, трудны для распознавания, поскольку за ними скрывается целый комплекс испытываемых человеком чувств. Еще один фактор, влияющий на эффективность распознавания эмоций, — возможность фальсифицировать эмоциональное поведение. Культурные требования, личные представления о допустимом или любые соображения, вытекающие из восприятия человеком актуальной ситуации (желание что-то скрыть или демонстрация чувства, которого на самом деле нет), — эти и другие факторы оказывают влияние на выражение человеком своих эмоций. Автоматическая идентификация образов

(текст, звук, лицо, персона, предметы и т. д.) с помощью компьютера является одним из наиболее важных направлений развития технологий искусственного интеллекта, позволяющая дать ключ к пониманию особенностей работы человеческого интеллекта. Исследование методов автоматического распознавания эмоций позволяет дать компьютеру возможность оценивать настроение человека, для этого в работе используется алгоритм распознавания эмоций, представленный на рис. 1.

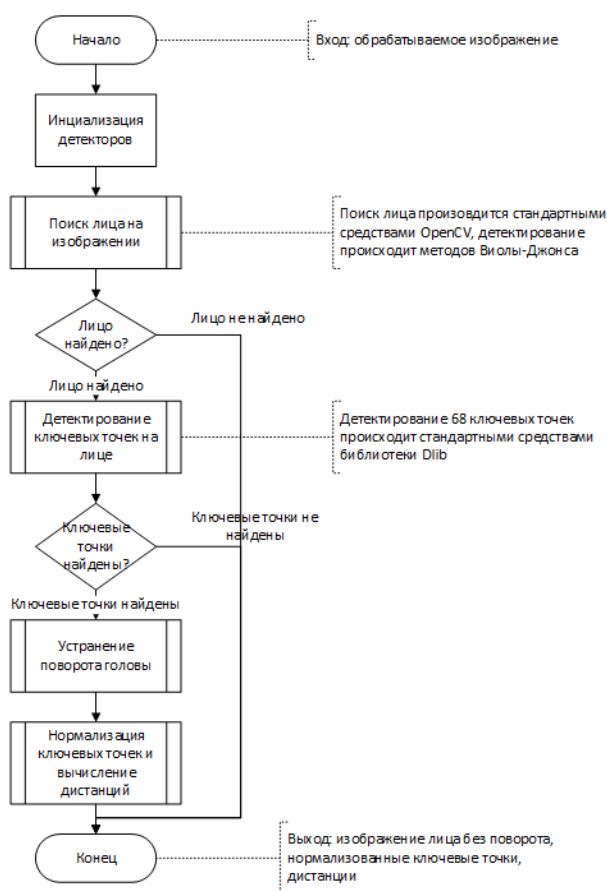


Рисунок 1. Основной алгоритм распознавания эмоций

3. Подготовка выборки для обучения

Для обучения алгоритмов машинного обучения была использована выборка The Extended Cohn-Kanade Database (СК+) [9], состоящая из 11 061 фотографий, с разрешением 640×490 пикселей, в формате *.png. В базе СК+ размечено 623 структуры с эмоциями. Выборка разбита на 8 классов эмоций: 1 — злость, 2 — презрение, 3 — отвращение, 4 — страх, 5 — счастье, 6 — грусть, 7 — удивление, 8 — нор-

мальное. В дальнейшем для краткости классы эмоций будут обозначаться соответствующими номерами.

Исходная база переработана таким образом, чтобы минимизировать ошибки распознавания. Для методов SVM и ANN для каждого класса эмоций и человека было выбрано изображение с максимально «выраженной» эмоцией и для каждого изображения выполнено зеркальное отображение для увеличения объема выборки.

Результирующий объем переработанной базы составил 2016 фотографий. На рис. 2 представлено распределение объема изображений по каждому классу.

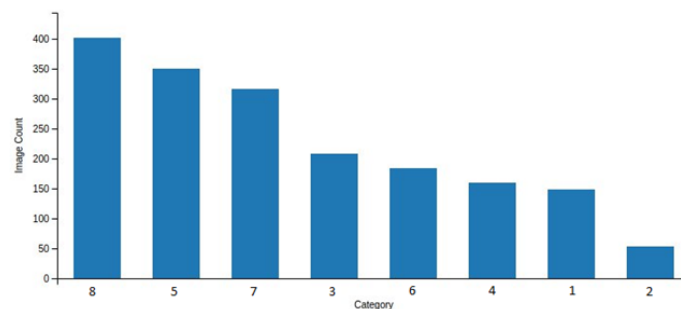


Рисунок 2. Диаграмма распределения объема данных по классам

Неравномерность распределения изображений по классам связана с ограниченностью объема выборки для каждого класса в исходной базе СК+, классы 5, 7, 8 имеют больший объем данных в исходной базе, чем классы 1, 2, 3, 4, 6. Пример изображений, используемых в базе, приведен на рис. 3.



Рисунок 3. Пример изображений из базы, классы расположены слева направо (крайнее левое изображение соответствует 1 классу, крайнее правое 8 классу) (The Extended Cohn-Kanade Database)

Для дальнейшей работы выборка была нормализована. Под нормализацией подразумевается изображение, в котором лицо расположено без поворотов, наклонов и 95% изображения составляет само лицо.

Пример нормализованных изображений для каждого класса эмоций приведен на рис. 4. Размер изображений после нормализации фиксирован и равен 80×80 пикселей.

К нормализованному изображению лица применяются методы детектирования ключевых точек, в результате чего на выходе получают 68 ключевых точек в си-

стеме координат нормализованного изображения. Для нормализации данных все координаты приводятся к диапазону $[0, 1]$.



Рисунок 4. Пример изображений из нормализованной базы, классы расположены слева направо (крайнее левое изображение соответствует 1 классу, крайнее правое 8 классу)

Вычисление нормализованных ключевых точек производилось по формулам

$$x'_i = \frac{x_i}{width}; y'_i = \frac{y_i}{height}. \quad (1)$$

Вычисление дистанций между ключевыми точками производилось по формуле

$$d_i = \sqrt{(x_c - x'_i)^2 + (y_c - y'_i)^2}. \quad (2)$$

Для обучения метода CNN необходимо перевести все изображения из базы в монохромный вид. Пример изображений для обучения метода CNN приведен на рис. 5.



Рисунок 5. Пример монохромных изображений из нормализованной базы, классы расположены слева направо (крайнее левое изображение соответствует 1 классу, крайнее правое 8 классу)

4. Обучение системы

Обучение алгоритмов SVM, ANN производилось при помощи встроенных методом машинного обучения в библиотеке OpenCV.

Для обучения алгоритмов SVM, ANN на нормализованных ключевых точках и дистанция подготовленная база была разбита на две части: тренировочная и тестовая. Под тестирование было выделено 15% от общего объема выборки, что составляет 272 изображения, тренировочная база состоит из 1744 изображений.

4.1. Обучение метода SVM

Для обучения SVM использовался тип `C_SVC`, который позволяет проводить классификацию на 8 классов, используя мультипликатор `C` со значением, равным 0.5.

Тип ядра был выбран полиномиальный с параметрами: $\text{degree} = 0.001$, $\text{coef0} = 0.1$, $\text{gamma} = 0.00015$.

4.2. Обучение метода ANN

Архитектура сети для обучения метода ANN на нормализованных ключевых точках состоит из 4 слоев: входной слой из 44 нейронов, два промежуточных слоя по 88 нейронов, каждый и выходной слой из 8 нейронов. Используется сигмоидальная функция активации, обучение сети производится методом обратного распространения ошибки с параметром scale , равным 0.1. Максимальное количество итераций равно 1000.

4.3. Обучение CNN

Для обучения метода CNN была выбрана библиотека Caffe, так как она обладает высокой скоростью работы, большим количеством обучающих видео, легкостью освоения и возможностью использования в production-решениях.

Для обучения сети использовалась вся выборка СК+, процентное соотношение данных для тестирования осталось аналогично описанному ранее методу, равным 15% от всей выборки.

Стандартными средствами библиотеки было получено среднее изображение по нормализованной базе (см. рис. 6), и из каждого изображения вычиталось среднее, пример изображения после вычитания среднего приведен на рис. 7.

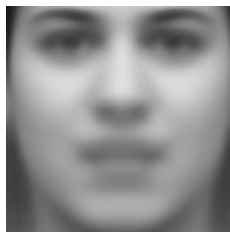
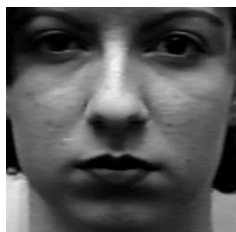
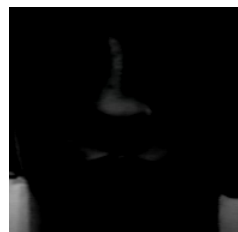


Рисунок 6. Среднее изображение по нормализованной базе



(a)



(б)

Рисунок 7. Пример изображения после вычитания среднего, изображение (a) исходное, изображение (б) получено после вычитания из изображения (a) среднего изображения

Рассмотрим ключевые моменты конфигурации каждого слоя.

Входной слой состоит из 2 элементов (train, test), каждый имеет тип: ImageData, т. е. на вход сеть принимает изображения. Изображения перечисляются в текстовом файле, где первый столбец — путь к изображению, второй столбец — класс эмоции, к которому относится изображения. Аналогичная структура для тестовых данных. Пример описания входного слоя приведен в листинге 1.

Листинг 1. Входной слой

```
layer {
  name: "emotion"
  type: "ImageData"
  top: "data"
  top: "label"
  image_data_param {
    source: "C:\\Caffe\\train.txt"
    batch_size: 4
    shuffle: true
    new_height: 80
    new_width: 80
  }
  transform_param {
    mirror: true
    crop_size: 75
    mean_file: "C:\\Caffe\\Full_Emotion\\mean.binaryproto"
  }
  include {
    phase: TRAIN
  }
}
```

Входные слои также включают атрибут transform_param. В данном атрибуте указываются трансформации, которым надо подвергнуть входные данные. В данном случае был указан mean_file — вычитание «среднего» изображения из входного, параметры new_width, new_height для изменения размера изображения до 80×80 пикселей и crop_size — равный 75 для обрезания изображения до размеров 75×75 пикселей случайным образом.

Основным рабочим блоком сети является последовательность слоев Convolution — BatchNorm — Sigmoid — Pooling. Слой Convolution отвечает за свертку изображения размером XX×XX пикселей с использованием ядра, равного 5, до изображения размером (XX−4)×(XX−4) пикселей. Слой BatchNorm служит для нормализации координат тензора в фиксированное распределение, одного для всех

слоев, чтобы избавиться от дополнительного обучения преобразования данных. Слой Sigmoid является активационным слоем и использует сигмоидальную функцию активации. Слой Pooling отвечает за сокращение изображения размером $XX \times XX$ пикселей с использованием ядра, равного 3, и свертки по среднему, до изображения размером $(XX/3) \times (XX/3)$ пикселей. Сеть содержит три повторяющихся блока, но с другими параметрами. Подбор параметров для этих слоев является эмпирическим. Пример рабочего блока приведен в листинге 2.

Листинг 2. Слой свертки и подвыборки

```
layer {
  name: "conv2"
  type: "Convolution"
  bottom: "Sigmoid1"
  top: "conv2"
  param {
    lr_mult: 1
  }
  convolution_param {
    num_output: 32
    pad: 2
    kernel_size: 5
    stride: 1
    bias_term: false
    weight_filler {
      type: "gaussian"
      std: 0.01
    }
  }
}
```

```
layer {
  name: "bn2"
  type: "BatchNorm"
  bottom: "conv2"
  top: "bn2"
  param {
    lr_mult: 0
  }
  param {
    lr_mult: 0
  }
  param {
    lr_mult: 0
  }
}
```



```
layer {
  name: "Sigmoid2"
  type: "Sigmoid"
  bottom: "bn2"
  top: "Sigmoid2"
}
layer {
  name: "pool2"
  type: "Pooling"
  bottom: "Sigmoid2"
  top: "pool2"
  pooling_param {
    pool: AVE
    kernel_size: 3
    stride: 2
  }
}
```

Полносвязный слой имеет тип InnerProduct. Выходной слой соединяется со слоем функцией потерь (SoftmaxWithLoss) и слоем точности (accuracy). Слой точности срабатывает только в тестовой фазе и показывает процент верно классифицированных изображений в тестовой выборке. Полная структура используемой сети представлена на рис. 8.

Пример выходного слоя приведен в листинге 3.

Листинг 3. Полносвязные и выходные слои

```
layer {
  name: "ip1"
  type: "InnerProduct"
  bottom: "pool3"
  top: "ip1"
  param {
    lr_mult: 1
    decay_mult: 1
  }
  param {
    lr_mult: 1
    decay_mult: 0
  }
  inner_product_param {
    num_output: 8
    weight_filler {
      type: "gaussian"
      std: 0.01
    }
  }
}
```

```
bias_filler {
  type: "constant"
}
}
}
layer {
  name: "accuracy"
  type: "Accuracy"
  bottom: "ip1"
  bottom: "label"
  top: "accuracy"
  include {
    phase: TEST
  }
}
layer {
  name: "loss"
  type: "SoftmaxWithLoss"
  bottom: "ip1"
  bottom: "label"
  top: "loss"
}
```

Параметры обучения в Caffe задаются через конфигурационный файл типа *.prototxt (см. листинг 4). Основными параметрами являются: net — путь к конфигурации архитектуры сверточной нейронной сети, test_interval — количество итераций, между которыми производится тестирование, snapshot — количество итераций, между которыми стоит сохранять состояния сети, solver_mode — режим, указывающий, на каком устройстве будет происходить обучение (GPU или CPU).

Листинг 4. Конфигурация параметров обучения

```
net:
"C:\Caffe\EmotionRecognition\er_full_sigmoid_train_test_bn.prototxt"
  test_iter: 10
  test_interval: 1000
  base_lr: 0.001
  iter_size: 2
  momentum: 0.9
  lr_policy: "step"
  gamma: 1
  stepsize: 5000
  display: 100
  max_iter: 100000
  snapshot: 10000
  snapshot_prefix: "train\er_full_sigmoid_bn_crop"
  solver_mode: GPU
```

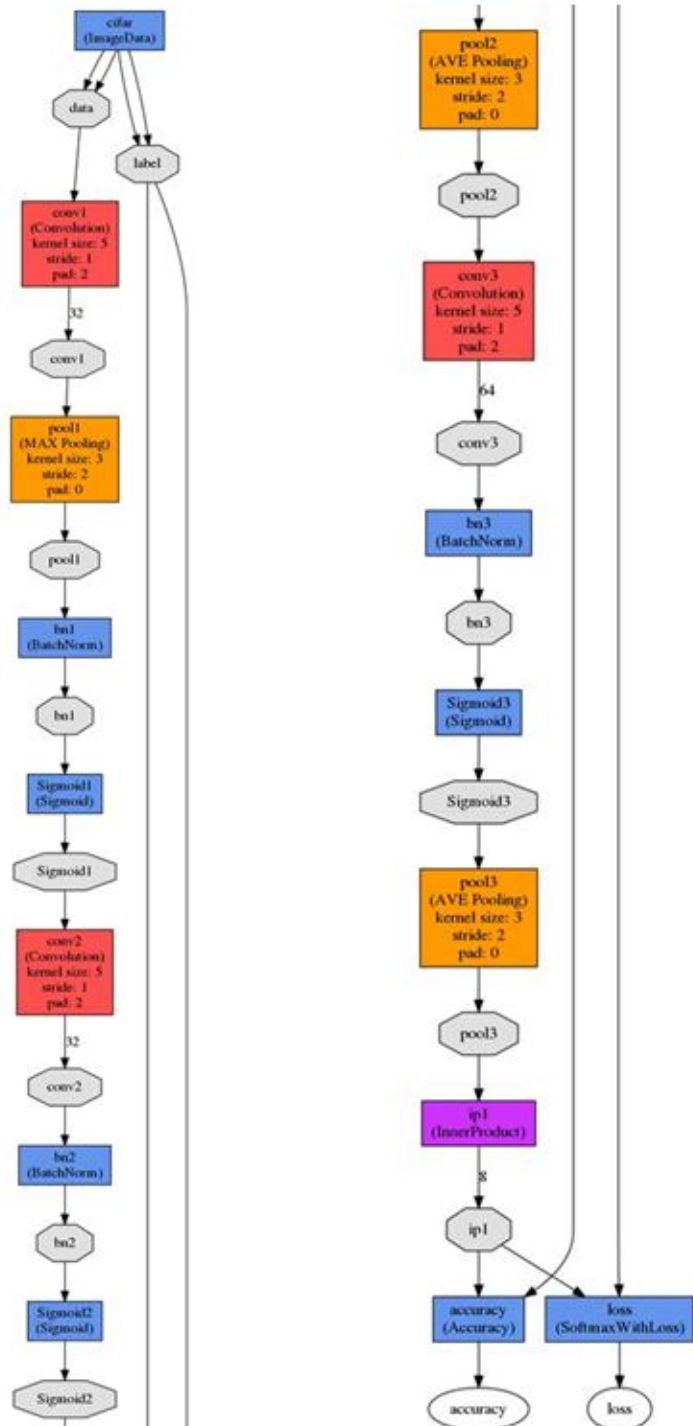


Рисунок 8. Полная структура используемой сверточной нейронной сети

5. Результаты эксперимента

Работа методов машинного обучения SVM и ANN проводилась на двух наборах данных: нормализованные ключевые точки и дистанции между ключевыми точками, а метода CNN — на монохромных нормализованных изображениях. Результаты экспериментов, проведенных для каждого метода машинного обучения с разными наборами данных, приведены в табл. 1–5.

Таблица 1. Результат тестирования метода SVM на нормализованных ключевых точках

Ожидаемая эмоция	Распознанная эмоция							
	1	2	3	4	5	6	7	8
1	30.62	0.00	22.50	0.00	0.00	31.25	0.00	15.62
2	3.70	0.00	3.70	0.00	16.66	11.11	0.00	64.81
3	0.00	0.00	31.77	0.00	17.77	14.77	0.00	35.69
4	0.00	0.00	0.00	37.07	17.53	0.00	12.69	32.69
5	0.00	0.00	3.33	0.00	93.33	0.00	0.00	3.33
6	0.00	0.00	9.37	6.13	9.37	48.88	0.00	26.25
7	0.00	0.00	0.00	4.16	0.00	4.17	91.67	0.00
8	0.00	0.00	0.56	1.11	4.33	1.67	0.00	90.94

Таблица 2. Результат тестирования метода SVM на данных дистанции

Ожидаемая эмоция	Распознанная эмоция							
	1	2	3	4	5	6	7	8
1	28.13	0.00	15.63	0.00	3.13	31.25	0.0	21.88
2	1.85	0.00	12.96	1.85	1.85	1.11	0.00	70.37
3	0.00	0.00	50.00	0.00	50.00	0.00	0.00	100
4	0.00	0.00	0.00	69.23	7.69	19.23	3.85	0.00
5	3.33	0.00	6.67	3.33	76.67	3.33	0.00	6.67
6	3.13	0.00	15.63	6.25	3.13	56.25	0.00	15.63
7	0.00	0.00	0.00	12.50	12.50	4.17	70.83	0.00
8	1.72	0.00	8.33	1.39	2.94	1.33	0.00	86.28

Таблица 3. Результат тестирования метода ANN на нормализованных координатах

Ожидаемая эмоция	Распознанная эмоция							
	1	2	3	4	5	6	7	8
1	37.50	0.00	3.12	0.00	9.37	21.87	28.12	0.00
2	0.00	27.78	0.00	1.85	7.40	33.30	29.62	0.00
3	0.00	0.00	21.30	0.00	0.00	34.80	21.70	22.20
4	0.00	3.84	0.00	54.10	3.84	5.380	23.80	9.04
5	0.00	0.00	0.00	0.00	50.00	16.40	14.50	19.10
6	0.00	0.00	0.00	0.00	0.00	71.87	8.12	20.00
7	0.00	0.00	0.00	3.33	0.00	4.20	77.91	14.36
8	0.00	1.38	0.00	4.17	2.70	0.00	3.61	88.14

Таблица 4. Результат тестирования метода ANN на данных дистанции

Ожидаемая эмоция	Распознанная эмоция							
	1	2	3	4	5	6	7	8
1	41.87	0.00	0.00	0.00	6.25	12.5	18.75	20.62
2	0.00	38.51	0.00	1.85	1.85	14.81	11.11	21.85
3	0.00	0.00	50.00	0.00	0.00	0.00	0.00	50.00
4	0.00	0.00	0.00	58.46	0.00	11.53	16.92	13.07
5	0.00	0.00	0.00	0.00	60.00	0.00	13.33	26.67
6	0.00	0.00	0.00	0.00	3.12	60.62	9.37	26.87
7	0.00	0.00	0.00	0.00	0.00	0.00	87.50	12.50
8	0.00	1.38	0.00	0.00	0.00	4.16	16.66	77.77

Таблица 5. Результат тестирования метода CNN на монохромных изображениях

Ожидаемая эмоция	Распознанная эмоция							
	1	2	3	4	5	6	7	8
1	97.73	0.50	0.00	0.00	0.70	0.00	1.10	0.00
2	3.30	91.01	1.51	1.10	0.00	0.00	0.00	3.10
3	1.00	0.00	71.70	16.83	4.30	0.00	3.00	3.20
4	3.00	0.00	17.70	72.40	6.00	0.00	3.60	3.30
5	3.90	0.00	0.00	0.00	88.41	1.80	0.00	5.91
6	0.70	0.00	0.00	0.00	0.30	98.40	0.60	0.03
7	5.40	0.81	0.41	0.82	3.70	2.10	86.30	0.50
8	0.00	2.10	0.00	0.00	4.10	0.40	0.00	93.40

В табл. 6 представлен сравнительный анализ полученных результатов. На рис. 9 оказано графическое представление результатов тестирования.

Таким образом, проанализировав данные, полученные по каждому тестовому примеру, можно сделать вывод о том, что методы SVM и ANN показали очень близкие результаты по точности классификации эмоций. Но метод ANN показал большую точность распознавания, чем метод SVM на данных дистанции на 5.33%, а на данных ключевых точек на 1%.

Таблица 6. Сводная таблица результатов тестирования

Метод	Данные	Точность распознавания, %
SVM	Ключевые точки	52.50
	Дистанции	54.67
ANN	Ключевые точки	53.50
	Дистанции	60.00
CNN	Монохромные изображения	87.41

Как можно заметить, точность классификации увеличивается, если использовать для обучения данные, основанные на дистанции между ключевыми точками. Для метода SVM точность увеличилась на 2.67%, а для метода ANN точность увеличилась на 6.50%.

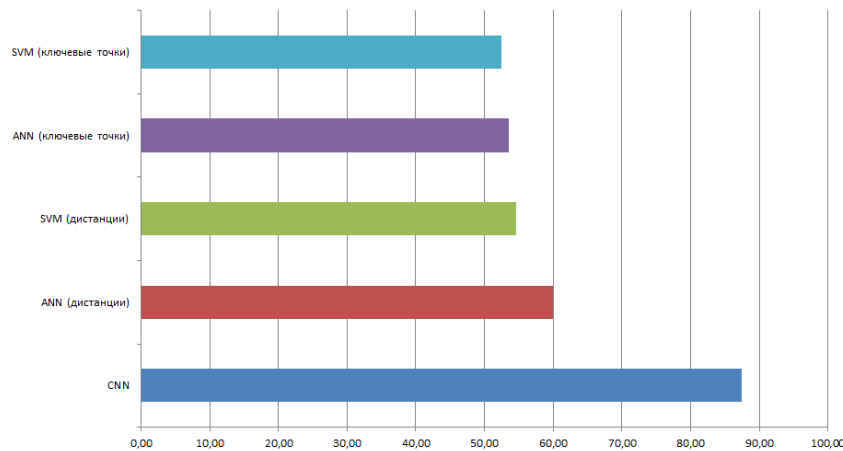


Рисунок 9. Диаграмма результатов тестирования методов машинного обучения на различных наборах данных

Метод CNN, тестируемый на нормализованных монохромных изображениях, показал точность классификации, равную 87.41%, данный показатель можно повысить, если использовать методы агрессивной аугментации над тренировочной базой для ее расширения, стандартизировать данные, также провести обучение на большем (>1000) количестве эпох, использовать dropout слои и скорректировать параметры сети.

6. Заключение

В работе были произведены разработка и сравнение наиболее популярных методов машинного обучения для задачи классификации эмоций и выявлено, что методы ANN и SVM на наборе данных из дистанций между ключевыми точками показывают более точным результат классификации, чем те же методы на наборе данных из нормализованных ключевых точек. Но тем не менее искусственные нейронные сети показывают более точный показатель распознавания, чем метод SVM. Оба метода имеют приемлемый уровень обнаружения. Метод CNN на наборе монохромных изображений показал наилучший результат из всех исследуемых методов.

Дальнейшее развитие исследования заключается в подготовке выборки, основанной на использовании данных с камеры глубины, проектировании архитектуры рекуррентной сверточной нейронной сети, в увеличении базовой выборки эмоций, нормализации ее объема по каждому классу, для более точного детектирования эмоций человека. Реализация метода распознавания эмоций, приведенного в данной статье, может быть использована в интерфейсе интеллектуальной системы для создания условий естественного и интуитивного человеко-машинного взаимодействия.

Литература

- [1] *Заболеева-Зотова А. В.* Задача создания системы автоматизированного распознавания эмоций // Открытые семантические технологии проектирования интеллектуальных систем : материалы Междунар. науч.-техн. конф. OSTIS-2012 (Минск, 16–18 февраля 2012). — Минск : БГУИР, 2012. С. 347–350.
- [2] *Rozaliev V. L., Orlova Yu. A.* Recognizing and Analyzing Emotional expressions in Movements // E-Learning Systems, Environments and Approaches. Theory and Implementation / Eds: P. Isaias et al. — Springer, 2015. P. 117–131.
- [3] *Заболеева-Зотова А. В.* Развитие системы автоматизированного определения эмоций и возможные сферы применения // *Открытое образование*. 2011. № 2. С. 59–62.
- [4] *Девятков В. В., Алфимцев А. Н.* Распознавание манипулятивных жестов // *Вестник МГТУ им. Н. Э. Баумана. Серия: Приборостроение*. 2007. Т. 68. Вып. 3. С. 56–75.
- [5] *Деметтиенко В. В.* Физические принципы построения систем безопасного мониторинга состояния человека-оператора : автореф. дисс. ... д-ра техн. наук. — М., 2010.
- [6] *Shan C., Gong Sh., McOwan Peter W.* Facial expression recognition based on Local Binary Patterns: A comprehensive study // *Image and Vision Computing*, 2009. No. 27. P. 803–816.
- [7] *Enikilopov S. N., Kusnetsova Yu. M.* The Task of Recognition of Violent Situations Using Automatic Systems and Methods of Artificial Intelligence // *Journal of Psychology and Law*. 2011. Vol. 2. P. 1–16.
- [8] *Yi J., Mao X., Chen L., Xue Y., Compare A.* Facial expression recognition considering individual differences in facial structure and texture // *IET Comp. Vision*. 2014. Vol. 8. P. 429–440.
- [9] *Королькова О. А.* Перцептивное пространство и предикторы различения эмоциональных экспрессий лица // *Российский журнал когнитивной науки*. 2014. № 1 (4). С. 82–97.
- [10] *Leibin B. M.* Expression of a person as an object of psychoanalytic research // The person's face as a means of communication. Interdisciplinary approach. 2012. P. 57.
- [11] *LeCun Y., Bengio Y.* Convolutional Networks for Images, Speech, and Time Series. The Handbook of Brain Theory and Neural Networks. — MIT Press, 2002. P. 276–279.
- [12] *Ekman P., Friesen W.* Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press. — Palo Alto, 1978. P. 56–75.
- [13] *Perakis P., Passalis G., Theoharis Th., Kakadiaris I. A.* 3D Facial Landmark Detection & Face Registration. — University of Athens, Greece, 2010.
- [14] *Эчагарай-Патрон Б. А., Кобер В. И., Карнаухов В. Н., Кузнецов В. В.* Метод распознавания лиц с использованием трехмерных поверхностей // *Информационные процессы*. 2016. Т. 16. № 2. С. 170–176.
- [15] *Lucey P.* The Extended Cohn-Kanade Dataset (CK+) : A complete dataset for action unit and emotion-specified expression // Proceedings of IEEE CVPR Workshop on Biometrics, 13–18 Jun 2010, San Francisco. — San Francisco : IEEE Computer Society, 2010. P. 94–101.

Авторы:

Александр Вадимович Катаев — кандидат технических наук, доцент кафедры «Системы автоматизированного проектирования и поискового конструирования», Волгоградский государственный технический университет

Дмитрий Александрович Астахов — аспирант кафедры «Системы автоматизированного проектирования и поискового конструирования», Волгоградский государственный технический университет

Use of Modern Machine Training Algorithms for the Task of Recognition of Emotions

D. A. Astakhov, A. V. Kataev

Volgograd state technical university, 28, Lenina avenue, Volgograd, 400005

e-mail: astakhov.d.a@gmail.com, alexander.kataev@gmail.com

Abstract. The work compares the results of the work of the machine learning algorithm SVM, ANN on a set of data from the normalized key points and the distances between them, as well as the CNN method on monochrome images. The process of preparing the base for training and experimental testing of the methods used is described.

Keywords: the system of recognition of emotions, key points, machine learning, face recognition, emotions.

References

- [1] Zableyeva-Zotova A. V. (2012) Zadacha sozdaniya sistemy avtomatizirovannogo raspoznavaniya emotsiy. In Proc. Otkrytyye semanticheskiye tekhnologii proyektirovaniya intellektual'nykh system. Materialy Mezhdunar. nauch.-tekhn. konf. OSTIS-2012. Minsk, 33. 347–350. [In Rus]
- [2] Rozaliev V. L., Orlova Yu. A. (2015) Recognizing and Analyzing Emotional expressions in Movements. In book: E-Learning Systems, Environments and Approaches. Theory and Implementation. ed. by P. Isaias et al. Springer, pp. 117–131.
- [3] Zableyeva-Zotova A. V. (2011) *Otkrytoye obrazovaniye*. 2:59–62. [In Rus]
- [4] Devyatkov V. V., Alfimtsev A. N. (2007) *Vestnik MGTU im. N. E. Baumana. Seriya: Priborostroyeniye*. 3:56–75. [In Rus]
- [5] Dementiyenko V. V. Fizicheskiye printsipy postroyeniya sistem bezopasnogo monitoringa sostoyaniya cheloveka-operatora. Tesis. M., 2010. [In Rus]
- [6] Shan C., Gong Sh., McOwan Peter W. (2009) *Image and Vision Computing*, 27:803–816.
- [7] Enikilopov S. N., Kusnetsova Yu. M. (2011) *Journal of Psychology and Law*. 2:1–16.
- [8] Yi J., Mao X., Chen L., Xue Y., Compere A. (2014) *IET Computer Vision*. 8:429–440.
- [9] Korol'kova O. A. (2014) *Rossiyskiy zhurnal kognitivnoy nauki*. 1(4):82–97. [In Rus]
- [10] Leibin B. M. (2012) Expression of a person as an object of psychoanalytic research. In book: The person's face as a means of communication. Interdisciplinary approach. P. 57.
- [11] LeCun Y., Bengio Y. (2002) Convolutional Networks for Images, Speech, and Time Series. The Handbook of Brain Theory and Neural Networks. USA, MIT Press. P. 276–279.
- [12] Ekman P., Friesen W. (1978) Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press. Palo Alto. P. 56–75.
- [13] Perakis P., Passalis G., Theoharis Th., Kakadiaris I. A. (2010) 3D Facial Landmark Detection & Face Registration. University of Athens, Greece.
- [14] Echeagaray-Patron B. A., Kober V. I., Karnaukhov V. N., Kuznetsov V. V. (2016) *Informatsionnyye protsessy*. 16(2):170–176. [In Rus]
- [15] Lucey P. (2010) The Extended Cohn-Kanade Dataset (CK+) : A complete dataset for action unit and emotion-specified expression. Proceedings of IEEE CVPR Workshop on Biometrics, 13–18 Jun 2010, San Francisco, pp. 94–101.