

УДК 004.8

Г.А. Кучук¹, Б.Г. Саатсазов²¹ *Национальный аэрокосмический университет имени Н.Е. Жуковского «ХАИ», Харьков*² *Национальный технический университет «ХПИ», Харьков*

РАСПОЗНАВАНИЕ ЧЕЛОВЕЧЕСКИХ ЭМОЦИЙ С ИСПОЛЬЗОВАНИЕМ НЕЙРОСЕТЕВЫХ ТЕХНОЛОГИЙ

*Проведен анализ существующих методов распознавания изображений и описано распознавание мимики человека как математическая задача. Большое внимание уделяется нейросетевым методам распознавания мимики, особенно многослойному перцептрону и глубинной сверточной нейронной сети (ГСНС). Разработан программный продукт, реализующий алгоритм распознавания человеческой мимики с использованием архитектуры ГСНС. Продукт создан на языке Python с использованием современных библиотек компьютерного зрения *dlib*, *open cv*, и библиотеки машинного обучения *tensor flow* от Google.*

Ключевые слова: *нейронная сеть, мимика лица человека, компьютерное зрение, Python, эмоции.*

1. Актуальность

На сегодняшний день еще не существует программного продукта, который мог бы идентифицировать мимику и эмоции человека со 100% точностью. Решение задачи распознавания несет в себе необходимость в обработке больших массивов данных. Целесообразно использовать для этого нейросетевые алгоритмы, так как они наиболее быстро и эффективно справляются с обработкой больших массивов данных. Такие алгоритмы могут выполнять широкий спектр задач коммерческого и некоммерческого характера. Примером применения таких алгоритмов может служить мониторинг выражений лица для проверки реакции человека на различные объявления, рекламные анонсы кинофильмов, книг, телешоу.

2. Анализ литературы

Для разработки системы, которая будет способна распознавать эмоции по выражению лица, были проведены предварительные исследования того как люди проявляют эмоции и была рассмотрена теория автоматической классификации изображений.

2.1. Человеческие эмоции

Ключевым аспектом в человеческом взаимодействии является универсальность выражения эмоций и языка человеческого тела. В 19 веке Чарльз Дарвин опубликовал общепроявляемые выражения лиц, которые играют важную роль в невербальной коммуникации.

В 1971 году Экман и Фризен заявил, что выражения лиц людей по всему миру универсальны и связаны с определенными эмоциями.

Люди развивают похожие мышечные движения, принадлежащие к определенному психическому состоянию, несмотря на их место рождения, расы, образование и т.д. Следовательно, если правильно смоделировать, эта универсальность может быть

очень удобной особенностью в человеко-машинных взаимодействиях: хорошо обученная система может понимать эмоции, независимо от того, кем является субъект.

Следует иметь в виду, что выражение лица не обязательно непосредственно переводится в эмоции, и наоборот. Выражение лица дополнительно является функцией, например, психического состояния, в то время как эмоции выражаются также языком тела и голосом. Таким образом, новые системы распознавания эмоций должны также включать эти последние два параметра. Однако это выходит за рамки этих исследований и останется рекомендацией для будущих исследований.

2.2. Методы классификации изображений

Рост доступной вычислительной мощности на компьютерах-компиляторах в начале XXI века дал толчок развитию алгоритмов, используемых для интерпретации изображений. В классической классификации изображений можно указать две стартовые точки. С одной стороны, предварительно запрограммированные экстракторы функций могут использоваться для аналитического разбиения нескольких элементов изображения, чтобы классифицировать отображаемый объект. Непосредственно против этого подхода, самообучающиеся нейронные сети предоставляют форму техники «черного ящика». В последней концепции сама система разрабатывает правила классификации объектов путем обучения по помеченным образцам данных.

Широкий обзор экстрактов аналитических признаков и подходов нейронной сети для распознавания лицевой экспрессии дается Фазелем и Луэттином. Можно сделать вывод, что к моменту написания, в начале двадцатого века, оба подхода работают одинаково хорошо. Однако, учитывая текущую доступность данных обучения и вычислительной мощности, ожидается, что эффективность моделей

на основе нейронной сети может быть значительно улучшена к настоящему времени. Некоторые недавние достижения будут перечислены ниже.

I. Прорывная публикация об автоматической классификации изображений в целом дается Крижевским и Хинтоном. Эта работа показывает глубинную нейронную сеть, которая напоминает функциональность зрительной коры человека. Используя саморазвитую коллекцию из 60000 изображений более 10 классов, называемую набором данных CIFAR-10, получена модель для категоризации объектов из изображений. Другим важным результатом исследования является визуализация изображений в сети, так что можно оценить, как модель разбивает фотографии.

II. В другой работе, в которой используется набор данных CIFAR-10, разработана очень широкая и глубокая сетевая архитектура в сочетании с поддержкой GPU для сокращения времени обучения. На популярных наборах данных, таких как MNIST, рукописные данные, китайские иероглифы, а также эмоции CIFAR-10, достигается почти человеческая производительность. Чрезвычайно низкие частоты ошибок значительно превзошли предыдущие современные результаты. Однако следует отметить, что сеть, используемая для набора данных CIFAR-10, состоит из 4 сверточных слоев с 300 картами каждый, 3 максимальных слоев объединения и 3 полностью связанных выходных уровней. В результате, хотя использовался графический процессор, время обучения составляло несколько дней.

III. В 2010 году введение ежегодного конкурса "Imagenet" способствовало проведению исследований по классификации изображений, а созданный в рамках этих исследований набор данных часто используется в публикациях с тех пор. В более поздней работе Крижевского, сеть с 5 сверточными, 3 максимальными слоями и 3 полностью связанными слоями обучена с 1,2 миллионами изображений с высоким разрешением на конкурсе ImageNet SVRC-2010. После внедрения технологий для сокращения переобучения, результаты являются многообещающими по сравнению с предыдущими моделями состояния. Кроме того, эксперименты проводятся с уменьшением размера сети, заявляя, что количество слоев может быть значительно уменьшено, а производительность падает лишь незначительно.

IV. Что касается распознавания лиц, это представляет собой глубокую сеть доверия (ГСД), специально предназначенную для использования с японскими женскими выражениями лиц (JAFPE) и расширенными базами данных Cohn-Kanade (CK+). Наибольшей особенностью сети является иерархическая концепция синтаксического анализа лица, то есть изображение передается через сеть несколько раз, прежде чем обнаруживать лицо, после этого

глаза, нос и рот, и наследует принадлежащую им эмоцию. Результаты сопоставимы с точностью, полученной другими методами в одной базе данных, такой как метод опорных векторов (SVM) и квантование вектора обучения (LVQ).

V. В другой работе по базе данных Канаде используется фильтрация Габора для обработки изображений и метод опорных векторов (SVM) для классификации. Фильтр Габора особенно подходит для распознавания образов в изображениях и, как утверждается, имитирует функцию визуальной системы человека. Точность распознавания эмоций высока, варьируя от 88% (гнев) до 100% (удивление). Однако большой недостаток подхода заключается в том, что требуется очень точная предварительная обработка данных, так что каждое изображение должно соответствовать строгому формату, прежде чем подавать его в кластер.

VI. Одно из последних исследований по распознаванию эмоций описывает нейронную сеть, способную распознавать расу, возраст, пол и эмоции из изображению лиц. Набор данных, используемый для последней категории, происходит от вызова распознавания выражения лица (FERC-2013). Четко организованная глубокая сеть, состоящая из 3 свернутых слоев, 1 полностью связанного слоя и некоторых небольших слоев между ними, получила среднее значение 67% от класса эмоций, что соответствует предыдущему состоянию дел Публикации по одному и тому же набору данных. Кроме того, в этом тезисе содержится ценный анализ эффективности рекламы, объединения и отсева.

Подчеркнутая другой литературой, наиболее перспективной концепцией анализа выражения лица является использование глубинных сверточных нейронных сетей. Однако сеть из пункта 2 считается слишком большой для нашего ограниченного количества доступных ресурсов обработки. Исходная сеть из пункта 3 также велика, но, как утверждается, в равной степени подходят меньшие версии. Кроме того, из-за их несколько аналитических и нетрадиционных подходов мы не будем оценивать методы из пункта 4 и 5. Следовательно, в следующем разделе три глубокие архитектуры в целом будут подвергнуты задаче классификации эмоций. Эти архитектуры получены, но не обязательно равны сетям, описанным в пунктах 1,3 и 6.

3. Подготовка к началу эксперимента

Для оценки трех подходов, упомянутых ранее на их способности распознавания эмоций, было разработано три сети на основе понятий из [7], [9] и [20]. В этом разделе описываются данные, используемые для обучения и тестирования, объясняются подробности каждой сети, и оцениваются результаты, полученные с помощью всех трех моделей.

3.1. Наборы данных

Нейронные сети и глубинные сети, в частности, известны их потребностью в больших количествах обучающих данных. Кроме того, выбор изображений, используемых для обучения является ключевым для большей части производительности окончательной модели. Это влечет за собой необходимость в большом и высококачественном наборе данных. Для рас-

познавания эмоций, несколько наборов данных доступны для исследований, колеблясь от нескольких сот фотографий высокого разрешения до десятков тысяч мелких изображений. Будут рассмотрены 3 набора данных: Исследование распознавания человеческих лиц (FERC-2013) [8], расширенная база Cohn-Kanade (CK+) [12] и база данных лиц Radboud (RaFD) [11] показаны на рис. 1.

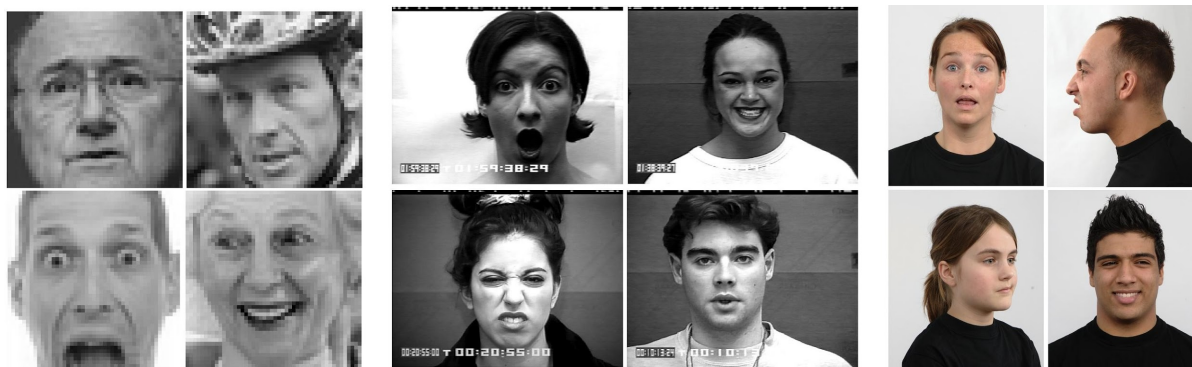


Рис. 1: Образцы из наборов данных FERC-2013 (слева), CK+ (в центре) и RaFD (справа)

Наборы данных отличаются, в основном, количеством, качеством и “чистотой” изображений. Набор данных FERC-2013, к примеру, содержит 32000 изображений низкого разрешения, в то время как RaFD содержит 8000 изображений высокого разрешения. Кроме того, стоит отметить, что выражения лиц в CK+ и RaFD являются “чистыми”, в то время как в наборе FERC-2013 эмоции на изображениях являются “грязными”, т.е. зашумленными. Это делает снимки из набора FERC-2013 более трудно дешифрируемыми, но, учитывая, большой размер набора данных, можно сказать, что этот набор данных является более выгодным для надежности модели.

Можно утверждать, что после того, как было произведено обучение на наборе данных FERC-2013, кадры из «чистых» наборов данных могут быть легко классифицированы, но не наоборот. Следовательно, для трех рассматриваемых сетей, обучение будет сделано с использованием 9000 образцов из набора данных FER-2013 (рис. 2) с 1000 новых образцов для проверки. В дальнейшем тестирование будет осуществляться с помощью 1000 изображений из набора RaFD, чтобы получить оценку превосходства над “чистыми” данными высокого разрешения. Этот набор имеет равномерное распределение по всем эмоциям.

3.2 Нейронные сети

Сети запрограммированы с использованием библиотеки TF Learn из набора Tensor Flow, работает на Python. Эта среда снижает сложность кода, так как должны быть созданы только нейронные слои, вместо того, чтобы создавать каждый нейрон отдельно. Программа также обратную связь в режиме

реального времени о ходе обучения и точности, и позволяет легко сохранять и повторно использовать модель после обучения. Более подробную информацию о этой структуре можно найти в работе [16].

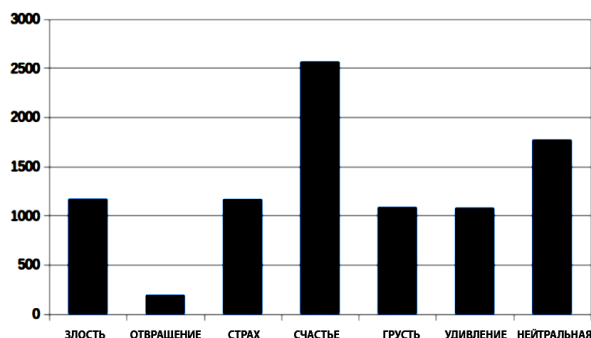


Рис. 2: Количество изображений различных эмоций в выборке

(А) Первая исследуемая сеть основана на ранее описанном исследовании Крижевского и Хинтона [9]. Это самая маленькая сеть из трех, что означает, что она имеет самые низкие вычислительные требования. Так как одно из будущих применений может быть в формате распознавания в режиме реального времени во встроенной системе, быстро работающие алгоритмы будут более предпочтительны. Сеть состоит из трех сверточных слоев и двух полностью соединенных слоев, в сочетании с максимально объединенными слоями для уменьшения размера изображений и отсеивания слоев, чтобы уменьшить вероятность переобучения. Гиперпараметры выбраны таким образом, что количество вычислений в каждом сверточном слое оставалось примерно таким же. Это гарантирует, что информация сохраняется во всей

сети. Обучение проводится с использованием различного числа сверточных фильтров для оценки их эффективности на производительность.

(В) В 2012 году сверточная сеть AlexNet была разработана для классификации изображений в более чем 1000 различных классов, используя 1,2 миллиона снимков из Image Net набора образцов. В связи с тем, что в данном исследовании модель должна различать только семь эмоций, и из-за наших ограниченных вычислительных ресурсов, размер исходной сети считаются слишком большими.

Следовательно, вместо 5 сверточных слоев мы применили 3, и в последующем 3 полностью подключенных слоя, число точек каждого из которых было уменьшено в 4096 до 1024. В то время как первоначальная сеть была разделена для параллельного обучения. Было замечено, что не было необходимости для меньшей версии. Сеть также использует локальную нормализацию для ускорения обучения и отсеивания слоев для уменьшения вероятности переобучения.

(С) Последний эксперимент выполняется на основе работы Гуди [7]. Поскольку это исследование также направлено на распознавание 7 эмоции с помощью набора данных FEREC-2013, архитектура должна быть хорошей отправной точкой для нашего исследования.

Первоначальная сеть начинается с входного слоя размером 48 на 48 нейронов, что соответствует размеру входных данных. Этот слой представлен одним сверточным слоем, слоем локальной нормализации контраста, и полносвязным слоем соответственно. Сеть завершает свою работу еще с двумя сверточными слоями и одним полносвязным слоем, соединенным с выходным слоем. Отсеивание было применено к полносвязному слою.

Для нашего исследования, второй полносвязный слой наносят, чтобы уменьшить число параметров. Это снижает вычислительную интенсивность сети, в то время как снижение производительности, как утверждают, всего 1-2%. Кроме того, скорость обучения регулируется. Вместо того, чтобы линейно уменьшать скорость обучения, как это сделано у Гуди [7], мы полагаем, что скорость обучения, предполагающее использование функции активации будет сходиться быстрее, так как функция активации увеличивает скорость обучения, когда градиент продолжает идти в в том же направлении.

3.3 Оценка результатов

Все сети обучаются 60 эпох с данными, указанными в разделе 3.1. На рис. 3 и в табл. 1 показаны различные детали тренировочного процесса и конечная модель. Для сети А, финальная точность по данным проверки составляет около 63%. Уже через 10 эпох, точность вырастает выше 60%, что указывает на быструю способность к обучению. Кроме

того, примечательно что корректировка размерности фильтра не имеет большого влияния на точность.

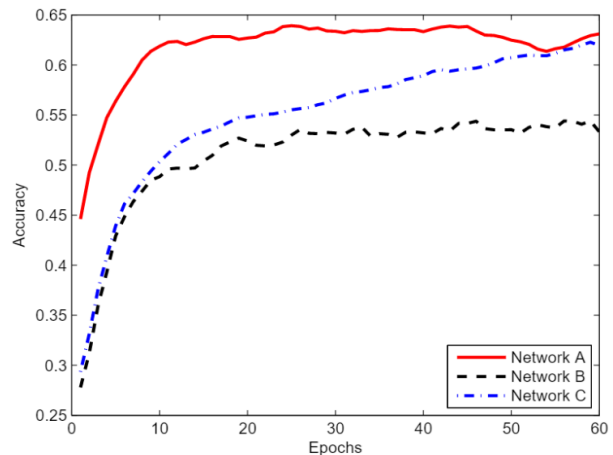


Рис. 3. Точность в проверяемом наборе данных во время обучения

Таблица 1
Детали процесса обучения сетей

Сеть	Точность		Размер
	Проверка	RaFD	
А	63%	50%	Маленький
В	53%	46%	Большой
С	63%	60%	Средний

Удивительно, что вторая, самая большая сеть, обучалась так же быстро, но при этом с точностью около 54%. По-видимому, уменьшение размера сети ломает многообещающую производительность исходной сети больше, чем ожидалось. Вместе с гораздо более высокой вычислительной интенсивностью и, следовательно, более медленной производительностью, эта модель не является достойным конкурентом двум другим архитектурам. Сеть С показывает несколько более медленную кривую обучения, но конечная точность в наборе валидации аналогична сети А. Требования к обработке находятся в зависимости от требований других сетей, поэтому на основе этого факта сеть А, по-видимому, является самой перспективной в подходе к нашей задаче распознавания эмоций.

Однако производительность сети С на дополнительном наборе тестов RaFD значительно лучше (60%), чем в сети А (50%). Это указывает на лучшие обобщающие возможности, что очень важно для будущих приложений. Следовательно, в следующем разделе статьи модель из сети С будет дополнительно исследована и протестирована.

4. Окончательная модель

Было выявлено, что последняя описанная сеть из подраздела 3.2 имеет наиболее перспективные показатели для практического применения. Обзор ее архитектуры показан на рис. 4.

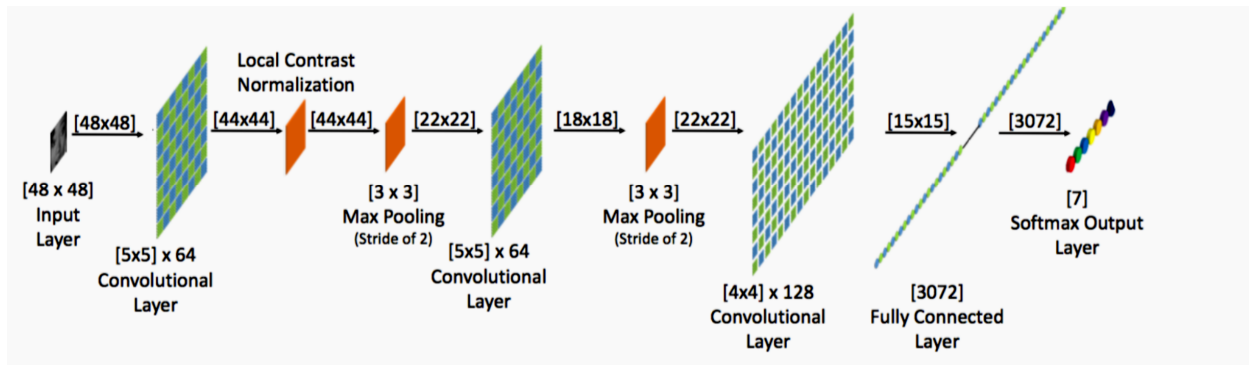


Рис. 4: Обзор сетевой архитектуры конечной модели

В попытке улучшить итоговую модель еще больше, сеть будет обучаться на более широком множестве, чем описанное ранее. Вместо 9000 снимков, обучение будет сделано с 20000 фотографий из набора данных FERC-2013. Соотношение эмоций, присутствующих в этом наборе приведено на рис. 5. Вновь составленный набор данных для обучения (2000 изображений) и тестовые наборы (1000 изображений). Новый тестовый набор состоит из данных FERC-и RaFD.

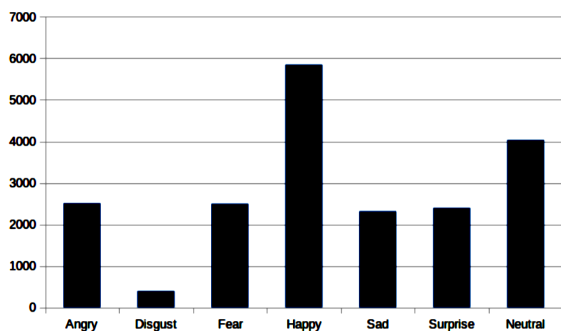


Рис. 5. Количество изображений каждой эмоции в финальном наборе данных для обучения

4.1. Результаты теста

Уровень точности конечной модели приведен в табл. 2. Во время всех проверок и испытаний точность была выше, чем во время предыдущих запусков. Это подтверждает, что большая размерность обучающей выборки и более длительное обучение могут улучшить производительность сети.

Таблица 2

Сеть	Точность сетей		RaFD
	FERC-2013		
	Проверка	Тест	
А	63%		50%
В	53%		46%
С	63%		60%
Итоговая	66%	63%	71%

Учитывая, что уровень развития сетей из предыдущих исследований получен на около 67% обучающей выборки, и имея в виду наши ограниченные, результаты являются достаточно хорошими.

Примечательной является точность на тестовом наборе RaFD, который содержит полностью отличные фотографии нежели в обучающей выборке. Это иллюстрирует мощные генерализационные способности конечной модели.

Чтобы увидеть, как эта модель работает для каждой, отдельно взятой эмоции, создана таблица, (рис. 6). Очень высокие показатели распознавания получены на изображениях с выражением «счастливый» (90%), «нейтральный» (80%), и «удивлен» (77%). Это на самом деле наиболее различимые выражения лиц. «Грустно», «страшно» и «гнев» часто были неправильно классифицированы и приняты за «нейтральное» выражение лица. Видимо, эти эмоции очень похожи друг на друга. Самая низкая производительность была получена на «грустный» (28%) и «страх» (37%). Наконец, следует отметить, что даже несмотря на то процент данных с выражением лица «отвращение» в обучающем наборе низка, уровень классификации является достаточно высоким.

РЕАЛЬНАЯ ЭМОЦИЯ	РАСПОЗНАННАЯ ЭМОЦИЯ						
	ЗЛОСТЬ	ОТВРАЩЕНИЕ	СТРАХ	СЧАСТЬЕ	ГРУСТЬ	УДИВЛЕНИЕ	НЕЙТРАЛЬНАЯ
НЕЙТРАЛЬНАЯ	0.04	0.01	0.03	0.07	0.04	0.02	0.80
УДИВЛЕНИЕ	0.03	0.00	0.07	0.06	0.02	0.77	0.06
ГРУСТЬ	0.12	0.03	0.10	0.08	0.28	0.00	0.39
СЧАСТЬЕ	0.01	0.00	0.00	0.90	0.00	0.02	0.07
СТРАХ	0.14	0.04	0.37	0.05	0.07	0.11	0.22
ОТВРАЩЕНИЕ	0.14	0.62	0.05	0.11	0.00	0.00	0.07
ЗЛОСТЬ	0.50	0.06	0.09	0.05	0.07	0.03	0.21

Рис. 6: Матрица производительности конечной модели (вертикально – вход, в горизонтальном – направлении на выходе)

Выводы

С использованием современных технологий машинного обучения и компьютерного зрения был разработан алгоритм для распознавания мимики лица человека. Было рассмотрено несколько вариантов распознавания мимики, а именно два, первый из которых оказался неудачным, что было обнаружено уже после создания и обучения сети.

Начальний алгоритм включав в себе рішення задачі класифікації з допомогою многослойного перцептрона із ста тридцяти шести входними нейронами, які приймали шістьдесят вісім точок антропометричної моделі обличчя. Експеримент виявився невдачним через те, що нейронна мережа не змогла знайти закономірностей.

На цьому етапі алгоритм був змінений на використання глибокої сверточної нейронної мережі, яка зараз вважається однією з найкращих мереж для розпізнавання та класифікації графічних зображень. Слабким місцем ГСНС може бути висока чутливість до різних перетворень зображень, наприклад афінних, тому було вирішено стандартизувати входні зображення афінного перетворення.

Таким чином, реалізацією алгоритму є програма на мові програмування Python з використанням сучасних бібліотек комп'ютерного зору `dlib`, `open cv`, та бібліотеки машинного навчання `tensor flow` від Google. Нейронна мережа навчена з використанням маркованого набору даних розміром 20000 зображень. Мережа здатна розпізнавати 7 базових мимічних проявів, серед яких злість, радість, печаль, здивування, нейтральність, страх та огоршення.

Точність у розпізнаванні становить в середньому близько 80%. Розпізнавання страху має високу ймовірність помилки, на відміну від розпізнавання посмішки, яку можна розпізнати майже з 99%-ною точністю.

Список літератури

1. T. Ahsan, T. Jabid and U.-P. Chong. Facial expression recognition using local transitional pattern on gaborfiltered facial images. *IETE Technical Review*, 30(1):47{52, 2013.
2. D. Ciresan, U. Meier and J. Schmidhuber. Multicolumn deep neural networks for image classification. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 3642{3649. IEEE, 2012.
3. C. R. Darwin. *The expression of the emotions in man and animals*. John Murray, London, 1872.
4. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei. *Image net: A large-scale hierarchical image database*. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conf. on*, pages 248{255. IEEE, 2009.
5. P. Ekman and W. V. Friesen. Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2):124, 1971.
6. B. Fasel and J. Luetthi. Automatic facial expression analysis: a survey. *Pattern recognition*, 36(1):259{275, 2003.
7. A. Gudi. *Recognizing semantic features in faces using deep learning*. arXiv preprint arXiv:1512.00743, 2015.
8. Kaggle. *Challenges in representation learning: Facial expression recognition challenge*, 2013.
9. A. Krizhevsky and G. Hinton. *Learning multiple layers of features from tiny images*, 2009.
10. A. Krizhevsky, I. Sutskever and G. E. Hinton. Image net classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097{1105, 2012.
11. O. Langner, R. Dotsch, G. Bijlstra, D. H. Wigboldus, S. T. Hawk and A. van Knippenberg. Presentation and validation of the radboud faces database. *Cognition and emotion*, 24(8):1377{1388, 2010.
12. P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews. The extended cohnkanade dataset (ck+): A complete dataset for action unit and emotion specified expression. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pages 94{101. IEEE, 2010.
13. Y. Lv, Z. Feng and C. Xu. Facial expression recognition via deep learning. In *Smart Computing (SMARTCOMP), 2014 International Conference on*, pages 303{308. IEEE, 2014.
14. J. Nicholson, K. Takahashi and R. Nakatsu. Emotion recognition in speech using neural networks. *Neural computing & applications*, 9(4): 290{296, 2000.
15. Open Source Computer Vision. *Face detection using haar cascades*. URL http://docs.opencv.org/master/d7/d8b/tutorial_py_face_detection.html.
16. TFLearn. *Tearn: Deep learning library featuring a higher-level API for tensorflow*. URL <http://tflearn.org/>.

Надійшла до редакції 11.05.2017

Рецензент: д-р техн. наук, проф. І.В. Рубан, Харківський національний університет радіоелектроніки, Харків.

РОЗПІЗНАВАННЯ ЛЮДСЬКИХ ЕМОЦІЙ З ВИКОРИСТАННЯМ НЕЙРОМЕРЕЖЕВИХ ТЕХНОЛОГІЙ

Г.А. Кучук, Б.Г. Саатсазов

Проведено аналіз існуючих методів розпізнавання зображень та описано розпізнавання міміки людини як математичну задачу. Велика увага приділяється нейромережевим методам розпізнавання міміки, особливо багатоваріаційному перцептрону та глибокій згортковій нейронній мережі (ГЗНМ). В результаті проведення науково-дослідницької роботи було створено програмний продукт, який реалізує алгоритм розпізнавання людської міміки з використанням архітектури ГЗНМ. Продукт було створено на мові комп'ютерного програмування Python з використанням сучасних бібліотек комп'ютерного зору `dlib`, `open cv`, та бібліотеки машинного навчання `tensor flow` від Google.

Ключові слова: нейронна мережа, міміка обличчя людини, комп'ютерний zip, python, емоції.

RECOGNITION OF HUMAN EMOTIONS USING NEURAL NETWORKS TECHNOLOGIES

H.A. Kuchuk, B.G. Saatsazov

The analysis of existing methods of image recognition is conducted and the recognition of human mimicry as a mathematical problem is described. Much attention is given to neural network methods of facial recognition, especially the multilayer perceptron and deep convolutional neural network (DCNN). As a result of this work, a software product was created that implements the algorithm for recognizing human facial expressions using the DCNN architecture. The product is created in the language of computer programming Python using modern libraries of computer vision `dlib`, `open cv`, and the library of machine learning `tensor flow` from Google.

Keywords: neural network, people facial mimics, computer vision, python, emotions.