

**В.М. Горяев** [V.M. Goryaev],  
**Г.А. Манкаева** [G.A. Mankaeva],  
**Т.Б. Очир-Горяева** [T.B. Ochir-Goryaeva],  
**А.Б. Мантусов** [A.B. Mantusov],  
**В.В. Лиджи-Гаряев** [V.V. Lidji-Garyaev],  
**А.Б. Корняков** [A.B. Kornyakov]

УДК 004.852  
DOI: 10.37493/2307-910X.2023.1.24

## **АНАЛИЗ РАСПОЗНАВАНИЯ ЭМОЦИЙ НА БАЗЕ МОДЕЛЕЙ ГЛУБОКОЙ КОНВОЛЮЦИОННОЙ НЕЙРОННОЙ СЕТИ**

## **ANALYSIS OF EMOTION RECOGNITION BASED ON DEEP CONVOLUTIONAL NEURAL NETWORK MODELS**

*Калмыцкий государственный университет имени Б. Б. Городовикова,  
г. Элиста, Россия, e-mail: goryaeff@mail.ru / Kalmyk State University named after B.B. Gorodovikov,  
Elista, Russia, e-mail: goryaeff@mail.ru*

### **Аннотация**

*Интерес к распознаванию эмоций на лице растет все больше, и внедряются новые алгоритмы и методы. Последние достижения в области контролируемого и неконтролируемого машинного обучения привели к прорывам в области исследований, и с каждым годом появляется все больше и больше точных систем. Однако, несмотря на значительный прогресс, обнаружение эмоций по-прежнему остается очень сложной задачей. Экспериментальные исследования эмоциональных процессов обычно основаны на наблюдении изображений с аффективным содержанием, включая выражения лица. В статье представлена реализация классификатора эмоциональных выражений путем обучения алгоритма обнаружения лиц на наборе данных АКDEF фронтальных изображений лиц ( $N = 4900$ ). Обученная конволюционная нейронная сеть, достигает средней точности обучения по пересечению над объединением 72,7%.*

*Цель исследования состояла в том, чтобы представить нормативные данные относительно распознавания шести выражений лица с эмоциональным содержанием, а также нейтрального выражения лица, с использованием задания с принудительным выбором. В данном исследовании согласие в распознавании между изображенными и выбранными участниками выражениями варьировалось от 70% (для гнева) до 76% (для счастья). В целом, результаты показывают высокий уровень распознавания представленных выражений лица, что свидетельствует о том, что АКDEF обеспечивает адекватные стимулы для исследований, изучающих распознавание выражений эмоций на лице. Пол и возраст участников оказал ограниченное влияние на распознавание эмоций, однако пол модели и возраст требует дополнительного рассмотрения.*

**Ключевые слова:** эмоции, распознавание эмоций, лицевые выражения, набор АКDEF, глубокая конволюционная нейронная сеть.

### **Abstract**

*Interest in facial emotion recognition is increasing and new algorithms and methods are being introduced. Recent advances in supervised and unsupervised machine learning have led to breakthroughs in research, and more and more accurate systems are emerging every year. However, despite significant progress, emotion detection remains a very*

*challenging task. Experimental studies of emotional processes are usually based on the observation of images with affective content, including facial expressions. This paper presents the implementation of an emotional expression classifier by training a face detection algorithm on a dataset of AKDEF frontal face images (N = 4,900). The trained convolutional*

*The purpose of the study was to present normative data regarding the recognition of six facial expressions with emotional content as well as neutral facial expressions, using a forced-choice task. In this study, agreement in recognition between the expressions depicted and chosen by the participants ranged from 70% (for anger) to 76% (for happiness). Overall, the results show a high level of recognition of the facial expressions presented, suggesting that the AKDEF provides adequate stimuli for studies examining the recognition of facial expressions of emotion. Participant gender and age had limited impact on emotion recognition, but model gender and age require additional consideration*

**Key words:** emotions, emotion recognition, facial expressions, AKDEF set, deep convolutional neural network.

## **Введение**

В наше время для распознавания эмоций чаще всего используют глубокое обучение. Техническое зрение основано на оптимизированной архитектуре, которая использует глубинные сепарабельные свертки для построения глубоких нейронных сетей. При вводе двух простых глобальных гиперпараметров обеспечивается эффективный компромисс между задержкой и точностью. Эти гиперпараметры (множитель ширины и множитель разрешения) позволяют разработчику выбрать модель нужного размера для своего приложения, основываясь компромиссах между ресурсами и точностью и показываем высокую производительность по сравнению с другими популярными моделями на классификации ImageNet. Конволюционные нейронные сети стали повсеместно использоваться в компьютерном зрении с тех пор, как AlexNet популяризировал глубокие конволюционные нейронные сети. Общая тенденция заключается в создании более глубоких и сложных сетей для достижения более высокой точности. Однако эти достижения в области повышения точности не обязательно делают сети более эффективными в отношении размера и скорости. В данной работе описывается эффективная архитектура сети и набор из гиперпараметров для построения моделей с низкой задержкой, которые могут быть легко согласованы с требованиями к проектированию приложений для технического зрения. Для распознавания эмоций чаще всего используют классификатор эмоциональных выражений в реальном времени. В данном проекте был предложен классификатор эмоциональных выражений с поддержкой веб-камеры (вдоль аффективной оси "счастлив/не счастлив"), делая акцент на методах глубокого обучения, включающих обнаружение объектов на основе сверточных нейронных сетей (CNN) с предварительно подготовленной архитектурой глубокого обучения. Результат работы программы предполагал, что полученный детектор объектов будет принимать изображения в качестве входных и возвращает в качестве выходных данных рамки на обнаруженных лицах и подписями названия преобладающей эмоции. Распознавание эмоций широко изучается как в теории, так и в приложениях во многих областях исследований, например, в нейронауке [1,2], биологии [3], психологии [4][5][6] и медицине [7]. Как правило, существует семь типов человеческих эмоций – гнев, отвращение, страх, счастье, грусть, печаль и удивление [8]. Для обнаружения человеческих эмоций существует несколько подходов с использованием различных сигналов поведения человека, таких как речевые сигналы [9], сигналы электроэнцефалограммы (ЭЭГ) [10] и изображения лица [11]. Другие исследователи [12] решали проблему классификации эмоций по речевым сигналам с шестью видами эмоций. В работах по обнаружению эмоций по изображениям человека обычно требуется три следующих этапа: обнаружение лица, выбор признаков из изображений лица и классификация эмоций по

извлеченным признакам. В данной работе используется подход к построению системы обнаружения эмоций в реальном времени с использованием одной камеры. Были рассмотрены семь видов эмоций.

### Материалы и методы

Для проекта была выбрана база данных усредненный KDEF (AKDEF), который представляет собой набор стандартизированных изображений, созданных из исходных изображений базы Karolinska Directed Emotional Faces (KDEF): 4 900 изображений, изображенные эмоции, контролируемая среда (студийная съемка). В проекте был использован набор с предварительно помеченными данными для обнаружения объектов: фронтальные изображения (размер изображений= 562 \* 762 пикселей в формате jpeg).



Рисунок 1. Пример изображения, взятого из набора данных лиц AKDEF

В процессе классификации в качестве признаков использовались шесть единиц действия (AU), вычисляемых устройством Kinect. В таблице 1 приведены примерные значения AU для выражений лица одного из участников. Изображения соответствуют эмоциональным состояниям (ЭС): нейтральное, радость, удивление, гнев, печаль, страх, грусть.

Шаг 1: Выбор изображения

Шаг 2: Распознавание лиц. Извлечение признаков. Преобразование лица в оттенки серого, изменение размера лица (матрица 96 \* 96), предсказание лица с использованием CNN

Шаг 3: Классификация эмоций. Наборы данных, модель CNN

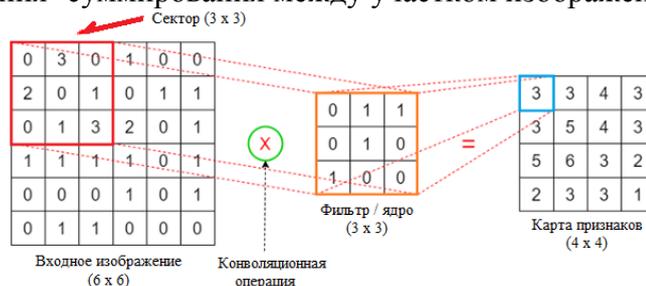
Шаг 4: прогнозирование эмоций

### Архитектура конволюционной нейронной сети (CNN)

Конволюционные нейронные сети (CNN) специально разработаны для работы с изображениями в отличие от многослойного перцептрона (MLP), поскольку, в сети задействовано большое количество параметров даже для небольших изображений. Чтобы использовать MLP для работы с изображениями, нам необходимо сплющить изображение. Если мы это сделаем, то пространственная информация (связи между близлежащими пикселями) будет потеряна. Таким образом, точность будет значительно снижена. CNN могут сохранять пространственную информацию, поскольку они воспринимают изображения в оригинальном формате. CNN могут значительно уменьшить количество параметров в сети.

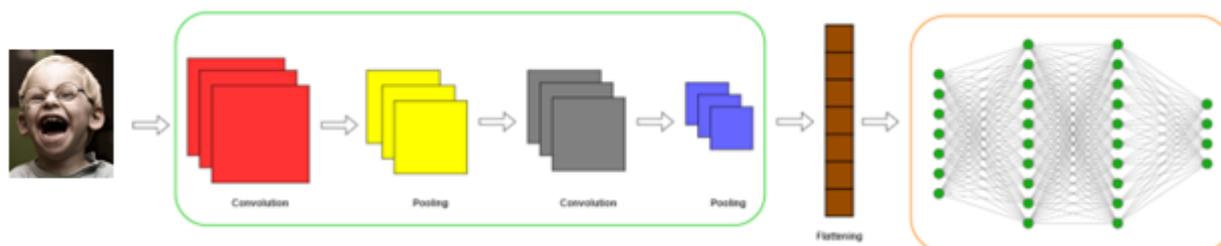
Архитектура CNN сложнее, чем архитектура MLP. В архитектуре CNN существуют различные типы дополнительных слоев и операций, и он принимает изображения в оригинальном формате. В CNN существует три основных типа слоев: Конволюционные слои, слои пулинга и полностью связанные (плотные) слои. Кроме того, после каждого конволюционного слоя и слоя с полным подключением добавляются слои активации. В CNN существует четыре основных типа операций: Операция свертки, операция объединения, операция сглаживания и операция классификации (или другая соответствующая операция) [13,14]. Первый слой в CNN - это сверточный слой, который принимает изображения на вход и начинает их обработку с задачей извлечь набор характеристик из изображения, сохраняя при этом связи между близлежащими пикселями. В конволюционном слое есть три элемента: входное изображение, фильтры и карта признаков (рис.2). Операция свертки

происходит в каждом этом слое. Операция свертки - это не что иное, как операция поэлементного умножения- суммирования между участком изображения и фильтром.



**Рисунок 2. Операция свертки между изображением и фильтром**

Операция свертки происходит между участком изображения и фильтром, на выходе получается карта признаков. Карта признаков хранит результаты различных операций свертки между различными секциями изображения и фильтром. Это будет входом для следующего слоя объединения. Количество элементов в карте признаков равно количеству различных участков изображения, которые мы получили, перемещая фильтры по изображению. На приведенной выше схеме показана операция свертки между участком изображения и одним фильтром. Количество шагов (пикселей), на которое мы сдвигаем фильтр по входному изображению, называется Stride. Сдвиг может быть выполнен как по горизонтали, так и по вертикали. Здесь мы используем Stride=1. Размер карты признаков меньше размера входного изображения. Размер карты признаков также зависит от Stride. Если мы используем Stride=2, размер еще больше уменьшится. Если в CNN есть несколько сверточных слоев, размер карты признаков еще больше уменьшится в конце, так что мы не сможем выполнять другие операции над картой признаков. Чтобы избежать этого, мы применяем Padding к входному изображению. Padding – это гиперпараметр, который надо настроить в конволюционном слое. Он добавляет дополнительные пиксели с нулевыми значениями к каждой стороне изображения. Это помогает получить карту признаков того же размера, что и входное изображение с задачей извлечь наиболее релевантные признаки, получив максимальное число или усреднив их, уменьшая при этом размерность выходного сигнала, полученного от предыдущих сверточных слоев, количество параметров в сети и устраняя шум, присутствующий в характеристиках, извлеченных предыдущими конволюционными слоями. В слое объединения есть три элемента: карта признаков, фильтр и объединенная карта признаков. Операция объединения происходит в каждом слое объединения. Операция объединения происходит между участком карты признаков и фильтром. На выходе получается объединенная карта характеристик. Существует два типа операций объединения: максимальное объединение и объединение по среднему значению. После применения объединения к карте особенностей количество каналов не изменяется. Это означает, что у нас одинаковое количество каналов в карте признаков и в объединенной карте признаков (ОКП). Если карта характеристик имеет несколько каналов, следует использовать фильтр с таким же количеством каналов. Операции объединения будут выполняться на каждом канале независимо. В отличие от сплющивания исходного изображения, при сплющивании объединенных карт сохраняются важные зависимости между пикселями (рис.3). В последних слоях входом является предыдущий слой.



**Рисунок 3. Общая архитектура CNN**

Полностью связанных слоев может быть несколько. Последний слой выполняет задачу классификации (или другую соответствующую задачу). В каждом полностью подключенном слое используется функция активации с задачей классификации обнаруженных признаков на изображении в метку класса. В типичной CNN слои располагаются в следующем порядке. CNN принимает на вход изображение как оно есть. Входное изображение проходит через ряд слоев и операций. Конволюционные и объединяющие слои необходимы для извлечения особенностей из изображения, сохраняя при этом важные зависимости между пикселями. Они также уменьшают размерность (количество пикселей) исходного изображения. Эти слои используются вместе как пары. Активация ReLU используется в каждом сверточном слое. Количество фильтров увеличивается в каждом конволюционном слое. Например, если использовать 16 фильтров в первом сверточном слое, то в следующем сверточном слое обычно используется 32 фильтра, и так далее. Первые несколько слоев фокусируются на менее важных паттернах (таких, как края) в данных изображения. Конечные слои находят более сложные паттерны (например, нос, глаза на изображении лица). Последний слой решает задачу классификации. Активация ReLU используется в каждом полностью связанном слое, за исключением последнего слоя, в котором используется активация Softmax для многоклассовой классификации (табл. 1). Слои в первой рамке вносят некоторые изменения в изображение. Во втором поле находится MLP. Между блоками находится упрощенный слой.

**Таблица 1. – Слой архитектуры нейросети**

Модель 1			Модель 2		
Layer (type) #	Output Shape	Param	Layer (type)	Output Shape	Param #
conv2d_4 (Conv2D)	(None, 48, 48, 64)	640	conv2d_8 (Conv2D)	(None, 48, 48, 64)	128
dropout (Dropout)	(None, 48, 48, 64)	0	batch_normalization	(Bat(None, 48, 48, 64)	256
max_pooling2d_4 (MaxPooling2D)	(None, 24, 24, 64)	0	dropout_5 (Dropout)	(None, 48, 48, 64)	0
conv2d_5 (Conv2D)	(None, 24, 24, 64)	36928	conv2d_9 (Conv2D)	(None, 48, 48, 128)	3856
dropout_1 (Dropout)	(None, 24, 24, 64)	0	batch_normalization_1	(Batch(None,48,48,128)	512
max_pooling2d_5 (MaxPooling2D)	(None, 12, 12, 64)	0	dropout_6 (Dropout)	(None, 48, 48, 128)	0
conv2d_6 (Conv2D)	(None, 12, 12, 128)	73856	conv2d_10 (Conv2D)	(None, 48, 48, 256)	19456
dropout_2 (Dropout)	(None, 12, 12, 128)	0	batch_normalization_2	(Batch(None, 48, 48, 256)	1024
max_pooling2d_6 (MaxPooling2D)	(None, 6, 6, 128)	0	max_pooling2d_8 (MaxPooling2D)	(None, 24, 24, 256)	0
conv2d_7 (Conv2D)	(None, 6, 6, 256)	295168	dropout_7 (Dropout)	(None, 24, 24, 256)	0
dropout_3 (Dropout)	(None, 6, 6, 256)	0	flatten_2 (Flatten)	(None, 147456)	0
max_pooling2d_7 (MaxPooling2D)	(None, 3, 3, 256)	0	dense_6 (Dense)	(None, 128)	874496
flatten_1 (Flatten)	(None, 2304)	0	batch_normalization_3	(Batch(None, 128)	512
			activation (Activation)	(None, 128)	

dense_3 (Dense) 590080	(None, 256)	0	dropout_8 (Dropout)	(None, 128)
dense_4 (Dense) 32896	(None, 128)	0	dense_7 (Dense)	(None, 256)
dropout_4 (Dropout) 0	(None, 128)	33024	batch_normalization_4	(Batch(None,256)
dense_5 (Dense) 903	(None, 7)	1024	activation_1 (Activation)	(None, 256)
=====		0	dropout_9 (Dropout)	(None, 256)
===		0	dense_8 (Dense)	(None, 7)
Total params: 1,030,471		1799	=====	
Trainable params: 1,030,471			===	
Non-trainable params: 0			Total params: 19,806,087	
			Trainable params: 19,804,423	
			Non-trainable params: 1,664	
Слои свертки, слои Pooldown, выпадающие слои и плотные слои			Слои конволюции, слой Pooldown, слои Dropout, пакетной нормализации и плотных слоев	

### Результаты и обсуждение

В методе на основе CNN лучше улавливает длительные временные характеристики контекста, но не очень – пиковые характеристики. Из-за этого проседают распознавания эмоций гнева, страха и удивления.

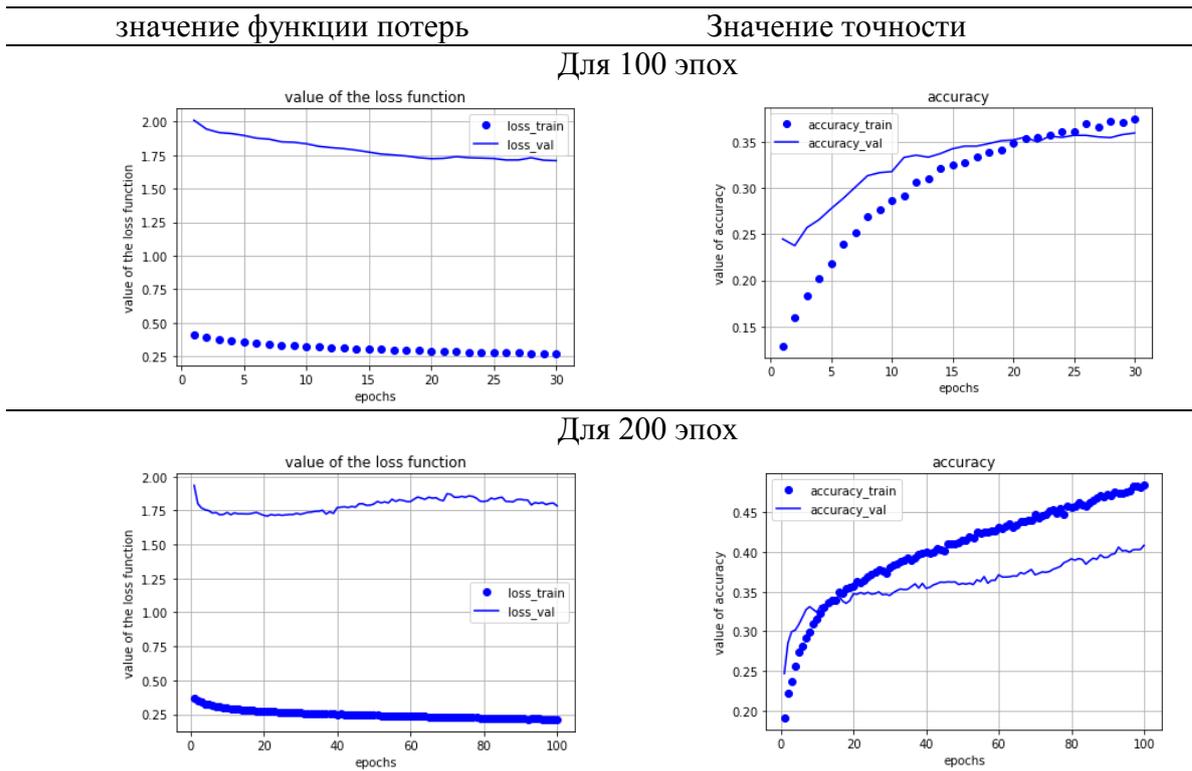
Таблица 2. Результаты на тестовых данных. Общая точность 57,2%

	Гнев	Страх	Радость	Нейтральное	Грусть	Удивление
Гнев	77.11	04.82	03.61	04.82	04.82	03.61
Страх	24.24	42.42	03.03	13.64	03.03	12.12
Радость	13.33	02.22	74.07	04.44	05.93	00.00
Нейтральное	14.94	00.57	08.05	53.45	18.39	02.87
Грусть	22.54	02.82	19.72	08.45	42.25	04.23
Удивление	25.00	17.86	07.14	17.86	14.29	14.29

Наша модель CNN была обучена на всех 4900 учебных изображениях в течение 100/200 эпох с параметрами по умолчанию [15]. Основной метрикой точности была метрика Accuracy- точность предсказанных классов эмоций. В каждую из эпох можно наблюдать за ходом обучения модели. Обычно точность данных обучения и проверки со временем увеличивается, также можно проанализировать историю модели. Для 100/200 эпох точность тестовых данных (24.49/59.23) чуть меньше, чем точность обучающих данных (25.13/89.02), при этом можно наблюдать, что потери при проверке увеличиваются с течением времени. Это также говорит о том, что модель плохо справлялась с данными проверки. Таким образом, надо двигаться дальше и вносить поправки в параметры. Для решения, ограниченного предложенным пространством поиска, скорость обучения была установлена на  $1e-3$ , использовался алгоритм оптимизации Adam [16]. Размер пакета был установлен на 64, размер ядер в конволюционных слоях был установлен на  $3 \times 3$ , а в качестве функции активации использовалась выпрямленная линейная единица (ReLU). Для классификации использовались 2 полностью связанных слоя, первый из которых имел 256 нейронов, а второй – 7 нейронов, по одному на каждую эмоцию в базе данных. Каждая модель отправляется на обучение в течение 100/200 эпох для проверки ее точности. Следует отметить, что лучшее решение из моделей показало точность валидации 63,22% с 200 эпох обучения. Схема лучшей модели представлена в табл.1, модель 2.



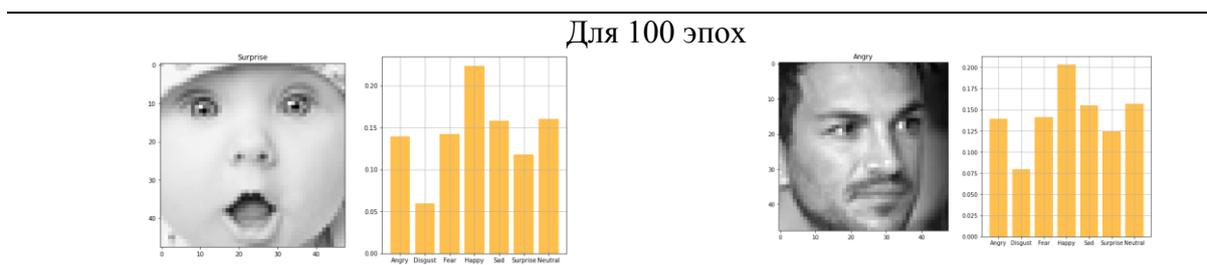
Таблица 4. – Сравнение качества прогноза нейросети для модели 2.



Train Accuracy: 99.77010488510132, 9ms/step - loss: 0.0060  
 Test Accuracy: 55.69796562194824, 9ms/step - loss: 2.9272  
 Validation Accuracy: 56.61744475364685, 9ms/step - loss: 2.7959

Модель имеет общее количество параметров 19,806,087. Для сравнения, модель VGG с общим количеством параметров от 133 до 144 миллионов параметров и размером в 10 раз больше достигла точности 72,7% на базе данных FER2013 [17], VGG-16 [18] использовали модели и варианты VGG, и поэтому общее количество параметров для их моделей находится в диапазоне 133-144 миллионов или выше. Для проверки общей производительности предложенная модель была обучена в течение 200 эпох на базе данных AKDEF. На этом этапе была установлена скорость обучения, выбран оптимизатор Adam со скоростью обучения 0,001, а размер партии был установлен на 64. После обучения модель получила точность 55.70% с потерей 2.93 на тестовых данных и точность 56.62% с потерей 2,79 на валидных данных AKDEF.

Таблица 5. – Тест для модели 1



Для 200 эпох

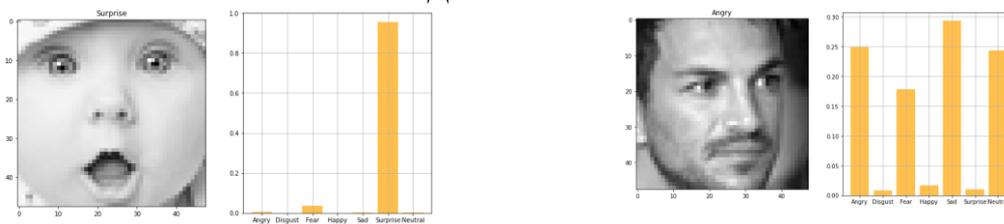
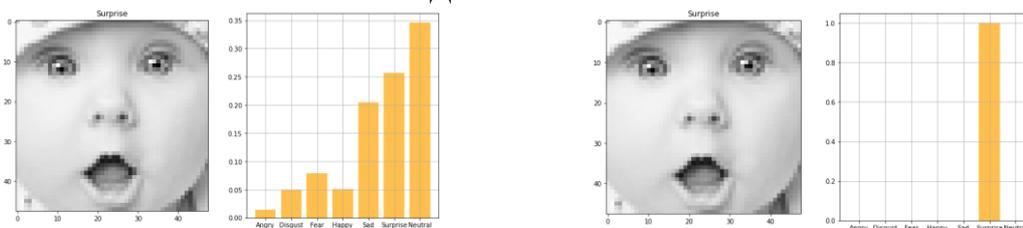
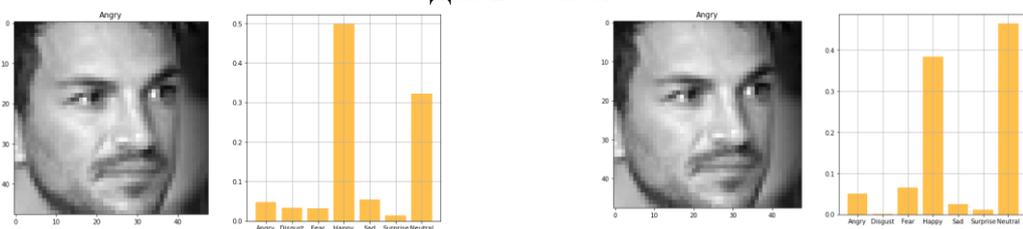


Таблица 6. – Тест для модели 2

Для 100 эпох



Для 200 эпох



Мы обучили модель для  $epoch = 100$  и  $epoch = 200$ . Более лучшие результаты были получены для 100 эпох в модели 1, что показывает на проблему переобучения данных на 200 эпохах. История Epochs показывает, что точность постепенно увеличивается, достигая максимальной точности как для обучающего, так и для проверочного набора, но в конце модель начинает переобучаться. При этом детектор объектов CNN достиг средней точности обучения 74%, максимальной точности для валидационных данных 59.23% и точности обучения 89,3% (набор обучающих примеров  $N = 4900$ ).



А. Нейтральное



Б. Гнев



В. Смех



Г. Печаль



Д. Удивление



Е. Страх

Рисунок 4. Граничные рамки с классом эмоций и временем обнаружения

Детектор объектов рисует рамки вокруг лиц, обнаруженных на изображениях, и присваивает вероятность того, что области, обведенные рамкой, принадлежат к классу "эмоции". Как оказалось, окончательная точность проверки CNN составила около 57%, что, конечно, немного. Потому что наши выражения обычно состоят из комбинации эмоций и использовать только одну метку для представления выражения может быть сложно. В этом случае, когда модель предсказывает неправильно, правильная метка часто является второй наиболее вероятной эмоцией, как показано в тестовых таблицах. Модель действительно хорошо справляется с классификацией положительных эмоций, что приводит к относительно высокой точности оценок для счастья и удивления. Точность счастья (happy) составляет 79,5%, что можно объяснить наличием наибольшего количества их примеров в обучающем наборе. Интересно, что удивление (surprise) имеет точность 69,1%, хотя у них наименьшее количество примеров в обучающем наборе. Производительность модели в среднем кажется слабее при отрицательных эмоциях. В частности, эмоция грусть (sad) имеет точность менее 40%. Модель часто ошибочно классифицировала злость, страх и нейтраль как грустные. Кроме того, он больше всего запутывается при прогнозировании грустных и нейтральных лиц, потому что эти две эмоции, вероятно, наименее выразительны.

### **Заключение**

Цель работы представить метод повышения точности модели конволюционной нейронной сети путем оптимизации гиперпараметров и ее архитектуры. Результаты эксперимента показали, что с помощью этого метода может быть получена компактная модель, основанная только на CNN с точностью 57%. Поверхностная оптимизация гиперпараметров имела целью показать, что эффективное решение может быть достигнуто в пространстве поиска, в котором предыдущие результаты считаются локальными минимумами. Результат является удовлетворительным, учитывая ограничения, наложенные на оптимизацию, которая проводилась на небольшом количестве гиперпараметров [19]. Наш детектор объектов не показал необычайно высоких результатов, хотя он продемонстрировал достаточно высокую точность для обучающих данных (модель 2, 200 эпох) – 99.77%. При этом надо учитывать, что выборки для наборов данных, как правило, для исследования включала большее количество женщин. Мы предложили подход, основанный на комбинировании моделей для распознавания эмоций в естественных условиях, и достигли конечной точности 51,7% на валидационном наборе. Результаты также показывают, что некоторые выражения лица распознавались легче, чем другие. Как и в оригинальном исследовании, положительные эмоции распознавались легче, чем негативные. Более конкретно, счастье было самой легко распознаваемой эмоцией, с высоким, как и в других известных исследованиях, с высоким процентом правильных ответов (97%). Наиболее распознаваемым выражение лица было нейтральное (69%), а презрение показало самый низкий уровень распознавания (24%).

Поскольку у нас не было достаточно большого обучающего набора, чтобы провести оптимальное разделение обучения и тестирования, мы не проверяли работу нашего детектора объектов на удержанном тестовом наборе. Поэтому вполне возможно, что наш детектор объектов может демонстрировать высокую степень чрезмерной подгонки к обучающему множеству. В этом случае уменьшение гибкости нашей модели (например, с помощью методов регуляризации) может снизить степень переподгонки в будущих итерациях нашей модели. В целом ясно, что распознавание, основанное только на распознавании с камеры, не дает необходимую точность, поэтому визуальное распознавание можно интегрировать с другими сигналами поведения человека, с такими как аудио информация, ЭКГ, электроэнцефалограммы и т.д.

## ЛИТЕРАТУРА

1. Liu Yisi, Sourina Olga and Nguyen Minh Khoa, "Real-Time EEG-Based Emotion Recognition and Its Applications", Transactions on Computational Science XII: Special Issue on Cyberworlds, 2011. Pp. 256-277.
2. Stickel Christian, Ebner Martin, Steinbach-Nordmann Silke, Searle Gig and Holzinger Andreas, "Emotion Detection: Application of the Valence Arousal Space for Rapid Biological Usability Testing to Enhance Universal Access", Lecture Notes in Computer Science, 2009. Vol. 5614. Pp. 615-624.
3. Александров А.А., Кирпичников А.П., Ляшева С.А., Шлеймович М.П. Анализ эмоционального состояния человека на изображении // Вестник технологического университета. 2019. Т. 22, № 8. С. 120-123.
4. D. Galati, R. Miceli and B. Sini, "Judging and coding facial expression of emotions in congenitally blind children", International Journal of Behavioral Development, 2001. Vol. 25, No 3. Pp. 268-278.
5. H. A. Elfenbein and N. Ambady, "On the universality and cultural specificity of emotion recognition: A meta-analysis", Psychological Bulletin, 2002, Vol. 128, No 2. Pp. 205-235.
6. Рюмина Е.В., Карпов А.А. Аналитический обзор методов распознавания эмоций по выражениям лица человека // Научно-технический вестник информационных технологий, механики и оптики. 2020. Т. 20, № 2. С. 163-176.
7. Степанова О.А., Ивановский Л.И., Хрящев В.В. Использование глубокого обучения и сверточных нейронных сетей для анализа выражения лица // DSPA: Вопросы применения цифровой обработки сигналов. 2018. Т. 8, № 4. С. 170-173.
8. Livingstone S.R., Russo F.A. The Ryerson audio-visual database of emotional speech and song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English // PLoS ONE. 2018. Vol. 13, No. 5. Pp. e0196391. DOI: 10.1371/journal.pone.0196391
9. Zangeneh Soroush, Morteza & Maghooli, Keivan & Setarehdan, Kamal & Motie Nasrabadi, Ali. (2017). A Review on EEG Signals Based Emotion Recognition. International Clinical Neuroscience Journal. 4. Pp. 118-129.
10. Vahid Kazemi and Josephine Sullivan, "One millisecond face alignment with an ensemble of regression trees", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014. Pp. 1867-1874.
11. Богданова Д. Р., Акушев А. Т. Распознавание эмоций по речевому сигналу // E-Scio. 2021. №. 6 (57). С. 43-52.
12. D. Bertero and P. Fung. A first look into a Convolutional Neural Network for speech emotion detection. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2017. Pp. 5115-5119, DOI: 10.1109/ICASSP.2017.7953131
13. Peres D. Emotional Thin-Slicing: A Proposal for a Short- and Long-Term Division of Emotional Speech. 2017. Pp. 591-595. 10.21437/Interspeech.2017-1719.
14. Горяев В.М., Бембитов Д.Б. и др. Прогнозирование степных пожаров с использованием данных дистанционного зондирования временных рядов Современные наукоемкие технологии. 2020. № 12-1. С. 15-19.
15. Baez John C., Dolan James. Higher-dimensional algebra and topological quantum field theory // Journal of Mathematical Physics. 1995. Vol. 36, No. 11. Pp. 6073-6105.
16. Аггарвал Ч. Нейронные сети и глубокое обучение: учебный курс: Пер. с англ. СПб.: ООО «Диалектика», 2020. 752 с.
17. Wang Kay Ngai et al. Emotion recognition based on convolutional neural networks and heterogeneous bio-signal data sources. Information Fusion. 2022. Vol. 77. Pp. 107-117.
18. Сикорский О.С. Обзор свёрточных нейронных сетей для задачи классификации изображений // Новые информационные технологии в автоматизированных системах. Москва, 2017. No 20. С. 37-42.

19. Goryaev V.M. et al. Development of a statistical forecast model to improve accuracy based on statistical analysis of weather historical data for the kalmyk region // IOP Conference Series: Earth and Environmental Science. International Symposium on Earth Sciences: History, Contemporary Issues and Prospects. 2019. Pp. 012058.

## REFERENCES

1. Liu Yisi, Sourina Olga and Nguyen Minh Khoa, "Real-Time EEG-Based Emotion Recognition and Its Applications", Transactions on Computational Science XII: Special Issue on Cyberworlds, 2011. Pp. 256-277.
2. Stickel Christian, Ebner Martin, Steinbach-Nordmann Silke, Searle Gig and Holzinger Andreas, "Emotion Detection: Application of the Valence Arousal Space for Rapid Biological Usability Testing to Enhance Universal Access", Lecture Notes in Computer Science, 2009. Vol. 5614. Pp. 615-624.
3. Aleksandrov A.A., Kirpichnikov A.P., Lyasheva S.A., Shleimovich M.P. Analiz ehmtotsional'nogo sostoyaniya cheloveka na izobrazhenii // Vestnik tekhnologicheskogo universiteta. 2019. Vol. 22, No. 8. P. 120-123.
4. D. Galati, R. Miceli and B. Sini, "Judging and coding facial expression of emotions in congenitally blind children", International Journal of Behavioral Development, 2001. Vol. 25, No. 3. Pp. 268-278.
5. H. A. Elfenbein and N. Ambady, "On the universality and cultural specificity of emotion recognition: A meta-analysis", Psychological Bulletin, 2002, Vol. 128, No. 2. Pp. 205-235.
6. Ryumina E.V., Karpov A.A. Analiticheskii obzor metodov raspoznavaniya ehmtotsii po vyrazheniyam litsa cheloveka // Nauchno-tekhnicheskii vestnik informatsionnykh tekhnologii, mekhaniki i optiki. 2020. T. 20, No. 2. P. 163-176.
7. Stepanova O.A., Ivanovskii L.I., Khryashchev V.V. Ispol'zovanie glubokogo obucheniya i svertochnykh neironnykh setei dlya analiza vyrazheniya litsa // DSPA: Voprosy primeneniya tsifrovoy obrabotki signalov. 2018. Vol. 8, No. 4. P. 170-173.
8. Livingstone S.R., Russo F.A. The Ryerson audio-visual database of emotional speech and song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English // PLoS ONE. 2018. Vol. 13, No. 5. Pp.0196391. DOI: 10.1371/journal.pone.0196391.
9. Zangeneh Soroush, Morteza & Maghooli, Keivan & Setarehdan, Kamal & Motie Nasrabadi, Ali. (2017). A Review on EEG Signals Based Emotion Recognition. International Clinical Neuroscience Journal. 4. Pp. 118-129.
10. Vahid Kazemi and Josephine Sullivan, "One millisecond face alignment with an ensemble of regression trees", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014. Pp. 1867-1874.
11. Bogdanova D. R., Akushev A. T. Raspoznavanie ehmtotsii po rechevomu signalu // E-Scio. 2021. No. 6 (57). P. 43-52.
12. D. Bertero and P. Fung. A first look into a Convolutional Neural Network for speech emotion detection. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2017. Pp. 5115-5119, DOI: 10.1109/ICASSP.2017.7953131.
13. Peres D. Emotional Thin-Slicing: A Proposal for a Short- and Long-Term Division of Emotional Speech. 2017. Pp. 591-595. 10.21437/Interspeech.2017-1719.
14. Goryaev V.M., Bembitov D.B. i dr. Prognozirovanie stepnykh pozharov s ispol'zovaniem dannykh distantsionnogo zondirovaniya vremennykh ryadov Sovremennye naukoemkie tekhnologii. 2020. No. 12-1. P. 15-19.
15. Baez John C., Dolan James. Higher-dimensional algebra and topological quantum field theory // Journal of Mathematical Physics. 995. Vol. 36, No 11. P. 6073-6105.
16. Aggarval CH. Neironnye seti i glubokoe obuchenie: uchebnyi kurs: Per. s angl. SPb.: OOO "Dialektika", 2020. 752 p.
17. Wang Kay Ngai et al. Emotion recognition based on convolutional neural networks and heterogeneous bio-signal data sources. Information Fusion. 2022. Vol. 77. Pp. 107-117.

18. Sikorskii O.S. Obzor svertochnykh neironnykh setei dlya zadachi klassifikatsii izobrazhenii // Novye informatsionnye tekhnologii v avtomatizirovannykh sistemakh. Moskva, 2017. No. 20. P. 37-42.

19. Goryaev V.M. et al. Development of a statistical forecast model to improve accuracy based on statistical analysis of weather historical data for the kalmyk region // IOP Conference Series: Earth and Environmental Science. International Symposium on Earth Sciences: History, Contemporary Issues and Prospects. 2019. Pp. 012058.

### ОБ АВТОРАХ / ABOUT THE AUTHORS

**Горяев Владимир Михайлович**, кандидат педагогических наук, доцент, Калмыцкий государственный университет имени Б.Б. Городовикова, Российская Федерация, Республика Калмыкия, г. Элиста, ул. Пушкина, д. 11, e-mail: [goryaeff@mail.ru](mailto:goryaeff@mail.ru)

**Vladimir M. Goryaev**, Cand. Sci. (Pedag.), Associate Professor, Kalmyk State University, Russian Federation, 358000, Republic of Kalmykia, Elista, Pushkin str., 11, e-mail: [goryaeff@mail.ru](mailto:goryaeff@mail.ru)

**Манкаева Галина Алексеевна**, старший преподаватель, Калмыцкий государственный университет имени Б.Б. Городовикова, Российская Федерация, Республика Калмыкия, г. Элиста, ул. Пушкина, д. 11, e-mail: [mankaeva.galina@yandex.ru](mailto:mankaeva.galina@yandex.ru)

**Galina A. Mankaeva**, Senior Lecturer, Kalmyk State University, Russian Federation 358000, Republic of Kalmykia, Elista, Pushkin str., 11, e-mail: [mankaeva.galina@yandex.ru](mailto:mankaeva.galina@yandex.ru)

**Очир-Горяева Тамара Борисовна**, кандидат экономических наук, доцент, Калмыцкий государственный университет имени Б.Б. Городовикова, Российская Федерация, Республика Калмыкия, г. Элиста, ул. Пушкина, д. 11, e-mail: [goryaeff@mail.ru](mailto:goryaeff@mail.ru)

**Tamara B. Ochir-Goryaeva**, Cand. Sci. (Econ.), Associate Professor, Kalmyk State University, Russian Federation 358000, Republic of Kalmykia, Elista, Pushkin str., 11, e-mail: [goryaeff@mail.ru](mailto:goryaeff@mail.ru)

**Мантусов Анатолий Бадмаевич**, кандидат педагогических наук, доцент, Калмыцкий государственный университет имени Б.Б. Городовикова, Российская Федерация, Республика Калмыкия, г. Элиста, ул. Пушкина, д. 11, e-mail: [Mantusov@kalmsu.ru](mailto:Mantusov@kalmsu.ru)

**Anatoliy B. Mantusov**, Cand. Sci. (Pedag.), Associate Professor, Kalmyk State University, Russian Federation 358000, Republic of Kalmykia, Elista, Pushkin str., 11, e-mail: [Mantusov@kalmsu.ru](mailto:Mantusov@kalmsu.ru)

**Манджиева Александра Нимяевна**, кандидат сельскохозяйственных наук, Доцент, Калмыцкий государственный университет имени Б.Б. Городовикова, Российская Федерация, Республика Калмыкия, г. Элиста, ул. Пушкина, д. 11, e-mail: [Mandzhieva@kalmsu.ru](mailto:Mandzhieva@kalmsu.ru)

**Alexandra N. Mandzhieva**, Cand. Sci. (Agric.), Associate Professor, Kalmyk State University, Russian Federation 358000, Republic of Kalmykia, Elista, Pushkin str., 11, e-mail: [Mandzhieva@kalmsu.ru](mailto:Mandzhieva@kalmsu.ru)

**Борисов Игорь Владимирович**, кандидат политических наук, старший преподаватель, Калмыцкий государственный университет имени Б.Б. Городовикова, Российская Федерация, Республика Калмыкия, г. Элиста, ул. Пушкина, д. 11, e-mail: [Borisov@kalmsu.ru](mailto:Borisov@kalmsu.ru)

**Igor V. Borisov**, Cand. Sci. (Polit.), Senior Lecturer, Kalmyk State University, Russian Federation 358000, Republic of Kalmykia, Elista, Pushkin str., 11, e-mail: [Borisov@kalmsu.ru](mailto:Borisov@kalmsu.ru)

Дата поступления в редакцию: 12.01.2023

После рецензирования: 18.02.2023

Дата принятия к публикации: 19.03.2023