

## **ИССЛЕДОВАНИЕ ТОЧНОСТИ НЕЙРОСЕТЕВОГО АЛГОРИТМА ОПРЕДЕЛЕНИЯ ЭМОЦИЙ ПРИ РАЗЛИЧНЫХ УГЛАХ ПОВОРОТА ЛИЦА**

студент 4 курса, 5ПИ группы Небышинец С. В.

*Научный руководитель - канд. техн. наук Садов В. С.*

Белорусский государственный университет

Минск, Беларусь

Аффективные вычисления — это изучение и разработка систем и устройств, которые могут распознавать, интерпретировать, обрабатывать и имитировать человеческие эмоции [3]. Целью работ в данной области является способность моделировать эмпатию. Машина должна интерпретировать эмоциональное состояние людей и адаптировать к ним свое поведение, давая соответствующую реакцию на эти эмоции.

Одним из ответвлений аффективных вычислений является определение эмоций по снимку лица человека. Данная область приобрела особую популярность в недавнее время из-за высокой доступности и точности систем распознавания эмоций по лицу человека. Данным системам нашли применение в таких отраслях как безопасность дорожного движения, маркетинг, образование и общественная безопасность. Одной из прикладных сфер применения систем распознавания эмоций по лицу является оценка качества мультимедиа продукта.

Разработка интерфейса таких систем с учетом пользовательского опыта связана с необходимостью постоянно переделывать интерфейс для того, чтобы на основании новых данных об отношении потребителя к продукту улучшить качество предоставляемых мультимедиа услуг. Сбор данных от пользователей об их опыте использования интерфейса зачастую занимает большое время и может не дать желаемого результата. В данном случае на помощь приходит система распознавания эмоций по лицу. Она определяет, в какие моменты использования интерфейса человек испытывал целевые эмоции, что значительно ускоряет сбор информации об опыте использования продукта и, следовательно, разработку интерфейса.

Однако у систем распознавания эмоций по лицу есть проблема – условия съёмки. Значительную роль в точности алгоритма играет угол поворота лица.

Процесс распознавания эмоций состоит из трех этапов: детектирование лица, выделение черт лица и классификация эмоции (рисунок 1).

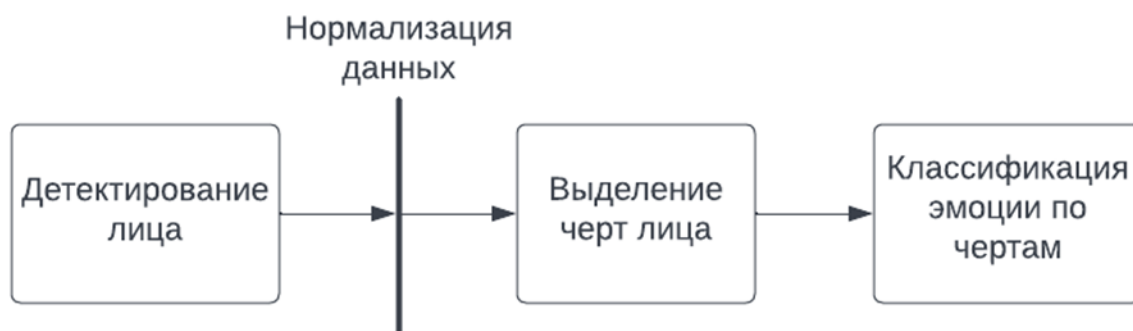


Рисунок 1. Общий алгоритм определения эмоций лица.

Первым этапом является детектирование лица с последующей нормализацией. Одним из самых распространенных методов детектирование лица является метод Виолы-Джонса. Данный метод основан на использовании признаков Хаара для обнаружения объектов, в частности, человеческого лица [2].

Следующим этапом является определение выражения лица. Любая эмоция лица определяется сокращением лицевых мышц. Данные сокращения мышц формируют определенное выражение лица, которые можно видеть в качестве отдельных признаков, таких как, например, приподнятые уголки губ, широко раскрытые глаза.

Последним этапом в определении эмоций лица является классификация эмоции лица. Описанного выражения лица достаточно для того, чтобы классифицировать эмоцию, изображенную на лице.

Самым распространенным и инструментом для определения выражения лица с последующей классификацией эмоции являются сверточные нейронные сети, использующие специализированные сверточные слои, позволяющие обучать нейронную сеть распознавать образы в двумерном пространстве [3].

В реализации алгоритма нейросетевого распознавания эмоций по лицу, рассматриваемого в данной работе, были применены метод Виолы-Джонса для детектирования лица и сверточная нейронная сеть для классификации 7 основных эмоций. Сверточная нейронная сеть имеет 30 слоев, 4 из которых сверточные слои. Общее количество тренировочных параметров – 4474759. Датасетом для

обучения послужил FER-2013, содержащий 30000 черно-белых изображений лица. Обучение модели заняло 25 часов.

Для непосредственной оценки точности работы данной модели по определению эмоций при различных углах поворота лица необходим датасет, содержащий изображение лица, информацию по настоящей эмоции и углах поворота по вертикали и горизонтали. Пример фотографии из собранного датасета, изображающей радость при углах наклона  $-15^{\circ}$  по вертикали и  $15^{\circ}$  по горизонтали представлен на рисунке 2..



Рисунок 2. Пример фотографии из собранного датасета, изображающий радость при углах наклона  $-15^{\circ}$  по вертикали и  $15^{\circ}$  по горизонтали

Сбор данных для разных углов поворота и наклона лица происходил на углах  $0, 15$  и  $30$  градусов по горизонтали и  $15, 0, -15, -30$  градусов по вертикали. Определение угла поворота лица происходило без использования программных инструментов, с помощью меток на стене, предварительно размеченных с использованием угломера, и фиксации тела и рук человека на одном месте.

Чтобы выбрать наиболее важные метрики для исследования эффективности алгоритма на различных углах поворота лица нужно принять во внимание то, что непосредственно для задачи оценки всплесков эмоций в большом количестве видеозаписи можно пренебречь ложно отрицательными оценками, так как количество видеозаписи велико и потеря нескольких всплесков не скажется на

общей эффективности работы алгоритма. Таким образом, полнота играет вторичную роль в оценке эффективности работы алгоритма по распознаванию эмоции потребителя.

Наиболее важную роль в оценке эффективности играет точность. Однако, нужно понимать и то, что точность не может быть единственной оценкой для данной задачи, так как во-первых выбор результата в итоге производится по топ-1 эмоции, во-вторых оценка лишь по точности может привести к тому, то из рассмотрения выходят ситуации чрезмерного забвения ложно отрицательных предсказаний [4].

Таким образом, помимо точности, важно оценивать, во-первых, относительность топ-1, во-вторых, ассигасу. Здесь ассигасу означает отношение всех верных предсказаний ко всем предсказаниям. Точность же определяется как отношение истинно-положительных предсказаний ко всем положительным предсказаниям.

Относительность топ-1 – это метрика, позволяющая оценивать, насколько близко к топ-1 находится настоящая эмоция. Этого можно достиг, оценивая относительно каждой эмоции среднее от того, во сколько раз вероятность настоящей эмоции меньше топ-1, где настоящая эмоция – это как раз оцениваемая эмоция. Математически эту метрику можно выразить следующим образом:

$$RT1 = \sum_N \frac{PT1}{T} , \quad (1)$$

где RT1 – оценка относительности топ-1, N – оцениваемая выборка, PT1 – оцененная на данной записи эмоция топ-1, T – истинная изображенная в записи эмоция.

Общая оценка ассигасу, точности и полноты по топ-1 по всем углам изображения, следующая: ассигасу – 39,34%, точность – 44,60%, полнота – 40,47%. С учетом семи оцениваемых классов эмоций и различных условий поворота лица результаты модели являются удовлетворительными.

Для угла поворота лица по вертикали 0° и по горизонтали 0°, то есть лицо анфас, точность – 61.34%, полнота – 65.24%, ассигасу – 62.86%. Как можно заметить, эффективность классификации при повороте лица анфас, то есть значение по-умолчанию, удовлетворительно на основании того, что как точность, так и полнота являются довольно высокими. Это подтверждается высоким

значением показателя ассурасу. Как можно заметить, общие результаты по всем углам поворота лица значительно хуже, чем при идеальных условиях.

Результаты оценки точности для каждого из исследуемых углов поворота лица представлены на рисунке 1. Под каждой фигурой числа внизу - углы поворота головы по вертикали и по горизонтали.

Исходя из проанализированных данных по точности, в краткой форме представленных на рисунке 2, можно увидеть несколько тенденций. Во-первых, сильное уменьшение точности при положительных углах наклона головы по вертикали. Во-вторых, при небольшом угле поворота лица вправо и даже при большом повороте лица вниз потери в точности незначительны по сравнению с потерями точности при повороте головы на 30 градусов вправо.

Результаты оценки относительности топ-1 для каждого из исследуемых углов поворота лица представлены на рисунке 2. Под каждой фигурой числа внизу - углы поворота головы по вертикали и по горизонтали.

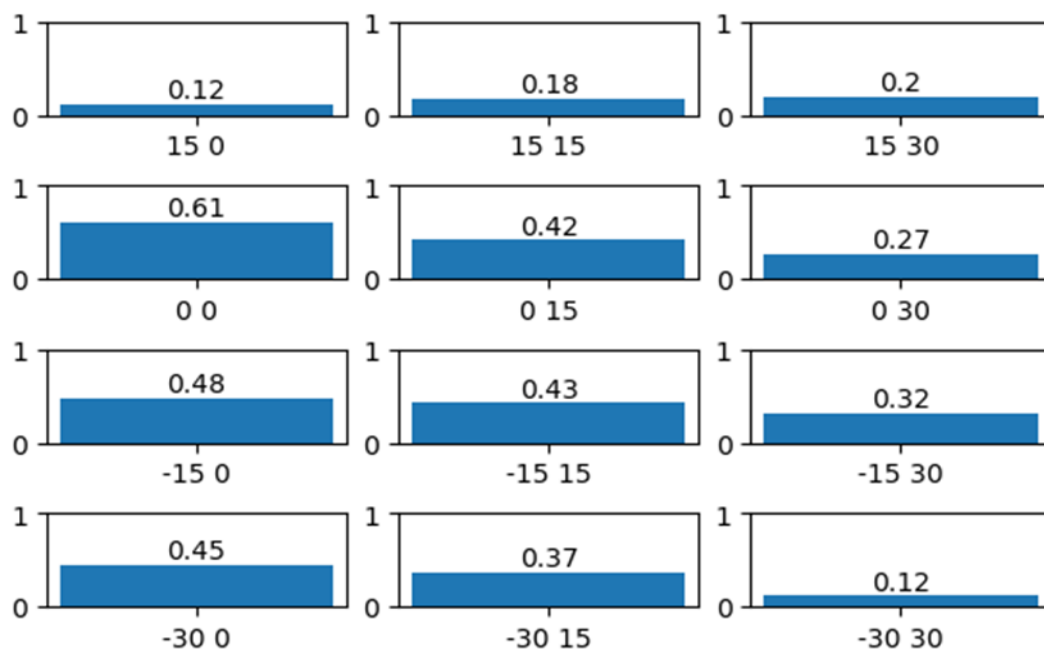


Рисунок 2. Результаты оценки точности для каждого из исследуемых углов поворота лица

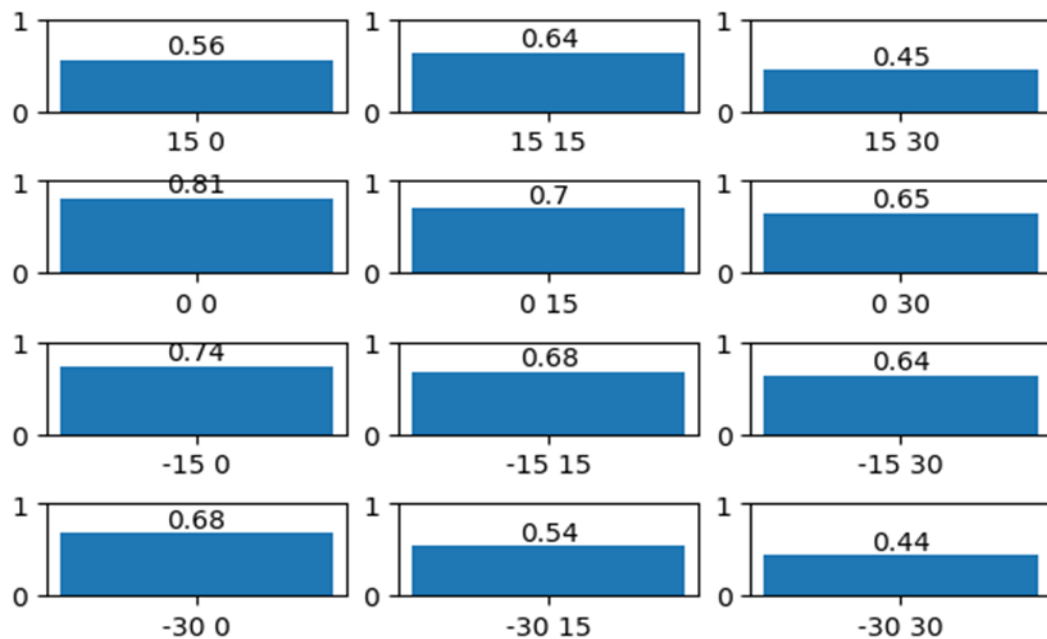


Рисунок 3. Результаты оценки относительности топ-1 для каждого из исследуемых углов поворота лица.

Принимая во внимание данные об оценке относительности топ-1, кратко представленные на рисунке 3, можно заметить, что в общем модель дает довольно высокую вероятность для истинной эмоции, но во многих случаях этого недостаточно для того, чтобы обойти эмоции, неправильно получившие более высокую вероятность. Таким образом, например, тенденция к уменьшению точности модели при поднятии человеком головы обретает новое обстоятельство – несмотря на очень низкую точность, модель выводила истинную эмоцию довольно высоко среди остальных. Это говорит о том, что при достаточной натренированности модели можно преодолеть эту проблему.

**Заключение.** Было проведено исследование эффективности работы алгоритма по распознаванию эмоций на различных углах поворота лица и были выделены тенденции эффективности работы: точность алгоритма по выбору эмоции по топ-1 падает при поднятии лица вверх, а также при больших поворотах лица в сторону; модель хуже выводит истинную эмоцию ближе к первому месту с большим поворотом лица от положения ан-фас.

## *Литература*

1. Mingjie Wang. Facial expression recognition based on CNN / Mingjie Wang - Vancouver, BC, Canada : Journal of Physics: Conference Series, 2020 - 15 p.
2. Paul Viola. Rapid Object Detection using a Boosted Cascade of Simple Features / Paul Viola, Michael Jones. - Cambridge, MA, USA : Accepted Conference on Computer Vision, 2001 - 9 p.
3. A. Landowska. Emotion recognition and its applications / A. Landowska, M. Szwoch, W. Szwoch, M.R. Wróbel, A. Kołakowska – Berlin, Germany : Springer International Publishing, 2014 – 13p
4. Mehdi S. M. Sajjad. Assessing Generative Models via Precision and Recall / Mehdi S. M. Sajjad, Olivier Bachem, Mario Lucic – Montreal, Canada : NeurIPS, 2018 – 10 p.