

# РАСПОЗНАВАНИЕ ПОЛА И ВОЗРАСТА В ТАРГЕТИРУЕМОЙ РЕКЛАМЕ

*Скавыш М.Г.*

*Институт информационных технологий БГУИР  
г. Минск, Республика Беларусь*

*Образцова О.Н. – канд. техн. наук, доцент*

В настоящее время системы обнаружения и распознавания лиц широко применяются для идентификации людей в биометрических системах безопасности или как дополнительная мера системы безопасности в местах большого скопления людей. Такие системы обеспечивают высокую точность распознавания, что позволяет повысить уровень безопасности какого-либо устройства или сервиса. Одной из наиболее применяемой технологией в таких системах являются нейронные сети. На этом применение сетей не заканчивается. Представьте, что заходя в магазин такая система подключена к камерам в магазине, и заходя в него система определяет ваш пол, возраст и эмоциональную реакцию от предложенной рекламы. В данной работе рассказывается как работает и обучается CNN сеть на загруженной выборке и применение обученной сети в режиме реального времени.

**Введение.** Искусственный интеллект – одна из самых популярных тенденций последнего времени. Машинное обучение и глубокое обучение составляют искусственный интеллект.

Диаграмма Венна (рисунок 1) объясняет взаимосвязь машинного обучения и глубокого обучения.

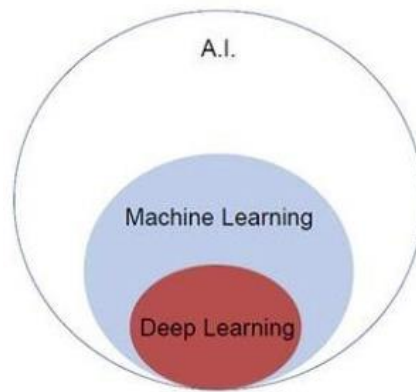


Рисунок 1 - Взаимосвязь машинного обучения и глубокого обучения

**Машинное обучение** – это наука о том, как заставить компьютеры действовать в соответствии с алгоритмами, разработанными и запрограммированными.

**Глубокое обучение** – это подполе машинного обучения, где соответствующие алгоритмы вдохновлены структурой и функциями мозга, называемыми искусственными нейронными сетями.

Вся ценность глубокого обучения сегодня заключается в контролируемом обучении или обучении с использованием маркированных данных и алгоритмов.

Каждый алгоритм глубокого обучения проходит один и тот же процесс. Он включает в себя иерархию нелинейного преобразования входных данных, которые можно использовать для создания статистической модели в качестве выходных данных.

**Материалы и методы.** В работе использовалась библиотека TensorFlow – открытая программная библиотека для машинного обучения, разработанная компанией Google для решения задач построения и тренировки нейронной сети с целью автоматического нахождения и классификации образов, достигая качества человеческого восприятия[1].

#### **Тензорная структура данных**

Тензорные элементы используются в качестве основных структур данных в языке TensorFlow. Тензорные элементы представляют соединительные ребра в любой блок-схеме, называемой графиком потока данных. Тензоры определяются как многомерный массив или список.

Тензоры обозначены следующими параметрами: ранг, форма, тип.

Единица размерности, описанная в тензоре, называется рангом. Он идентифицирует количество измерений тензора. Ранг тензора можно описать как порядок или n-размерность тензора.

Количество строк и столбцов вместе определяют форму Tensor.

Тип описывает тип данных, назначенный элементам Tensor.

TensorFlow включает в себя различные размеры

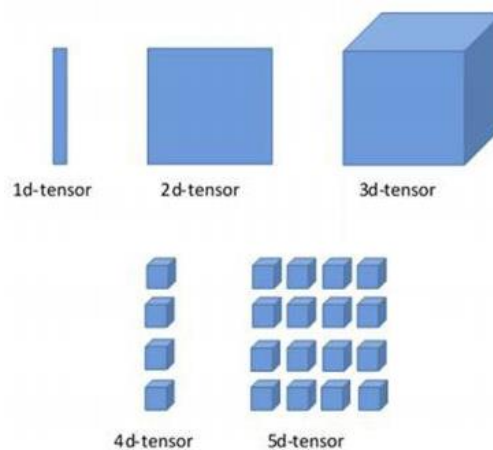


Рисунок 2 - Различные размеры TensorFlow

Сверточные нейронные сети предназначены для обработки данных через несколько уровней массивов. Этот тип нейронных сетей используется в таких приложениях, как распознавание изображений или распознавание лиц. Основное различие между CNN и любой другой обычной нейронной сетью состоит в том, что CNN принимает входные данные в виде двумерного массива и работает

непосредственно с изображениями, а не фокусируется на извлечении признаков, на котором сосредоточены другие нейронные сети.

Сверточная нейронная сеть использует три основных идеи:

- Местные соответствующие поля
- Свертка
- Объединение

Свертка состоит из ядра представляющего из себя матрицу весов. Ядро скользит над двумерным изображением, поэлементно выполняя операцию умножения с той частью входных данных, над которой оно сейчас находится, и затем суммирует все полученные значения в один выходной пиксель. Ядро повторяет эту процедуру с каждой локацией, над которой оно скользит, преобразуя двумерную матрицу в другую все еще двумерную матрицу признаков. Признаки на выходе являются взвешенными суммами (где веса являются значениями самого ядра) признаков на входе, расположенных примерно в том же месте, что и выходной пиксель на входном слое.

В примере, приведенном на рисунке 3, мы имеем  $5 \cdot 5 = 25$  признаков на входе и  $3 \cdot 3 = 9$  признаков на выходе. Для стандартного слоя (standard fully connected layer) мы бы имели весовую матрицу  $25 \cdot 9 = 225$  параметров, а каждый выходной признак являлся бы взвешенной суммой всех признаков на входе. Свертка позволяет произвести такую операцию с всего 9-ю параметрами, ведь каждый признак на выходе получается анализом не каждого признака на входе, а только одного входного, находящегося в примерно том же месте.

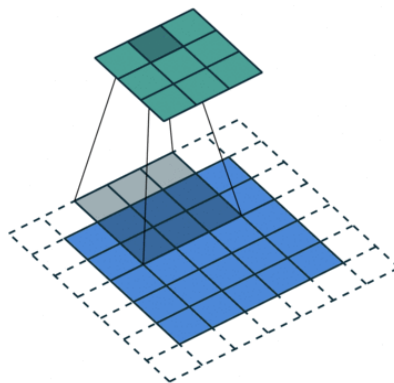


Рисунок 3 - Двумерная свертка

Часто бывает, что при работе со сверточным слоем, нужно получить выходные данные меньшего размера, чем входные. Это обычно необходимо в сверточных нейронных сетях, где размер пространственных размеров уменьшается при увеличении количества каналов. Один из способов достижения этого – использование субдискретизирующих слоев (pooling layer), например, принимать среднее/максимальное значение каждой ветки размером  $2 \times 2$ , чтобы уменьшить все пространственные размеры в два раза [3].

Каждый нейрон на входе связан с каждым нейроном на выходе, т.е. полностью связан (рисунок 4). Благодаря этой связи каждый нейрон на выходе будет использоваться не более одного раза.  $\sum_{i=0}^n xW + b$

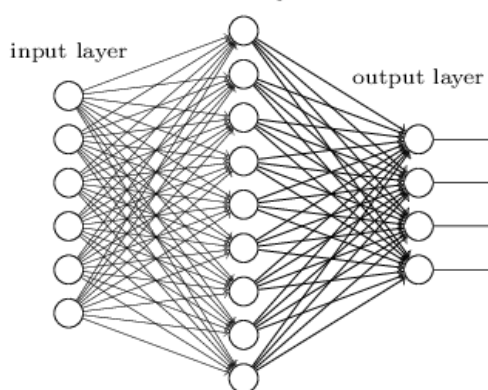


Рисунок 4 - Полностью связанный слой

Принцип работы CNN сети представлен на рисунке 5. Передавая входные данные через свертку, вы извлекаете пространственные объекты высокой размерности. Объединение обобщает пространственную информацию и уменьшает размерность. Затем это представление объектов передается через полностью связанные слои классификатору или регрессору.

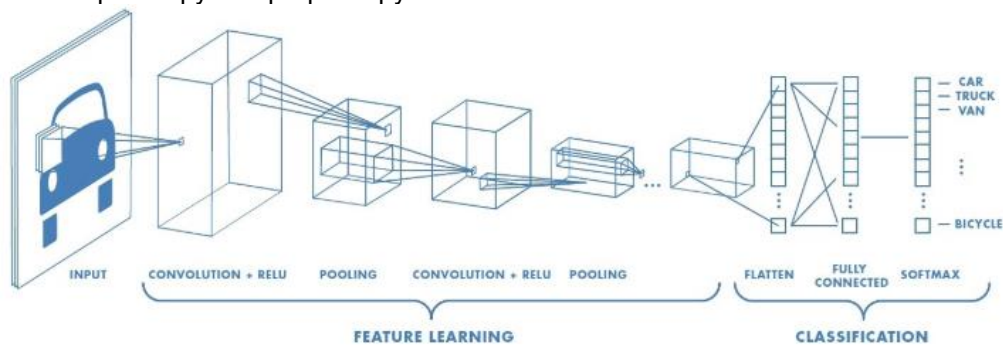


Рисунок 5 - Принцип работы CNN сети

**Результаты.** На рисунке 6 проиллюстрированы результаты работы программы. Модель была обучена с помощью набора данных WIDER FACE – это эталонный набор данных для распознавания лиц, изображения которого выбираются из общедоступного более широкого набора данных. Данная выборка характеризуется высокой степенью вариабельности [4].

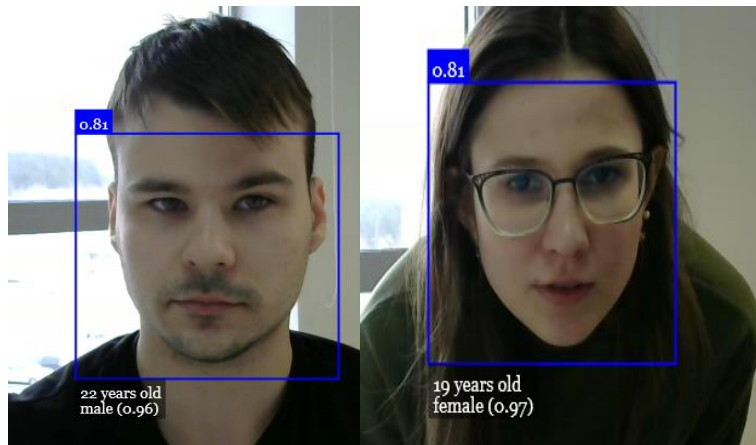


Рисунок 5 - Определение пола и возраста в видеопотоке

Данные полученные можно передать на сторонний сервис, который будет передавать необходимую рекламу в соответствии полу и возрасту клиента.

**Заключение.** Сверточные нейронные сети позволяют компьютерному зрению работать как с простыми задачами, так и со сложными продуктами и услугами, начиная от распознавания лиц и заканчивая улучшением медицинских диагнозов. Но на этом их спектр не заканчивается, как видно из предложенного выше их можно применить практически во всех сферах деятельности человека начиная с медицины и заканчивая сферами развлекательных услуг.

**Список использованных источников:**

3. Basics of machine learning with TensorFlow [Электронный ресурс] <https://www.tensorflow.org/resources/learn-ml>.
4. Как работает сверточная нейронная сеть: архитектура, примеры, особенности [Электронный ресурс] <https://neurohive.io/ru/osnovy-data-science/glubokaya-svertochnaja-nejronnaja-set/>.
5. Deep Learning CNN's in Tensorflow [Электронный ресурс] <https://hackernoon.com/deep-learning-cnns-in-tensorflow-with-gpus-cba9ef0acc2>
6. WIDER FACE: A Face Detection Benchmark [Электронный ресурс] <http://shuoyang1213.me/WIDERFACE/>