|  |  |
| --- | --- |
| **СОГЛАСОВАНО**  Доцент кафедры  ИАНИ ННГУ, к.ф.-м.н.  \_\_\_\_\_\_\_\_\_\_\_\_\_ Д.А. Яшунин  «\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_2019 г. | **УТВЕРЖДАЮ**  Профессор кафедры  ИАНИ ННГУ, д.т.н.  \_\_\_\_\_\_\_\_\_\_\_\_\_ Н.В. Старостин  «\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_2019 г. |

**Пояснительная записка № 1**

**«Метод распознавания лиц на фотографии с использованием и OpenVINO Deep Learning Framework и SphereFace»**

**Этап 2. Подготовка обзора выбранных подходов к решению   
задачи и построение технологического стека**

**НИР «Разработка и реализация программного обеспечения   
для распознавания лиц на групповых фотографиях»**

**(Шифр ПО «FaceRecognition»)**

Ответственный исполнитель

\_\_\_\_\_\_\_\_\_\_\_\_ A.C. Филатова

«\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_2019 г.

**Н. Новгород 2019**

В пояснительной записке в рамках этапа обзора выбранных подходов к решению задачи «FaceRecognition» и построения технологического стека представлено описание фреймворка SphereFace для распознавания лиц; описание функции SoftMax, которая используется в классификаторах при тренировки нейронных сетей, а также ее модификация в рамках метода SphereFace; обобщенный алгоритм тренировки сети.

Оглавление

[1 Описание фреймворка SphereFace 4](#_Toc4447200)

[2. Математическое обоснование функции потерь 4](#_Toc4447201)

[2.1. Применение функции SoftMax в задаче классификации 4](#_Toc4447202)

[2.2. Пересмотр функции SoftMax loss 5](#_Toc4447203)

[2.3. Введение углового запаса в SoftMac loss 6](#_Toc4447204)

[3. Обучение и тренировка 8](#_Toc4447205)

[Список используемых источников 9](#_Toc4447206)

# Описание фреймворка SphereFace

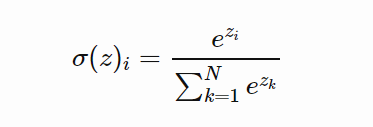
SphereFace – фреймворк для глубокого распознавания лиц. В общем случае пайплайн содержит три этапа - обнаружение, выравнивание и распознавание лиц. Для нашей дальнейшей работы нас будет интересовать только этап распознавания лиц.

Данный метод основан на сопоставлении вектора в пространстве гиперсферы объекту (в нашем случае – изображению лица).

Этот метод предлагает функцию A-Softmax (angular softmax), которая позволяет сверточным нейронным сетям исследовать угловые отличительные особенности. Геометрически функция A-Softmax может рассматриваться как наложение дискриминационных ограничений на гиперсферное многообразие, которое по сути совпадает с тем, что изображения также лежат на многообразии.

# Математическое обоснование функции потерь

## Применение функции SoftMax в задаче классификации

Функция SoftMax, функция мягкого максимума, часто используется в нейронных сетях в качестве функции активации при решении задачи классификации. SoftMax задается следующей формулой:

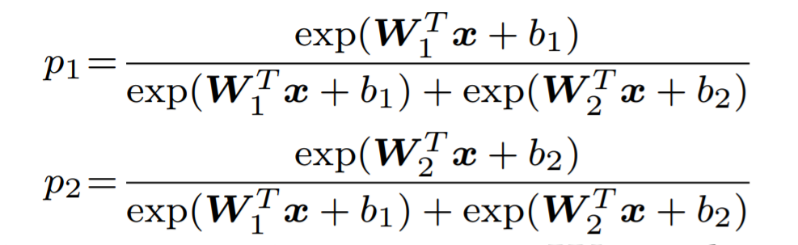
где zi – значение на выходе из i-го нейрона до активации, а N – общее количество нейронов в слое.

Эта функция очень часто используется для задач классификации.

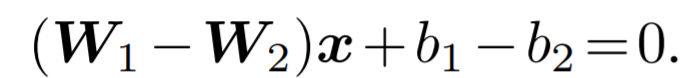
Рассмотрим задачу классификации, в которой объект принадлежит только одному классу. Значения на выходе из слоя могут трактоваться как вероятность принадлежности объекта заданному классу. Это означает, что каждое значение должно быть в диапазоне от 0 до 1, и сумма всех значений должна равняться единице. Также функция должна быть дифференцируема, чтобы можно было применять градиентные методы обучения нейронной сети.

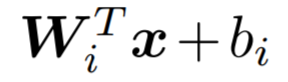
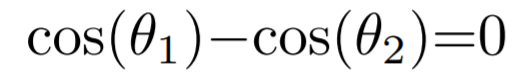
Достоинство функции SoftMax в том, что она использует экспоненту. За счет этого решается проблема с отрицательными выходными значениями, т.к. экспонента всегда положительна. Кроме того, экспонента значительно увеличивает большие значения. Если на выходе из нейрона 1, то экспонента равна 2.7, если 2 – экспонента равна 4.7, а если на выходе 4, то экспонента получается 54.6. Именно поэтому функции SoftMax удобно применять для задач классификации, т.к. она позволяет трактовать выходные значения нейронов как вероятность принадлежности данному классу, а также обеспечивает, чтобы только одно выходное значение было близко к единице за счет применения экспоненты.

## Пересмотр функции SoftMax loss

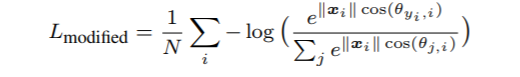
В случае двух классов апостериорные вероятности, полученные при Softmax loss равны:

где х - вектор изученных признаков, Wi и bi - веса и смещения последнего полностью связного слоя, соответствующего классу i, соответственно.

Предсказанная метка будет присвоена первому классу, если p1> p2, и второму классу, если p1<p2. Тогда функция, отвечающая за разделение между двумя классами будет иметь вид: 

можно представить в виде  , где θi - это угол между Wi и х. Нормализуем веса и занулим смещения слоев, тогда границей решения становится (угловая биссектриса вектора W1 иW2)

Хотя приведенный выше анализ построен на основе двоичной классификации, обобщить анализ для случая нескольких классов тривиально.

Функция Softmax loss может быть записана как:

где θj, i (0≤θj, i ≤π) - угол между вектором Wj и xi.

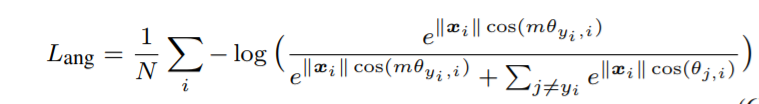
## Введение углового запаса в SoftMac loss

Из предыдущего анализа потерь Softmax известно, что границы принятия решений могут сильно повлиять на распределение признаков, поэтому основная идея состоит в том, чтобы манипулировать границами принятия решений, чтобы получить угловой запас.

Пусть x – вектор признаков из первого класса, тогда чтобы правильно классифицировать х необходимо, чтобы .

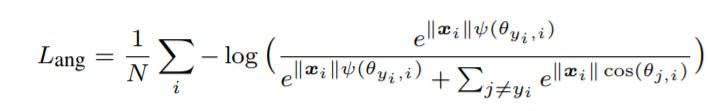
Если вместо этого потребовать , то это сделает решение более строгим. Границей решения будет являться .

Таким образом измененная функция Softmax loss будет иметь следующий вид:



, где (θyi, i) должен находиться в диапазоне [0,π/m].

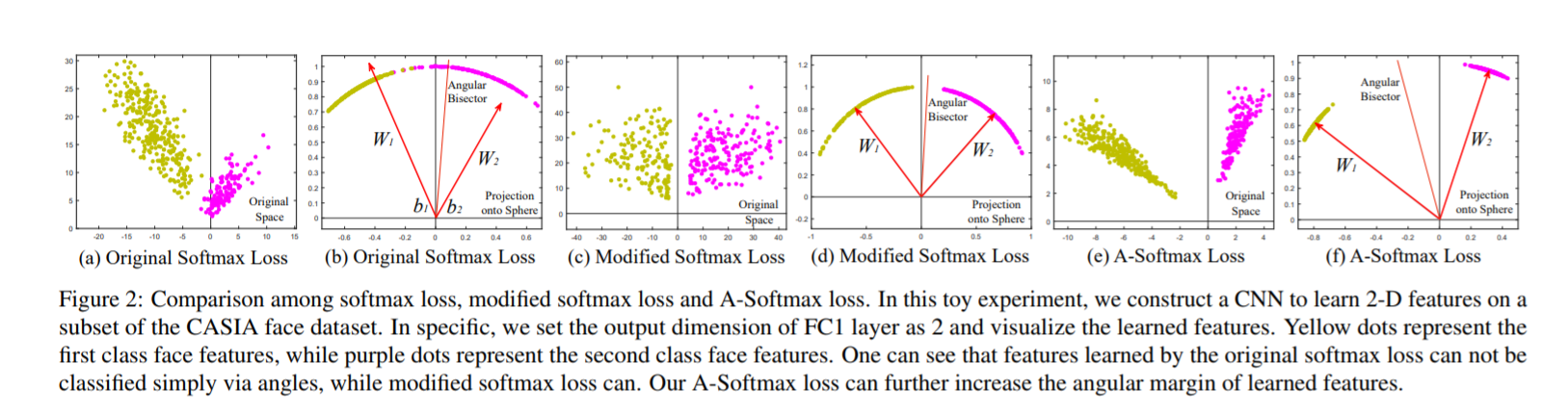
Чтобы избавиться от этого ограничения и сделать его оптимизируемым в CNN, расширяется диапазон определения cos (θyi, i), обобщив его на монотонно убывающую функцию угла ψ (θyi, i), которая должна быть равна cos (θyi , i) в [0,π/m]. Итоговая функция A-Softmax loss имеет вид:

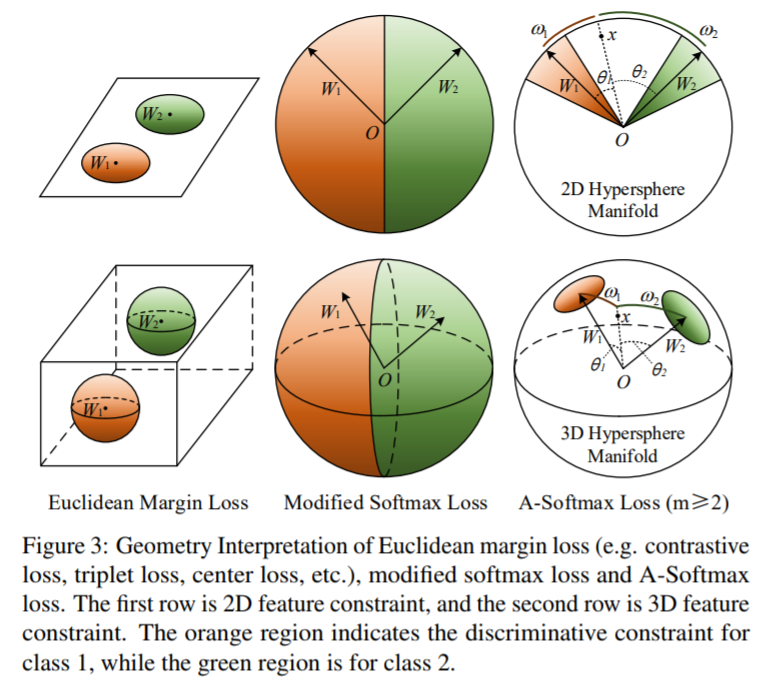


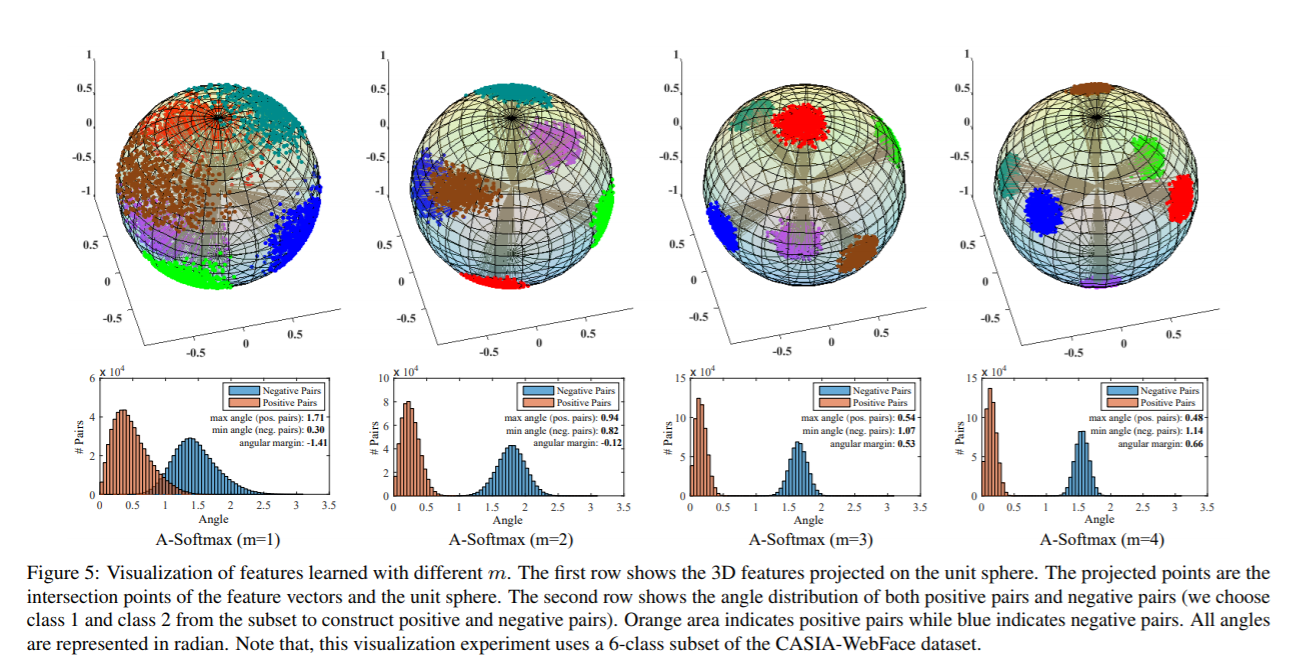
, где  , 

Следовательно, A-Softmax Loss - это угловое расстояние между категориями выборки, что делает функцию принятия решений более строгой и более различимой. С увеличением m угловое расстояние также увеличивается.

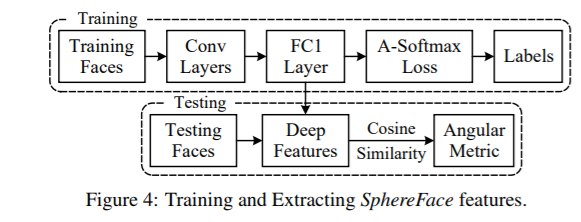
Контролируемые функцией A-Softmax, CNN изучают черты лица с геометрически интерпретируемым угловым запасом. Поскольку потери ASoftmax требуют Wi = 1, bi = 0, классификация зависит только от углов между выборкой x и Wi. Так что х можно отнести к тождеству с наименьшим углом. Параметр m добавляется с целью изучения углового поля между различными идентификаторами.







# Обучение и тренировка

Caffe используется для реализации функции потерь A-Softmax и CNN. Общая структура для обучения и извлечения функций SphereFace показана на рисунке:

Глубокие векторы признаков извлекаются из вывода слоя FC1. Для всех экспериментов окончательное представление тестируемого лица получается путем объединения ее исходных векторов признаков и горизонтальных отраженных элементов. Оценка (метрика) вычисляется по косинусному расстоянию двух признаков. Классификатор ближайшего соседа и пороговое значение используются для идентификации и проверки лица, соответственно.

# Список используемых источников

1. Техническое задание на НИР «Разработка и реализация программного обеспечения для распознавания лиц на групповых фотографиях» (Шифр ПО «FaceRecognition»)
2. <https://arxiv.org/pdf/1704.08063.pdf>
3. <https://github.com/wy1iu/sphereface>
4. K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In CVPR, 2016
5. https://www.rankred.com/face-recognition-algorithms-techniques/