

Прикладные задачи компьютерного зрения

Распознавание и поиск похожих объектов

Поиск похожих объектов в базе




Аутентификация




Face ID




Findface



PRODUCTS INDUSTRIES SOLUTIONS PARTNERS PRESS DOCS CONTACTS

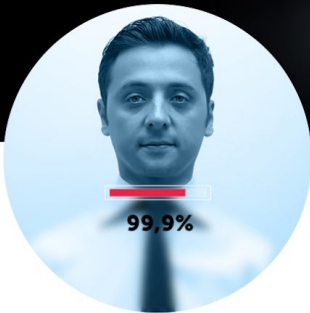




Face Detection

40 pixels Eye Distance


Reliably detect human faces in a photo, along with their coordinates and sizes.



Verification (1:1)

More than 99% Accuracy

Compare one face with another to identify if it is the same person or not.



Identification (1:N)

Store billions of people in the gallery.

Find a person by face in a pre-loaded collection of faces.

<https://findface.pro/en/>

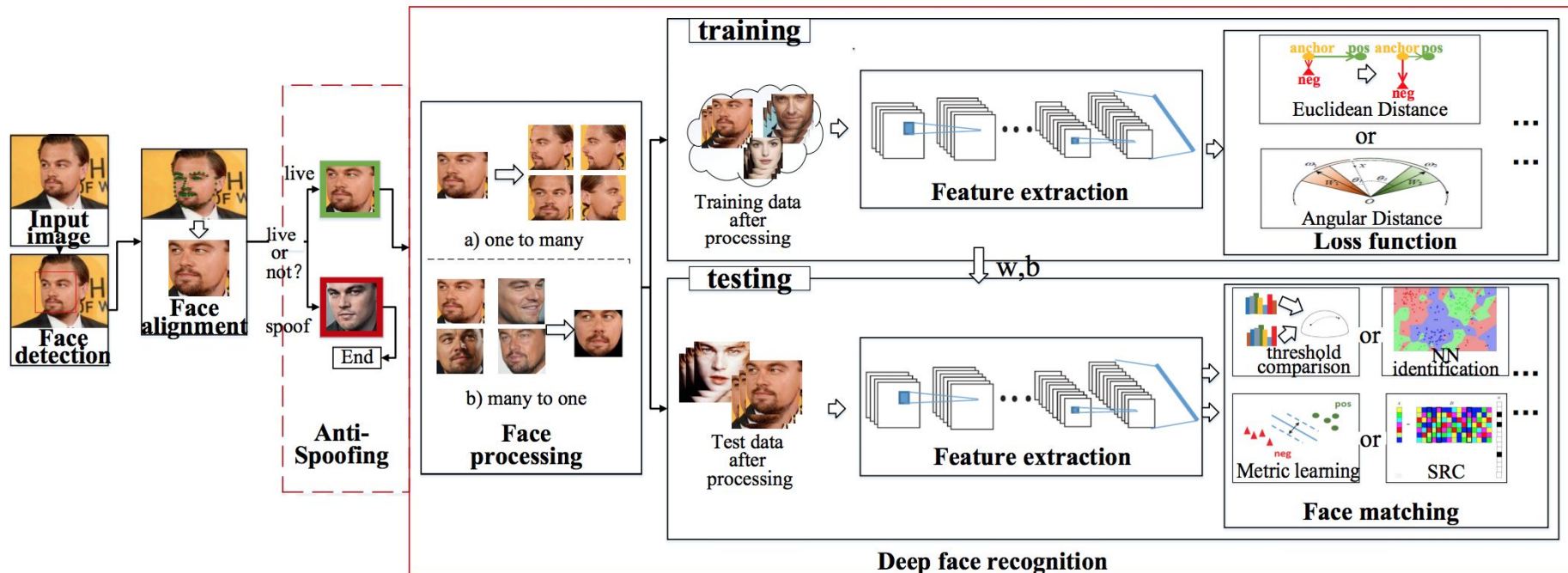
О чем пойдет речь

1. Особенности задачи и варианты решения
2. Дескрипторы (признаки) изображения
3. Сиамские сети
4. Функция потерь и процесс построения модели

Особенности задачи

- Ракурс наблюдения объекта может варьироваться
- Изменение освещенности и окружения
- Большой размер датасета
- Не все объекты могут быть доступны во время обучения

Решение задачи



Решение задачи

- Свести изображения в признаковое пространство меньшей размерности
- Такое представление должно быть *дискриминативно*, т.е. позволять различать положительные и отрицательные примеры
- При построении модели необходимо учитывать как положительные, так и отрицательные примеры

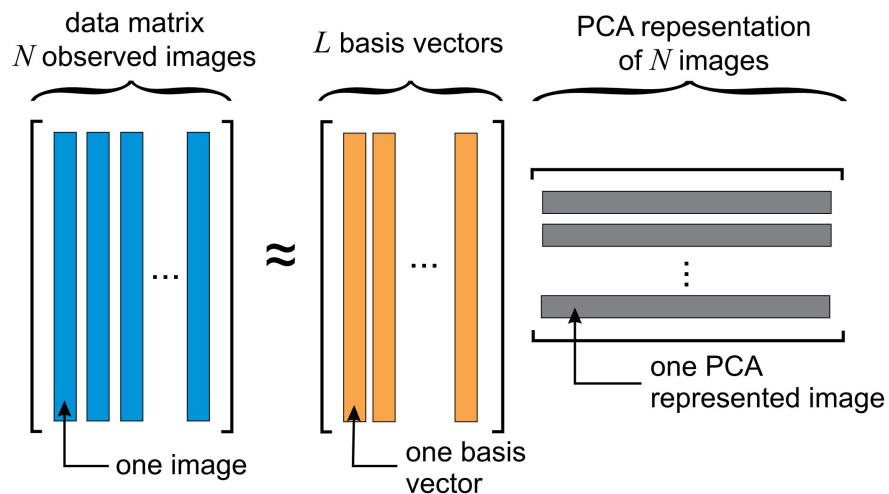
Дескриптор (признаки) изображения

РСА - анализ главных компонент (Eigenface)

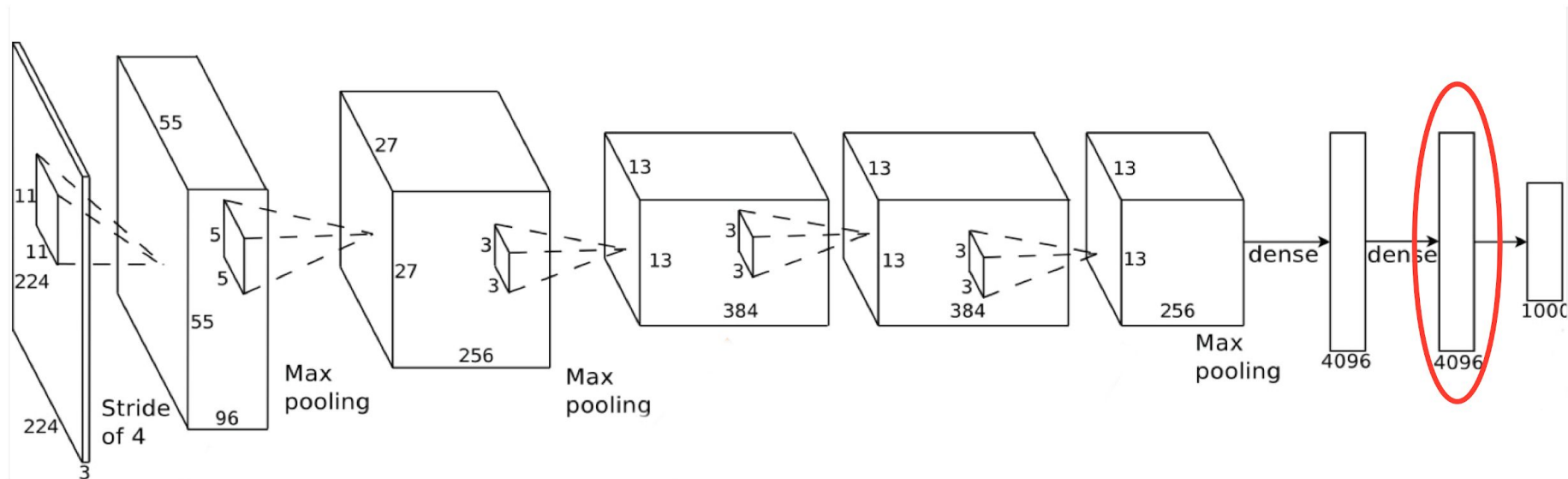


<https://en.wikipedia.org/wiki/Eigenface>

PCA - анализ главных компонент (Eigenface)

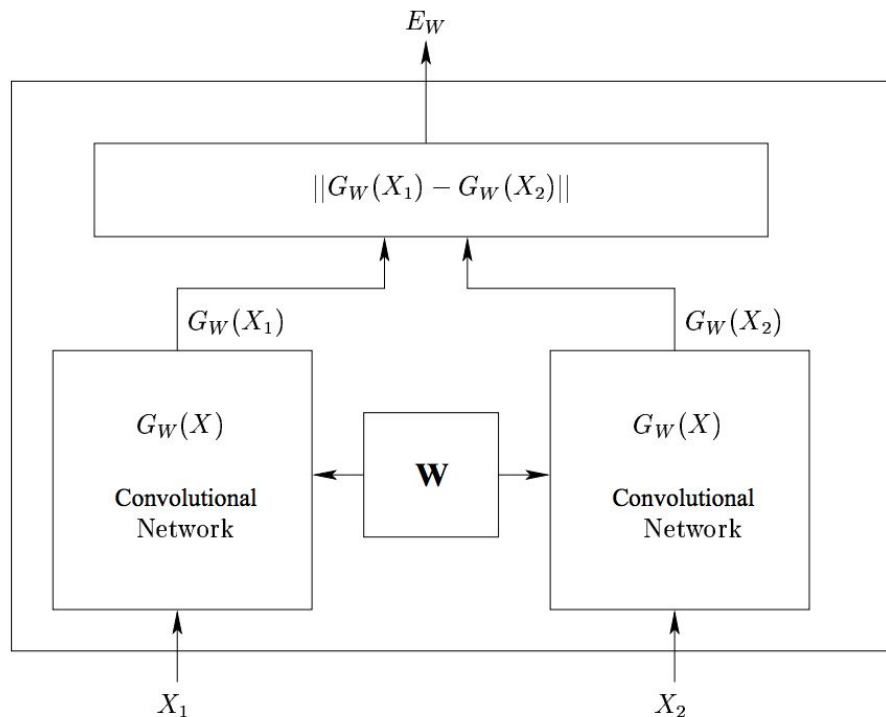


Выделение признаков на основе нейронных сетей



Сиамские нейронные сети

Сиамские нейронные сети



Архитектура сиамской сети. Параметры \mathbf{W} одинаковы во входных блоках.

Сиамские нейронные сети



Визуализация активаций сверточной сиамской сети

Metric Learning

Metric Learning

- Softmax
- Contrastive Loss
- Triplet Loss
- Cosine Similarity

Contrastive Loss

$$\mathcal{L} = y_{ij} \max(0, \|f(x_i) - f(x_j)\|_2 - \epsilon^+) + (1 - y_{ij}) \max(0, \epsilon^- - \|f(x_i) - f(x_j)\|_2)$$

- Рассматриваем пары изображений i и j
- $y_{ij} = 1$, если на изображениях один объект и $y_{ij} = 0$, если объекты различаются
- $f(x_i)$ - функция получения дескриптора (признаков) изображения
- ϵ^+ и ϵ^- - требуемые отступы в расстояниях для положительных и отрицательных примеров

Triplet Loss

$$\|f(x_i^a) - f(x_i^p)\|_2^2 + \alpha < \|f(x_i^a) - f(x_i^n)\|_2^2$$

- Выбирается якорный пример x^a
- Для якоря подбираются положительный x^p и негативный пример x^n
- В процессе обучения предпочтение отдается сложным негативным примерам
- Функция потерь требует, чтобы расстояние между положительным примером было меньше расстояния между отрицательным примером на величину α или больше

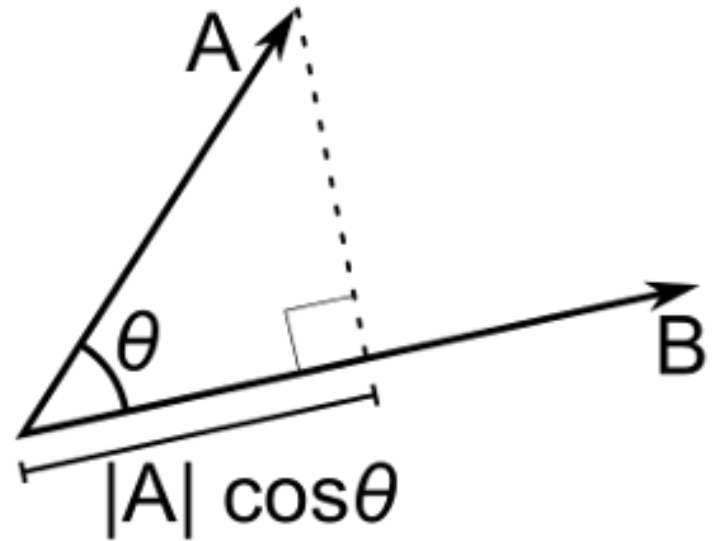
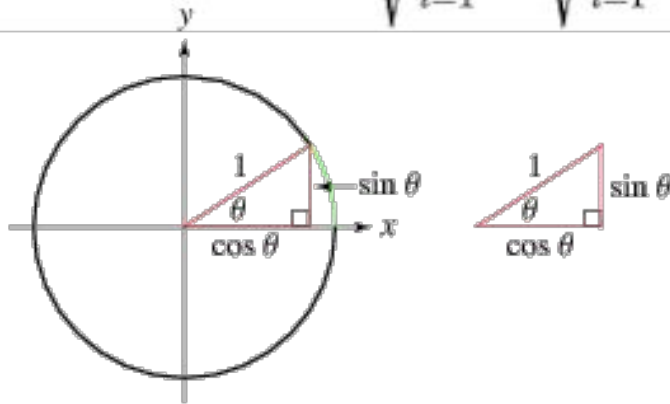
Center Loss

$$\mathcal{L}_C = \frac{1}{2} \sum_{i=1}^m \|x_i - c_{y_i}\|_2^2$$

- Минимизация расстояния до центра кластера
- Уменьшает вариативность расстояний внутри кластера

Cosine Similarity

$$\cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$



L-Softmax

$$\|W_1\| \|x\| \cos(m\theta_1) > \|W_2\| \|x\| \cos(\theta_2)$$

- θ_1 - угол между положительным примером
- θ_2 - угол между отрицательным примером
- m - требуемый марджин (отступ)
- W_1, W_2 - веса полносвязного слоя для класса 1 и 2 соответственно
- x - признаки объекта

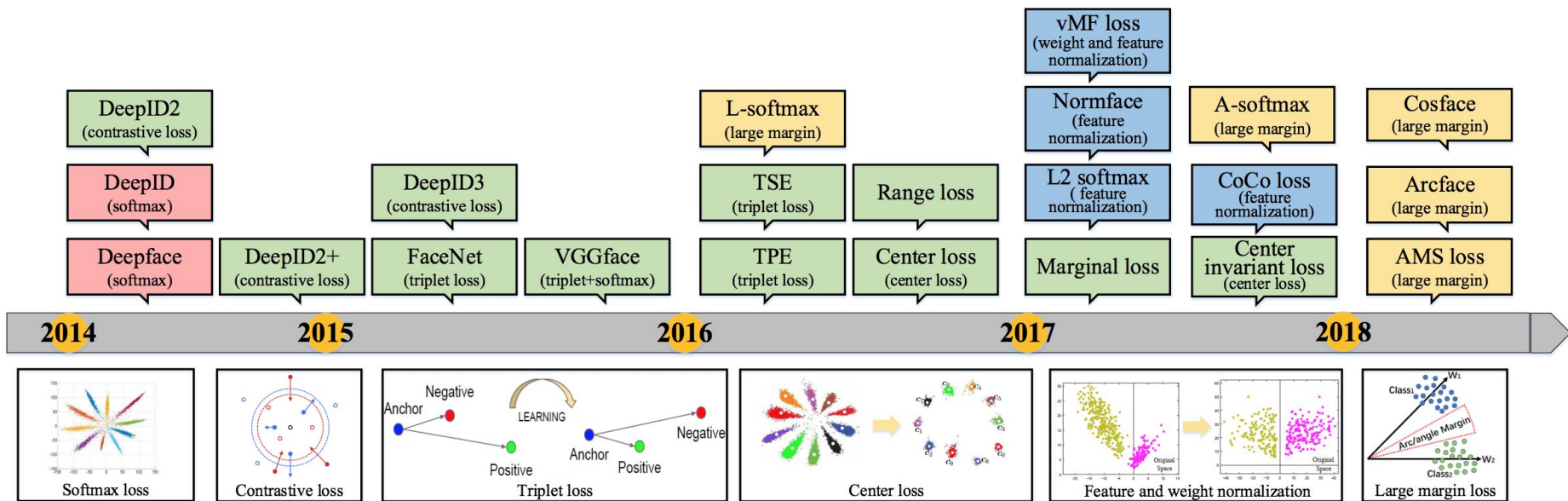
L-Softmax

$$\mathcal{L}_i = -\log \left(\frac{e^{\|W_{yi}\| \|x_i\| \varphi(\theta_{yi})}}{e^{\|W_{yi}\| \|x_i\| \varphi(\theta_{yi})} + \sum_{j \neq y_i} e^{\|W_{yi}\| \|x_i\| \cos(\theta_j)}} \right)$$

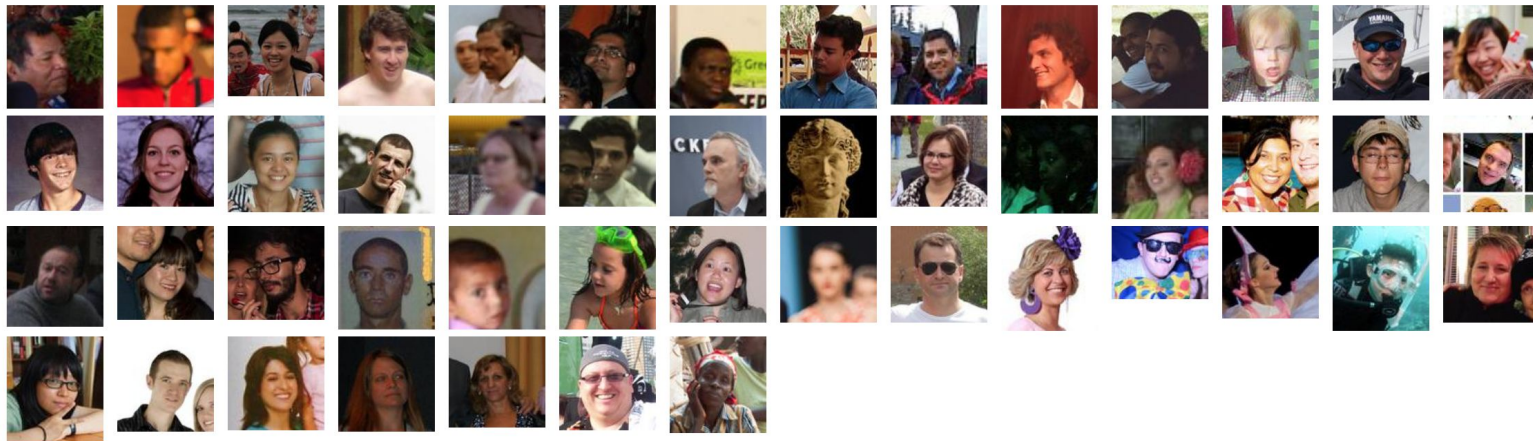
$$\varphi(\theta) = (-1)^k \cos(m\theta) - 2k, \theta \in \left[\frac{k\pi}{m}, \frac{(k+1)\pi}{m} \right]$$

- Тк нас интересует диапазон углов $0..\pi$, то необходимо корректно учитывать смену знака $\cos(m*\theta)$

История развития подходов



MegaFace Dataset



Distractors

1 Million Photos

690,572 Unique Users

Training Set

4.7 Million Photos

672,057 Unique Identities

7 Mean photos / person (3 min, 2469 max)

Test Sets

FaceScrub Celebrities

FGNet Age-invariant non-celebrities

<http://megaface.cs.washington.edu>

Другие датасеты по распознаванию лиц

- [AR Face Database](#)
- [Labeled Faces in the Wild](#)
- [The MUCT Face Database](#)
- [Yale Face Database](#)

Полезные ссылки

- [Learning a Similarity Metric Discriminatively, with Application to Face Verification](#)
- [Dimensionality Reduction by Learning an Invariant Mapping](#)
- [Deep Face Recognition: A Survey](#)
- [FaceNet: A Unified Embedding for Face Recognition and Clustering](#)