Real time planetary scale face recognition system
Agenda

- Face recognition market;
- Our algorithm;
- MegaFace challenge;
- FindFace — large scale search engine;
- Accuracy and timing at scale;
- Real-life applications.
The global facial recognition market

Source: Technavio; Stratistics MRC; Marketsandmarkets

The market embraces software, hardware and relevant services (consulting, cloud computing, etc.)
# Applications of face recognition

<table>
<thead>
<tr>
<th>MAJOR APPLICATIONS</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GOVERNMENTAL</strong></td>
<td></td>
</tr>
<tr>
<td>Face recognition in a crowd</td>
<td>Identification of blacklisted people in airports and other public places</td>
</tr>
<tr>
<td>Road safety</td>
<td>Detection of traffic violations (e.g. bikers not wearing helmets, etc.)</td>
</tr>
<tr>
<td>Identification of suspects</td>
<td>Identification/Identity verification of suspects in police stations</td>
</tr>
<tr>
<td><strong>COMMERCIAL</strong></td>
<td></td>
</tr>
</tbody>
</table>
| Night clubs, casinos | Identification of blacklisted people  
 | | Recognition of visitors’ mood, average age and gender ratio to be displayed to prospective visitors |
| Shopping | Advising products/services (food, clothing, etc.) based on specific characteristics |
| Financial sector* | Advising relevant services (banking)  
 | | Customer identity verification  
 | | Payer identity verification and assessment of specific customer parameters |
| Dating services | Searching for matches with pre-specified parameters/looking like someone else; photo-based profile search |
| Businesses | Identification of employees and time tracking software |
| Other | Driver drowsiness detection systems, automatic face sorting in Disneyland photos, etc. |
The advantages of the algorithm

- Best result in the Megaface contest — better than Google’s FaceNet;
- Proven efficiency on very large databases;
- 73% rank-1 on 1M dataset;
- 70+% rank-10 accuracy on 300M+ photos from the vk.com social network;
- Extremely low computational requirements (person’s features vector takes less than 1 Kb);
- Query time less than a second (based on 300M+ photos).
Face Recognition Pipeline

Face representation:

Detector → Normalizer → Feature extractor

Scenarios:

Identification

Verification
Deep Learning

Image recognition

ILSVRC top-5 error on ImageNet

* Source: devblogs.nvidia.com

Speech recognition

* Source: Microsoft’s speech group

Image captioning


Text analysis

Source: T.Mikolov. Efficient Estimation of Word Representations in Vector Space

AlexNet architecture

![AlexNet Architecture Diagram]

Loss function

- Multinomial logistic regression
  \[ J(\theta) = -\frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{c} y^{(i)}_j \log \frac{\exp(\theta_j^T x^{(i)})}{\sum_{k=1}^{c} \exp(\theta_k^T x^{(i)})} \]

- Triplet loss
  \[ \sum_{i=1}^{N} \left[ \| f(x_i^a) - f(x_i^p) \|^2 - \| f(x_i^a) - f(x_i^n) \|^2 + \alpha \right]_+ \]

- Contrastive loss
  \[ \frac{1}{2} \| f_i - f_j \|^2 \quad \text{if } y_{ij} = 1 \]
  \[ \frac{1}{2} \max(0, m - \| f_i - f_j \|^2)^2 \quad \text{if } y_{ij} = -1 \]

- 20M face photos for training;
- 3 weeks training on 3 GPUs NVidia Titan Black.

AlexNet architecture

- Robustness to different shooting conditions (perspective, age, emotions so on..);
- Compact face representation (up to 16 floats);
- Ability to reusing for training additional classifier (gender, race, ...).
LFW

- Images collected from the web;
- 13K photos, 5K people;
- The only constraint on faces is that they were detected by the Viola-Jones detector;
- A number of algorithms achieve near to perfect accuracy;
- Need some bigger dataset!

Example errors:

**False negative:**

**False positive:**
Megaface challenge

- Autumn 2015
- 1M identities
- Identification and Verification scenarios
- More than 100 teams participated
Identification Scenario

**FaceScrub**
- 80x50 photos
  - Identity 1
  - Identity 2
  - Identity 3
  - Identity 80

**MegaFace**
- 1M unlabeled photos

80 identities

**MegaFace + 1**
- search photo 1
- search photo 2
- search photo 3
- search photo 50
Identification Results

![Graph showing identification rates for different algorithms](image)

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Set 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTechLAB - facex_large</td>
<td>73.300%</td>
</tr>
<tr>
<td>Google - FaceNet v8</td>
<td>70.496%</td>
</tr>
<tr>
<td>Beijing Faceall Co. - FaceAll_Norm_1600</td>
<td>64.803%</td>
</tr>
<tr>
<td>Beijing Faceall Co. - FaceAll_1600</td>
<td>63.977%</td>
</tr>
<tr>
<td>Barebones_FR - cnn</td>
<td>59.363%</td>
</tr>
<tr>
<td>NTechLAB - facex_small</td>
<td>58.218%</td>
</tr>
<tr>
<td>3DIVI Company - tdvm6</td>
<td>33.705%</td>
</tr>
<tr>
<td>Joint Bayes</td>
<td>3.021%</td>
</tr>
<tr>
<td>LBP</td>
<td>2.326%</td>
</tr>
</tbody>
</table>

---

*uses large training set*
Neural Net vs Human

Comparable to human abilities on small datasets;

Scales much better and outperforms human in a large scale recognition problem.
FindFace.ru

- Search among profile photos in vk.com - largest Russian social network
- 250M photos in index
- 90M people in index
- Search time 0.5 s
- 50 RPS on 5 AWS machines.
Accuracy at scale

Rank-1 accuracy

<table>
<thead>
<tr>
<th>Number of photos</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 M</td>
<td>73 %</td>
</tr>
<tr>
<td>250 M</td>
<td>60 %</td>
</tr>
</tbody>
</table>

Graph showing the change in rank-1 accuracy with the number of photos and number of distractors.
## Timings

<table>
<thead>
<tr>
<th></th>
<th><strong>GPU</strong></th>
<th></th>
<th><strong>CPU</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nvidia GTX TITAN Black</td>
<td>514 hours x 3 GPUs</td>
<td>Intel Core i7-5930K</td>
<td></td>
</tr>
<tr>
<td>Neural Net Training</td>
<td>-</td>
<td>150 ms</td>
<td>-</td>
<td>150 ms</td>
</tr>
<tr>
<td>Face Detection</td>
<td>-</td>
<td>143 ms</td>
<td>-</td>
<td>130 ms</td>
</tr>
<tr>
<td>Feature extraction</td>
<td>8.96 ms</td>
<td></td>
<td>143 ms</td>
<td></td>
</tr>
<tr>
<td>Search time</td>
<td>-</td>
<td>130 ms</td>
<td>-</td>
<td>130 ms</td>
</tr>
</tbody>
</table>
FindFace in real-life

Russian police use findface.ru for searching criminals. We’ve got a lot of emails from them about their experience. E.g. police from Udmurt Republic in more than 50% of cold cases found suspect by photo in vk.com.

Findface.ru increases the percentage of solved crimes.
FindFace in real-life

In St. Petersburg: two teenagers decide to fire a newly built house.

Their faces were filmed on the hidden camera in the elevator.

After the video had appeared in the Internet, people quickly found accounts of these hooligans in the largest Russian social network VK.com with all additional information: where they live, what school they're attending etc.

All information was transferred to the police.
Robustness to occlusions

- 86.25%
- 75.65%
- 62.80%
- 44.00%
- 30.00%
Age and Gender

Gender recognition

- SVM classifier above feature vectors
- 99.5% accuracy
FindFace.pro - b2b cloud platform

- Upload your dataset up to 1 billion photos;
- Identification and verification scenarios;
- Gender, Age, Emotion;
- Scale to any number RPS you need;
- Extend to 5B people dataset.
Interaction pattern for enterprise

Cameras
Data acquisition

Our software
SDK

Customer’s servers

Storing data in customer’s Private Cloud.
Contacts

Artem Kukharenko

kukharenko@ntechlab.ru

+7 (499) 110-22-54