VISIONLABS

Cross-Domain Face Recognition Solution Based On GPU-Powered Deep Learning and Inference

Alexander Khanin, Founder & CEO | Talk ID: DC7127
VisionLabs

About:
VisionLabs is a team of Computer Vision and Machine Learning experts. Founded in 2012, today we offer face recognition products and custom Computer Vision solutions globally.

Products:
The flagship products of VisionLabs are LUNA SDK face recognition engine, LUNA PLATFORM biometric data management system and LUNA CLOUD face recognition service. We also focus on AR/VR field with Face.DJ product for creating the photorealistic 3D face model from a single selfie.

Customers:
We are proud to serve the industry leaders and SMBs.

Collaboration:
We are happy to work together with the global technology leaders.
Outstanding facts

Our face recognition algorithm was independently tested by NIST as a part of the ongoing FRVT and is TOP rated for accuracy and speed.

We are the only company in the TOP 5 according to the LFW test result having commercially available algorithm with the same performance characteristics.

More than 500,000 IP cameras are streaming video to our products daily with the largest single Customer implementation of 76,000+ IP cameras.

Our products process more than 58B face recognition requests per year globally with the largest single Customer implementation of 27B requests per year.
LUNA SDK is a face recognition engine developed by VisionLabs. It enables efficient and accurate processing of faces in images and live video streams.

LUNA’s face recognition pipeline includes several key modules: face detection, face alignment, facial descriptor extraction, face matching, facial attribute classification and face spoofing prevention.

LUNA SDK is entirely developed in C++ and enables

- Optimized memory handling
- Extensive multithreading
- High performance of all modules
- Cross-platform integration
- Standard C++ API
LUNA PLATFORM is a biometric data management system for facial verification and identification. The platform offers a great flexibility to create scenarios of varying complexity for integrated facial recognition.

LUNA PLATFORM offers RESTful Web API to its clients for command transfer and data communication. The primary application-level protocol is HTTP 1.1, and the data is communicated within session(s) initiated by a HTTP-request.

Examples of possible third-party systems include:

- ACS
- ERP
- ABS
- CRM, acting as gateways
LUNA CLOUD

Experience the unmatched quality and speed of LUNA CLOUD solution for controlled and uncontrolled face recognition. Connect unlimited number of cameras, no matter where they are located and benefit from our unique cross-domain face recognition technology for server, WEB-browsers, mobile OS, embedded devices.

- Video storage in the cloud, locally (on an SD-card or hard drive), or both
- Three ways to record to a local archive: continuously, scheduled, or by activation
- Flexible management of access rights
- Work with any data-transfer networks, including 3G and satellite Internet
- Ready integration with ACS, fire alarm systems, POS, GPS trackers and more than 225 applications and devices through IFTTT
Face recognition accuracy

- TPR at FPR $10^{-3}$ $\approx 99.17\%$
- TPR at FPR $10^{-6}$ $\approx 97.88\%$
- TPR at FPR $10^{-8}$ $\approx 91.29\%$

256 bytes Face Template size
Cross-platform face recognition

- Set of front end modules for face detection, best shot selection and normalization

- Stable cross-platform operation in flexibly configurable end solution architectures

- Ultra-fast GPU powered backend face recognition platform with extremely low latency
Cross-domain face recognition

Various input data options are supported for face recognition and matching:

- Live video stream and video archive
- Photo ID
- Mobile and WEB camera face captures
Deep Learning: Training

10M+ face images

- Access to data from multiple device types, face capture scenarios with varying quality settings and environment conditions guarantees robust algorithm operation

- Intense use of NVIDIA GPUs for training; parallel training on 4-8 GPUs gives faster training and better accuracy with larger batch sizes
Cross-domain face recognition: Example

- Matching across 1.5M face database
- 1K+ correct matches in 3 days of surveillance footage
Attributes Detection

Age recognition

Results for 50k test set:
VLNet MAE = 2.21 years
Shallow Net MAE = 2.67 years

Gender recognition

VLNet:
AUC=0.9994  EER=1.03%
Shallow net:
AUC=0.9988  EER=1.51%

Dataset with 40 attributes

% of people in happy mood over time

Saturday

Evening

Morning
DNN Compression

Original network

Pruned & quantized network

32 bits for weight
Prune & quantize

5 bits for weight
Compress

- Compression without network retraining
  - Easy to integrate
  - Compression rate up to ~15x

- Compression with network retraining
  - More difficult to integrate
  - Compression rate up to ~30x

Example:
- Initial face recognition model: 347 MB; ROC AUC = 0.99710
- Model after compression (without retraining): 41 MB; ROC AUC = 0.99719
DNN Acceleration on GPU

- Proprietary DNN inference framework
- Graph compilation for inference (GPU memory savings)
- int8 inference support (3x speedup)
- Fast approximate nearest neighbours search using GPUs (5-10x speedup)
DNN Acceleration on ARM

- Highly optimized routines for typical CNN layers
- Support for quantized 8-bit inference
- Faster than NNPACK library (no support for quantized inference)
  - 2.5x improvement on face recognition on 32 bits
  - 6x improvement on face detection (small net) on 32 bits
- Faster than Arm Compute library:
  - on average 2x improvement on 32 bits
  - on average 3x improvement on 8 bits
Face Detection Acceleration

- Cascaded CNN-based face detector
- State-of-the-art accuracy on FDDDB benchmark
- Full model size <2Mb
- Robust to face rotations, lighting, occlusions, glasses, scale variations, ...
- Significantly faster compared to fully-convolutional methods (Yolo2, SSD)

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Still image</th>
<th>Video frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPU (TESLA P4)</td>
<td>9ms</td>
<td>3ms</td>
</tr>
<tr>
<td>CPU (Intel core i5, 1 core)</td>
<td>22ms</td>
<td>6ms</td>
</tr>
<tr>
<td>ARM (Apple iPhone 7, 1 core)</td>
<td>52ms</td>
<td>22ms</td>
</tr>
</tbody>
</table>
Face Landmarks Acceleration

Better accuracy
VisionLabs

Face Alignment Network (FAN)

Better speed
- FAN runtime – 200 ms on TitanX GPU
- VisionLabs runtime – 5 ms on iPhone 7
## NIST FRVT Ongoing

<table>
<thead>
<tr>
<th>#</th>
<th>Developer Name</th>
<th>Corporate HQ</th>
<th>NIST FRVT Ongoing Listed</th>
<th>FIDO Certified</th>
<th>NVIDIA Metropolis</th>
<th>3D Support (RGB-D)</th>
<th>Liveness Detection</th>
<th>Template (bytes)</th>
<th>Template (ms)</th>
<th>Matching (ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>VisionLabs</td>
<td>Netherlands</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>264</td>
<td>265</td>
<td>799</td>
</tr>
<tr>
<td>2</td>
<td>Company A*</td>
<td>France</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>116</td>
<td>41</td>
<td>1042</td>
</tr>
<tr>
<td>3</td>
<td>Company B*</td>
<td>Lithuania</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>2718</td>
<td>881</td>
<td>369140</td>
</tr>
<tr>
<td>4</td>
<td>Company C*</td>
<td>Russia</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>4096</td>
<td>274</td>
<td>634</td>
</tr>
<tr>
<td>5</td>
<td>Company D*</td>
<td>Cyprus</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>6744</td>
<td>330</td>
<td>82524</td>
</tr>
<tr>
<td>6</td>
<td>Company E*</td>
<td>China</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>2058</td>
<td>356</td>
<td>17799</td>
</tr>
<tr>
<td>7</td>
<td>Company F*</td>
<td>Russia</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>1330</td>
<td>782</td>
<td>83072</td>
</tr>
<tr>
<td>8</td>
<td>Company G*</td>
<td>China</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>4130</td>
<td>672</td>
<td>37848</td>
</tr>
<tr>
<td>9</td>
<td>Company H*</td>
<td>China</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>10</td>
<td>Company I*</td>
<td>China</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>11</td>
<td>Company J*</td>
<td>Spain</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>12</td>
<td>Company K*</td>
<td>Germany</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>6144</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>13</td>
<td>Company L*</td>
<td>Japan</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>14</td>
<td>Company M*</td>
<td>Japan</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>15</td>
<td>Company N*</td>
<td>USA</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

# NIST FRVT Ongoing

## False Non-Match Rate / Top-Rank Position

<table>
<thead>
<tr>
<th>#</th>
<th>Developer Name</th>
<th>Corporate HQ</th>
<th>FMR</th>
<th>Visa 0.0001</th>
<th>Mugshot 0.0001</th>
<th>Webcam 0.0001</th>
<th>Selfie 0.0001</th>
<th>Wild 0.0001</th>
<th>Child 0.01</th>
<th>Avg. Rank (less better)</th>
<th>Race Invariant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>VisionLabs</td>
<td>Netherlands</td>
<td></td>
<td>0.014/2</td>
<td>0.020/3</td>
<td>0.001/4</td>
<td>0.017/7</td>
<td>0.397/3</td>
<td>0.556/6</td>
<td>4.17</td>
<td>✓</td>
</tr>
<tr>
<td>2</td>
<td>Company A*</td>
<td>France</td>
<td></td>
<td>0.033/12</td>
<td>0.028/6</td>
<td>Fail/39</td>
<td>Fail/39</td>
<td>0.931/33</td>
<td>0.901/37</td>
<td>27.67</td>
<td>✗</td>
</tr>
<tr>
<td>3</td>
<td>Company B*</td>
<td>Lithuania</td>
<td></td>
<td>0.044/18</td>
<td>0.046/19</td>
<td>0.001/7</td>
<td>0.017/6</td>
<td>0.817/26</td>
<td>0.533/3</td>
<td>13.17</td>
<td>✗</td>
</tr>
<tr>
<td>4</td>
<td>Company C*</td>
<td>Russia</td>
<td></td>
<td>0.020/3</td>
<td>0.030/9</td>
<td>Fail/36</td>
<td>Fail/36</td>
<td>0.492/5</td>
<td>0.549/4</td>
<td>15.50</td>
<td>✗</td>
</tr>
<tr>
<td>5</td>
<td>Company D*</td>
<td>Cyprus</td>
<td></td>
<td>0.021/4</td>
<td>0.023/4</td>
<td>Fail/34</td>
<td>Fail/34</td>
<td>0.324/2</td>
<td>0.459/1</td>
<td>13.17</td>
<td>✗</td>
</tr>
<tr>
<td>6</td>
<td>Company E*</td>
<td>China</td>
<td></td>
<td>0.030/9</td>
<td>0.039/14</td>
<td>0.010/15</td>
<td>0.063/17</td>
<td>0.725/18</td>
<td>0.746/16</td>
<td>14.83</td>
<td>✗</td>
</tr>
<tr>
<td>7</td>
<td>Company F*</td>
<td>Russia</td>
<td></td>
<td>0.013/1</td>
<td>0.019/2</td>
<td>0.011/16</td>
<td>0.012/3</td>
<td>0.948/36</td>
<td>0.762/21</td>
<td>13.17</td>
<td>✗</td>
</tr>
<tr>
<td>8</td>
<td>Company G*</td>
<td>China</td>
<td></td>
<td>0.021/5</td>
<td>0.017/1</td>
<td>0.000/1</td>
<td>0.012/1</td>
<td>0.431/4</td>
<td>0.586/8</td>
<td>3.33</td>
<td>✗</td>
</tr>
<tr>
<td>9</td>
<td>Company H*</td>
<td>China</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Company I*</td>
<td>China</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Company J*</td>
<td>Spain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Company K*</td>
<td>Germany</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Company L*</td>
<td>Japan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Company M*</td>
<td>Japan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Company N*</td>
<td>USA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*list of companies-participants: www.nist.gov/sites/default/files/documents/2017/10/03/frvt_report_2017_10_03.pdf
Face recognition for target groups

The Visionlabs algorithm appears most insensitive to country effects

TPR @FPR 1e-4

<table>
<thead>
<tr>
<th>Group</th>
<th>TPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aged</td>
<td>85.8%</td>
</tr>
<tr>
<td>Asian</td>
<td>76.1%</td>
</tr>
<tr>
<td>African</td>
<td>81.4%</td>
</tr>
<tr>
<td>Hindu</td>
<td>88.4%</td>
</tr>
</tbody>
</table>
GPU powered interbank credit fraud prevention

Credit Bureau Fraud Prevention biometric service is used by 40+ banks and microfinance organizations today processes more than 5M identity verification requests monthly
GPU powered biometric identification platform

Biggest in Eastern Europe bank biometric identification platform with already existing database of 110M+ Customers processes ~ 1000 simultaneous identification requests (1-to-110 million matching) every second with allowed latency of 1 second
GPU powered city-wide security video surveillance

Criminal suspect search in city-wide (9.5M inhabitants) video surveillance system in uncontrolled environment works both in real-time video stream processing and video archive processing scenarios.
GPU powered B2C security video surveillance system

Nationwide B2C security video surveillance system hosted at Amazon AWS consisting of 250,000+ cameras that handles close to 300,000 hours of video data each day to detect humans present at the entrance to Customer houses.
Alexander Khanin  
CEO, VisionLabs  
E-mail: a.khanin@visionlabs.ai  
Web: www.visionlabs.ai