TO RECOGNIZE FAMILIES IN THE WILD:
A MACHINE VISION TUTORIAL

Joseph Robinson¹, Ming Shao², Yun Fu¹

1. Northeastern University, Boston, MA, USA
2. University of Massachusetts, Dartmouth, MA, USA
ABOUT JOE ROBINSON (PRESENTER)

- **Northeastern University, Boston, MA, USA**
- 4th year PhD
- *Advisor: Yun Fu*
- DHS Fellow

**Related Research**

- **Families In the Wild Database**
  - Published work in ACM MM, IEEE FG, PAMI

https://www.jrobsvision.com
https://web.northeastern.edu/smilelab/fiw/
ABOUT JOE ROBINSON (PRESENTER)

- Related Experience
  - Organized 3 data challenges
  - Organized and hosted related workshops

About Joe Robinson

Organized 3 data challenges
Organized and hosted related workshops

https://www.jrobsvision.com
To Recognize Families In the Wild: A Machine Vision Tutorial

ABOUT JOE ROBINSON (PRESENTER)

- Related research, internships, collaborations

https://www.jrobsvision.com

MIT Lincoln Laboratory
https://www.ll.mit.edu/

TRECVID

Raytheon BBN Technologies

http://www.stresearch.com/

IARPA Programs


Janus

Odin
OUTLINE

• Overview of visual kinship recognition
• Deep dive: kinship recognition research (2010-2017)
• Families In the Wild Database
• Break
• More applications & modern-day deep learning approaches
• Rapid DNA Testing (DHS)
• RFIW Data Challenges
• Closing Remarks
KINSHIP RECOGNITION: OVERVIEW

- Kin relationship predictable from DNA testing
  - Expensive
  - Complex
  - Time consuming

- Images accessible everywhere
  - Most images are human centric

- Two questions to consider
  - Who are these people
  - What are their relationships

- Visual kinship recognition researched for nearly a decade
  - Challenging task: varying age, ethnicity, expression, pose, more

What can we do by utilizing these images???
KINSHIP RECOGNITION: FORESEEABLE USE CASES

Search Investigations

Modern-Day Refugee Crisis

Genealogy Services/Research

Social Media Recommendations

To Recognize Families In the Wild: A Machine Vision Tutorial
Missing Children

To Recognize Families In the Wild: A Machine Vision Tutorial
What is visual kinship recognition?

Griffin Family

Simpson Family
Recognize Relationship Type

Mom

Dad

Children

Siblings
Fine-grain categorization

Abraham  Mona
  /    \
 /     \
Herb  Abbie  Homer
  /    \
 /     \
Lisa  Bart  Maggie  Ung

Jackie  Clancy
  /    \
 /     \
Marge  Patty  Selma
KINSHIP RECOGNITION: LOOKING BACK

2010  2012  2014  2015  2017

To Recognize Families In the Wild: A Machine Vision Tutorial

Download link: [http://chenlab.ece.cornell.edu/projects/KinshipVerification/KinshipVerification.zip](http://chenlab.ece.cornell.edu/projects/KinshipVerification/KinshipVerification.zip)
CORNELL KIN DATASET (2010)

- 150 parent/child pairs from the Internet (celebrities, public figures, etc.)
To Recognize Families In the Wild: A Machine Vision Tutorial

## Cornell Kin Dataset (2010)

- **List of targeted features**

| Color cues:          | Left eyebrow length | Left eyebrow thickness | Left eyebrow size | Left eyebrow subwindow | Right eye width | Left eye height | Left eye size | Right eye width | Right eye height | Right eye size | Left eye subwindow | Right eye subwindow | Left ear subwindow | Right ear subwindow | Left cheek width | Left cheek height | Left cheek height | Left cheek height | Left cheek height | Right cheek width | Right cheek width | Right cheek height | Chin shape | Distance Features | Statistical Features |
|----------------------|---------------------|------------------------|-------------------|------------------------|------------------|----------------|--------------|----------------|------------------|------------------|----------------|--------------------|--------------------|------------------|-------------------|------------------|------------------|------------------|------------------|-------------------|----------------|------------------|---------------------|
| Skin color           |                     |                        |                   |                        |                  |                |              |                |                  |                  |                |                    |                    |                  |                  |                  |                  |                  |                  |                  |                  |                   |            | Eye distance        | Histogram of Gradients |
Color cues:

<table>
<thead>
<tr>
<th>Skin color</th>
<th>Hair color</th>
</tr>
</thead>
</table>

- Mode filter at predetermined locations

![Skin color](image1)

![Hair color](image2)
List of targeted features

<table>
<thead>
<tr>
<th>Color cues:</th>
<th>Facial part features:</th>
<th>Distance Features</th>
<th>Statistical Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin color</td>
<td>Left eye width</td>
<td>Eye distance</td>
<td>Histogram of Gradients</td>
</tr>
<tr>
<td>Hair color</td>
<td>Left eye height</td>
<td>Nostril to mouth</td>
<td></td>
</tr>
<tr>
<td>Left eye color</td>
<td>Left eye size</td>
<td>Left temple to eye corner</td>
<td></td>
</tr>
<tr>
<td>Right eye color</td>
<td>Right eye width</td>
<td>Left eye comer to nose</td>
<td></td>
</tr>
<tr>
<td>Left eyebrow color</td>
<td>Right eye height</td>
<td>Right temple to eye corner</td>
<td></td>
</tr>
<tr>
<td>Right eyebrow color</td>
<td>Left eyebrow width</td>
<td>Right eye comer to nose</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Left eyebrow height</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Left eyebrow thickness</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Left eyebrow size</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Right eyebrow width</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Right eyebrow height</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Left eyebrow thickness</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Left eyebrow size</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Right eyebrow size</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Left eyebrow subwindow</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Right eyebrow subwindow</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mouth width</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Upper lip thickness</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lower lip thickness</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mouth subwindow</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Left cheek width</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Left cheek height</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Right cheek width</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Right cheek height</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chin shape</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
List of targeted features

- Euclidean distance normalized by the width of the face (BB)

Distance Features

- Eye distance
- Nostril to mouth
- Left temple to eye corner
- Left eye corner to nose
- Right temple to eye corner
- Right eye corner to nose

Width (normalization factor)
CORNELL KIN DATASET (2010)

- **Pictorial Structure**
  - Local appearance
    - Part models
    - Parts ≠ feature detection
  - Global geometry
    - Not necessarily fully connected graph
  - Joint optimization
    - Combine appearance and geometry **without hard constraints**
      - “Stretch and fit”
      - Qualitative
KINSHIP RECOGNITION: LOOKING BACK


Download link: [http://www1.ece.neu.edu/~yunfu/research/Kinface/Kinface.htm](http://www1.ece.neu.edu/~yunfu/research/Kinface/Kinface.htm)
UB KINSHIP DATABASE (2012)

- Genetic Observations

Most of children look like their parents at young ages

How to USE this to IMPROVE visual kinship recognition performance?

Key: Transfer Learning
UB KINSHIP DATABASE (2012)

- Child – old parent (Target)
- Child – young parent (Source)

IDEA – Transfer the knowledge from child to the young parent and then to the parent at an older age
UB KINSHIP DATABASE (2012)
- 200 parent-child pairs; 400 subjects; 600 images (young/old parent)
- Collected from the Internet (Illuminations, expressions, etc.)

Old Parents

Young Parents

Children
To Recognize Families In the Wild: A Machine Vision Tutorial

UB KINSHIP DATABASE (2012)

- Son-Father: 93 cases, 46.5%
- Son-Mother: 12 cases, 6%
- Daughter-Father: 77 cases, 38.5%
- Daughter-Mother: 18 cases, 9%
UB KINSHIP DATABASE (2012)
- 200 parent-child pairs; 400 subjects; 600 images (young/old parent)
- Collected from the Internet (Illuminations, expressions, etc.)

- Son-Father: 93
- Son-Mother: 12
- Daughter-Father: 77
- Daughter-Mother: 18
UB KINSHIP DATABASE (2012)
- Different Races
- Gender and relationships
UB KINSHIP DATABASE (2012)

- 200 parent-child pairs; 400 subjects; 600 images (young/old parent)
- Collected from the Internet (Illuminations, expressions, etc.)

![Bar Chart]

- Male: 71.5%
- Female: 66%
- Asian: 28.5%
- Non-Asian: 34%
UB KINSHIP DATABASE (2012)

A pair of inquire images

True or False?

UB KinFace Database

To Recognize Families In the Wild: A Machine Vision Tutorial
GENEALOGICAL FACE RECOGNITION

- Partition face in local regions (5 layers)
- Feature extraction by Gabor (8 directions and 5 scales)
- Select most discriminative features

\[
\text{max} \left\{ \text{cum}_{i+1}(R_{CvsO}) - \text{cum}_i(R_{CvsO}) \right\},
\]
Transfer subspace learning

Subspace where
- Child – young parent still discriminative
- Child – young parent and child – old parent share same distribution

Solved iteratively via gradient descent
Converges well to a local minima

\[ W = \arg\min_{W \in \mathbb{R}^{d \times D}} \{ F(W) + \lambda D_W(P_L || P_U) \} \]

- \( W \) -- projection matrix
- \( F(W) \) -- subspace learning function, i.e., LDA
- \( D_W \) -- Bergman Divergence
EXPERIMENTAL RESULTS

- Pairwise kinship verification
- Metric: metric learning
- TSL: Transfer subspace learning
- Structure: anthropometric model
- Local Gabor: feature extraction method proposed

<table>
<thead>
<tr>
<th>Method</th>
<th>Feature</th>
<th>5-fold</th>
<th>Leave-1-out</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pairwise</td>
<td>Structure</td>
<td>51.00%</td>
<td>52.75%</td>
</tr>
<tr>
<td>Metric</td>
<td>Structure</td>
<td>53.50%</td>
<td>50.50%</td>
</tr>
<tr>
<td>TSL</td>
<td>Structure</td>
<td>52.75%</td>
<td>52.75%</td>
</tr>
<tr>
<td>Pairwise</td>
<td>Local Gabor</td>
<td>50.25%</td>
<td>59.75%</td>
</tr>
<tr>
<td>Metric</td>
<td>Local Gabor</td>
<td>51.25%</td>
<td>52.50%</td>
</tr>
<tr>
<td>TSL</td>
<td>Local Gabor</td>
<td>56.50%</td>
<td>69.67%</td>
</tr>
</tbody>
</table>
EXPERIMENTAL RESULTS

- HUMAN vs MACHINE

• **TSL 2Dom:**
  - child + young parent & child + old parent

• **TSL 3Dom:**
  - child, young parent, old parent

• **Baseline 1:**
  - Human performance without training

• **Baseline 2:**
  - Human performance with training

![Leave-one-out results graph]
GENEALOGICAL FACE RECOGNITION

- VALID or not?
- Feature selected
- Kinship recognition
  - 200 child – old parent pairs
  - 200 child – young parent pairs

Most of children look like their parents at young ages
WHAT WAS LEARNED (2012)

- Local **Gabor outperforms** structure feature
- Machine may perform better than humans based on small sample set
- Kinship verification is challenging
- Transfer learning based method performs better than pairwise one

**Future work**

- **Specific classifiers** for 4 different relations
- Determine **impact race has** on kinship verification
- Include sibling and **other relationship types**
- **Meda-data** in family album

Download: [http://www.kinfacew.com/download.html](http://www.kinfacew.com/download.html)
KinWild Databases (2014)

**KinFaceW I**
- Acquired from different photos
- 156 (F-S) 134 (F-D) 116 (M-S) 127 (M-D)
  >500 in total

**KinFaceW II**
- Most from same photos
- 250 pairs for each type
  1,000 in total
Facial Recognition Schemes

Local Binary Pattern Histograms

**Overview**

- Uses local (texture) descriptors.
- Thresholds each pixel with neighboring pixels to generate a binary pattern.

\[
LBP(x_c, y_c) = \sum_{p=0}^{p-1} 2^p s(i_p - i_c)
\]

Summation around center pixel

Where \(i_p, i_c\) are intensity of center pixel & \(s = sign(x) = \begin{cases} 1 & x \geq 0 \\ 0 & else \end{cases}\)

Each \((x_c, y_c)\) has neighbors \((x_p, y_p)\) that are positioned by

\[
x_p = x_c + R \cos \left( \frac{2\pi p}{p} \right) \\
y_p = y_c - R \sin \left( \frac{2\pi p}{p} \right)
\]
Facial Recognition Schemes

Overview

- Uses local (texture) descriptors.
- Thresholds each pixel with neighboring pixels to generate a binary pattern.
Facial Recognition Schemes

Local Binary Pattern Histograms

**Overview**

- Histograms are nearly identical, regardless of the large variations in light.
EVALUATIONS

THE KINSHIP VERIFICATION IN THE WILD EVALUATION

Organizers: Jiwen Lu

Summary: Over the past two decades, many face image analysis problems have been investigated in computer vision. Representative examples include face alignment, face recognition, age estimation, facial behaviour analysis, gender classification and ethnicity recognition. Recent advances in face analysis have shown that it is possible to infer the kin relation of persons from their facial images. While kinship verification from facial images is an interesting and challenging problem, the performance of existing kinship verification approaches are still far from satisfying, especially when face images are captured in the wild. In this evaluation, we are interested in going a step further and evaluate the performance of possible solutions for facial kinship verification in the wild. There are many potential applications for kinship verification such as family album organization, missing parent/child search, and social media analysis.


http://www.kinfacew.com/results.html#NRML
To Recognize Families In the Wild: A Machine Vision Tutorial

<table>
<thead>
<tr>
<th>Method</th>
<th>F-D Subset</th>
<th>M-D Subset</th>
</tr>
</thead>
<tbody>
<tr>
<td>SILD (LBP)</td>
<td>78.22</td>
<td>71.13</td>
</tr>
<tr>
<td>SILD (HOG)</td>
<td>80.46</td>
<td>74.94</td>
</tr>
<tr>
<td>Polito</td>
<td>85.30</td>
<td>86.30</td>
</tr>
<tr>
<td>LIRIS</td>
<td>83.04</td>
<td>82.74</td>
</tr>
<tr>
<td>ULPGC</td>
<td>71.23</td>
<td>70.01</td>
</tr>
<tr>
<td>NUAA</td>
<td>86.25</td>
<td>82.96</td>
</tr>
<tr>
<td>BIU (LBP)</td>
<td>86.90</td>
<td>79.25</td>
</tr>
<tr>
<td>BIU (HOG)</td>
<td>85.80</td>
<td>79.75</td>
</tr>
<tr>
<td>NUA</td>
<td>87.50</td>
<td>86.70</td>
</tr>
<tr>
<td>Polito</td>
<td>80.64</td>
<td>84.98</td>
</tr>
<tr>
<td>ULPGC</td>
<td>70.85</td>
<td>80.89</td>
</tr>
</tbody>
</table>

Reproduce all results with sample code: [http://www.kinfacew.com/codes/NRML.zip](http://www.kinfacew.com/codes/NRML.zip)
<table>
<thead>
<tr>
<th>Methods</th>
<th>F-S</th>
<th>F-D</th>
<th>M-S</th>
<th>M-D</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>SILD (LBP) [2]</td>
<td>78.22</td>
<td>69.40</td>
<td>66.81</td>
<td>70.10</td>
<td>71.13</td>
</tr>
<tr>
<td>SILD (HOG) [2]</td>
<td>80.46</td>
<td>72.39</td>
<td>69.82</td>
<td>77.10</td>
<td>74.94</td>
</tr>
<tr>
<td>Polito [4]</td>
<td>85.30</td>
<td>85.80</td>
<td>87.50</td>
<td>86.70</td>
<td>86.30</td>
</tr>
<tr>
<td>LIRIS [4]</td>
<td>83.04</td>
<td>80.63</td>
<td>82.30</td>
<td>84.98</td>
<td>82.74</td>
</tr>
<tr>
<td>ULPGC [4]</td>
<td>71.23</td>
<td>70.85</td>
<td>58.52</td>
<td>80.89</td>
<td>70.01</td>
</tr>
<tr>
<td>NUAA [4]</td>
<td>86.25</td>
<td>80.64</td>
<td>81.03</td>
<td>83.93</td>
<td>82.96</td>
</tr>
<tr>
<td>BIU [4]</td>
<td><strong>86.90</strong></td>
<td><strong>76.48</strong></td>
<td><strong>73.89</strong></td>
<td><strong>79.75</strong></td>
<td><strong>79.25</strong></td>
</tr>
</tbody>
</table>

Reproduce results with sample code: [http://www.kinfacew.com/codes/NRML.zip](http://www.kinfacew.com/codes/NRML.zip)

To Recognize Families In the Wild: A Machine Vision Tutorial
KINWILD CHALLENGE: KINWILD II (2015)

To Recognize Families In the Wild: A Machine Vision Tutorial
<table>
<thead>
<tr>
<th>Methods</th>
<th>F-S</th>
<th>F-D</th>
<th>M-S</th>
<th>M-D</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>NRML (LBP) [3]</td>
<td>79.20</td>
<td>71.60</td>
<td>72.20</td>
<td>68.40</td>
<td>72.85</td>
</tr>
<tr>
<td>NRML (HOG) [3]</td>
<td>80.80</td>
<td>72.80</td>
<td>74.80</td>
<td>70.40</td>
<td>74.70</td>
</tr>
<tr>
<td>BIU (LBP) [4]</td>
<td>84.24</td>
<td>79.45</td>
<td>75.98</td>
<td>77.04</td>
<td>79.18</td>
</tr>
<tr>
<td>BIU (HOG) [4]</td>
<td>87.51</td>
<td>80.82</td>
<td>79.78</td>
<td>75.63</td>
<td>80.94</td>
</tr>
</tbody>
</table>
**KINSHIP RECOGNITION: LOOKING BACK**

- 2010
- 2012
- 2014
- 2015
- 2017


Download [http://chenlab.ece.cornell.edu/projects/KinshipClassification/Family101_150x120.zip](http://chenlab.ece.cornell.edu/projects/KinshipClassification/Family101_150x120.zip)
FAMILY 101 (2015)

- 101 Families
  - Age Varying
  - Family Tree Structure
Family Classification

- One-to-many classification
- At least 6 members (564 families)
- Leave 1 member out train, test on left-out
Methodology
INDEPENDENCE OF FACIAL PARTS

top hair  small hair  forehead  eyebrows  eyes  face
side hair  nose  cheeks  mustache  mouth  chin
To Recognize Families In the Wild: A Machine Vision Tutorial

CLASSIFICATION

Reconstruction error for part $p$ from family $j$

$$R_j^{(p)} = \left\| y^{(p)} - D_{G_j}^{(p)} \alpha_j^{(p)} \right\|_2^2$$

1. Choose three representative parts with smallest possible residues $R$.

2. Rank the normalized residues for all families on these three parts.

3. Sum the ranks and use the highest rank.
EXP 1: NO. OF FAMILIES

3 family members for training
2 family members for testing
30 images/person for training/testing

Average accuracy (%) vs. No. of Families

- Proposed
- SRC
- KNN
- SVM
- Random
EXP 2: NO. OF PEOPLE FOR TRAINING

To Recognize Families In the Wild: A Machine Vision Tutorial
EXP 2: NO. OF PEOPLE FOR TRAINING

20 families randomly selected
30 images/person for training/testing

Average accuracy (%) vs No. of train people

- Proposed
- SRC
- KNN
- SVM
- Random

To Recognize Families In the Wild: A Machine Vision Tutorial

Download [http://parnec.nuaa.edu.cn/xtan/data/TSKinFace.html](http://parnec.nuaa.edu.cn/xtan/data/TSKinFace.html)
TS KIN FACE DATASET (2015)

- Made up of tri-subject pairs
TS KIN FACE DATASET (2015)
- Made up of tri-subject pairs

502 Father / Mother-Son

512 Father / Mother-Daughter
TS KIN FACE DATASET (2015)

Verification

SBM

Feature Selection

Feature Descriptor

Facial Image Partition

To Recognize Families In the Wild: A Machine Vision Tutorial
To Recognize Families In the Wild: A Machine Vision Tutorial

**TS KIN FACE DATASET (2015)**

![Graph showing ROC curves for various methods]

- **Concatenated+SVM**
- **Sparse Group Lasso**
- **NRML**
- **Gated autoencoder**
- **ITML**
- **LMNN**
- **LMNN**
- **ABM**
- **SBM**
- **ABM-FS**
- **SBM-FS**
- **ABM-block-FS**
- **SBM-block-FS**
- **RSBM-block-FS**
**KINSHIP RECOGNITION: LOOKING BACK**

- Joseph P. Robinson, Ming Shao, Yue Wu, Hongfu Liu, Timothy Gillis, Yun Fu. *Visual Kinship Recognition of Families in the Wild*, IEEE Transactions on pattern analysis and machine intelligence (SI): Computational Face, 2018

Download [https://web.northeastern.edu/smilelab/fiw/download.html](https://web.northeastern.edu/smilelab/fiw/download.html)
Why Not Seen in Practice?

- Challenging Task
- Insufficient data distribution
- Lack labeled data

A Large-scale Kinship Dataset
FAMILIES IN THE WILD (FIW)

Goals

- Collect family photos for 1,000 families from around the globe

The following criteria must be met:

- ~10 family photos
- >2 family members (i.e., tree structure)
- Multiple samples for most members
- Variety of photo types (i.e., entire family, siblings at different ages, profile pics)
Compiled candidate family list

<table>
<thead>
<tr>
<th>F0001</th>
<th>Einstein, Albert</th>
</tr>
</thead>
<tbody>
<tr>
<td>F0002</td>
<td>Lee, Bruce</td>
</tr>
<tr>
<td>F0999</td>
<td>Jackson, Michael</td>
</tr>
<tr>
<td>F1000</td>
<td>Spanish Royal Family</td>
</tr>
</tbody>
</table>

Family ID (FID)
- F0001
- F1000
## Photo and metadata collection

<table>
<thead>
<tr>
<th>PID</th>
<th>FID</th>
<th>SURNAME</th>
<th>URL</th>
<th>METADATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>P00001</td>
<td>F001</td>
<td>Abbas.Mahmouq</td>
<td><a href="http://www.jpost.com/HttpHandlers/Madmoud">http://www.jpost.com/HttpHandlers/Madmoud</a> and wife Amina</td>
<td></td>
</tr>
<tr>
<td>P00002</td>
<td>F001</td>
<td>Abbas.Mahmouq</td>
<td><a href="http://media.israelhayom.co.il/2014/Madmoud">http://media.israelhayom.co.il/2014/Madmoud</a> and wife Amina</td>
<td></td>
</tr>
<tr>
<td>P00003</td>
<td>F001</td>
<td>Abbas.Mahmouq</td>
<td><a href="http://img.mako.co.il/2014/06/15/50/Madmoud">http://img.mako.co.il/2014/06/15/50/Madmoud</a> and wife Amina</td>
<td></td>
</tr>
<tr>
<td>P00004</td>
<td>F001</td>
<td>Abbas.Mahmouq</td>
<td><a href="http://images1.ynet.co.il/PicServer4/Madmoud">http://images1.ynet.co.il/PicServer4/Madmoud</a> and wife Amina</td>
<td></td>
</tr>
<tr>
<td>P00005</td>
<td>F001</td>
<td>Abbas.Mahmouq</td>
<td><a href="http://www.menassat.com/files/imagedefacto">http://www.menassat.com/files/imagedefacto</a> Palestinian president Mahmoud Abbas (centre) and his sons Yasser (right)Tareq (l)</td>
<td></td>
</tr>
<tr>
<td>P00006</td>
<td>F001</td>
<td>Abbas.Mahmouq</td>
<td><a href="http://jewsdownunder.files.wordpress/YPasser">http://jewsdownunder.files.wordpress/YPasser</a> R and Brother Tareq L</td>
<td></td>
</tr>
<tr>
<td>P00007</td>
<td>F001</td>
<td>Abbas.Mahmouq</td>
<td><a href="http://www.fpnp.net/uploads/albums/PSMahmoudTareqYasser">http://www.fpnp.net/uploads/albums/PSMahmoudTareqYasser</a></td>
<td></td>
</tr>
<tr>
<td>P00008</td>
<td>F001</td>
<td>Abbas.Mahmouq</td>
<td><a href="http://astandforjustice.org/photos/PMadmoud">http://astandforjustice.org/photos/PMadmoud</a> with wife Amina and random ppl</td>
<td></td>
</tr>
<tr>
<td>P00009</td>
<td>F002</td>
<td>Abbott.Greg</td>
<td><a href="https://theconservativetreehouse.file">https://theconservativetreehouse.file</a></td>
<td>Greg Abbott and his family</td>
</tr>
<tr>
<td>P00010</td>
<td>F002</td>
<td>Abbott.Greg</td>
<td><a href="http://s3.amazonaws.com/static.text">http://s3.amazonaws.com/static.text</a></td>
<td>Greg Abbott and his family at his rally</td>
</tr>
<tr>
<td>P12404</td>
<td>F100</td>
<td>DeNiro.Robert</td>
<td><a href="http://i.dailymail.co.uk/i/pix/2014/01/Wife">http://i.dailymail.co.uk/i/pix/2014/01/Wife</a> and children</td>
<td></td>
</tr>
<tr>
<td>P12405</td>
<td>F100</td>
<td>DeNiro.Robert</td>
<td><a href="http://fm.cnnbc.com/applications/cnnct">http://fm.cnnbc.com/applications/cnnct</a>     Actor Robert De Niro and his wife Grace Hightower</td>
<td></td>
</tr>
<tr>
<td>P12410</td>
<td>F100</td>
<td>DeNiro.Robert</td>
<td><a href="http://i.lv3.hbo.com/assets/images/Robert">http://i.lv3.hbo.com/assets/images/Robert</a> De Niro (right), with his father, Robert De Niro, Sr.</td>
<td></td>
</tr>
<tr>
<td>P12411</td>
<td>F100</td>
<td>DeNiro.Robert</td>
<td><a href="http://66.media.tumblr.com/25303ccf/Robert">http://66.media.tumblr.com/25303ccf/Robert</a> De Niro with his wife Grace and daughter at the 2014 White House Correspondents</td>
<td></td>
</tr>
<tr>
<td>P12412</td>
<td>F100</td>
<td>DeNiro.Robert</td>
<td><a href="http://8pic.ir/images/a200cqvxxwww/Young">http://8pic.ir/images/a200cqvxxwww/Young</a> DeNiro</td>
<td></td>
</tr>
<tr>
<td>P12414</td>
<td>F100</td>
<td>DeNiro.Robert</td>
<td><a href="http://www.thisisnotporn.net/wordpr/Young">http://www.thisisnotporn.net/wordpr/Young</a> DeNiro</td>
<td></td>
</tr>
</tbody>
</table>
To Recognize Families In the Wild: A Machine Vision Tutorial

Data Collection

F0001 Jackson, Michael

F0703 Lee, Bruce

F1000 Damon, Matt

PID₁

Brandon (Bruce’s son)

PID₂

Bruce, wife Linda, and children

Shannon, Brandon

PID₃

Bruce and Linda
LABOR INTENSIVE!

To Recognize Families In the Wild: A Machine Vision Tutorial
To Recognize Families In the Wild: A Machine Vision Tutorial

**Family-level**

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>2</th>
<th>1</th>
<th>Bruce</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>Linda</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>Shannon</td>
<td>F</td>
</tr>
</tbody>
</table>

1-Parent of;
2-Spouse of;
4-Child of

**Family Tree**

Shannon    Linda    Bruce
To Recognize Families In the Wild: A Machine Vision Tutorial

**Ground Truth Labels**

**Family-level**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>1</td>
<td>Bruce</td>
<td>M</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>Linda</td>
<td>F</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>0</td>
<td>Shannon</td>
<td>F</td>
</tr>
</tbody>
</table>

**Image-level**

**Increasing Age**

1. **Bruce**
2. **Linda**
3. **Shannon**

To Recognize Families In the Wild: A Machine Vision Tutorial
Semi-Automatic Labeling

Data Preparation

1. Data Collection
   - Web
     - Family List
       - F0001 Damon, Matt
       - F0703 Lee, Bruce
       - F1000 Jackson, Michael
   - Unlabeled Data
     - Collect PIDs + Metadata
   - Labeled Data
     - Get FID and MID
     - Brandon (Bruce’s son)
     - Bruce, wife Linda, and children
     - Shannon

2. Data Preparation
   - Unlabeled Data
     - Detect Faces + NER
     - PID1: Brandon
     - PID2: Bruce, Linda, Shannon, Brandon
     - PID3: Bruce, Linda
     - 1 Face, 1 Name
     - Create SVM
     - 4 Faces, 4 Names
     - 4 SVMs
     - 3 Faces, 2 Names
     - 2 SVMs (top scores)

3. Label Generation
   - Unlabeled Faces
   - Semi-Supervised Clustering
   - Side information leads to more accurate clustering
   - Unlabeled Faces
   - Semi-Supervised Clustering
   - Side information leads to more accurate clustering

4. Label Validation
   - Labeled Data
     - Encode Faces + Train
     - SVMs
     - Faces
     - CNN
     - Face Encodings
     - One-vs-rest SVMs
     - Use deep face encodings to train an SVM for each MID
   - Update SVMs for each label found
   - New Labeled Faces
     - GUI allows clusters to be validated in a quick and simple manner

To Recognize Families In the Wild: A Machine Vision Tutorial
1. Data Collection

Unlabeled Data
Collect PID + Metadata

F001 Damon, Matt
F0703 Lee, Bruce
F1000 Jackson, Michael

Labeled Data
Get PID and MID

PID1 Brandon
PID2 Bruce and Linda
PID3 Bruce, Linda, Shannon, Brandon

2. Data Preparation

Unlabeled Data
Detect Faces + NER

PID1
PID2
PID3

1 Face, 1 Name
Create SVM
4 Faces, 4 Names
4 SVMs
3 Faces, 2 Names
2 SVMs (top scores)

Labeled Data
Encode Faces + Train SVMs

Faces CNN Face Encodings One-vs-rest SVMs

Use deep face encodings to train an SVM for each MID

3. Label Generation

Unlabeled Faces
Semi-Supervised Clustering

Side information leads to more accurate clustering

Labeled Data
Update SVMs for each label found

Align faces with names and label pairs of high confidence

4. Label Validation

Label Validation Tool

Updated Family Tree
New Labeled Faces

GUI allows clusters to be validated in a quick and simple manner

To Recognize Families In the Wild: A Machine Vision Tutorial
**Semi-Automatic Labeling Scheme (Version 2)**

1. Align faces with names and label pairs of high confidence.

2. Use deep face encodings to train an SVM for each MID.

**Detect Faces + NER**

- **PID₁**: Brandon
- **PID₂**: Bruce, with Linda, Shannon, and Brandon
- **PID₃**: Bruce and Linda

**Train classifiers & mine labels**

- Faces
- CNN
- Face Encodings
- One-vs-rest SVMs

To Recognize Families In the Wild: A Machine Vision Tutorial
Specs

- Very Deep Architecture
  - Trained weights on LFW (~2.6M faces, 2,622 IDs)

- Network Details
  - “Very small” convolution filters (i.e., 3x3)
  - Convolutional stride of 1
  - ReLu non-linearity
  - 3 Fully-Connected Layers

Labelled Faces in the Wild (LFW)
SEMI-AUTOMATIC LABELING SCHEME (VERSION 2)

https://github.com/davidsandberg/facenet

To Recognize Families In the Wild: A Machine Vision Tutorial
To Recognize Families In the Wild: A Machine Vision Tutorial
To Recognize Families In the Wild: A Machine Vision Tutorial

SEMI-AUTOMATIC LABELING SCHEME (VERSION 2)

Input  Output  Pairwise Similarity

Jon Gruden  Jon Gruden  0.82

Kenneth Carlsen  0.74  0.68
To Recognize Families In the Wild: A Machine Vision Tutorial
SEMI-AUTOMATIC LABELING SCHEME (VERSION 2)
SEMI-AUTOMATIC LABELING SCHEME (VERSION 2)

Detect Faces + NER

1. Align faces with names and label pairs of high confidence

Train classifiers & mine labels

2. Use deep face encodings to train an SVM for each MID

To Recognize Families In the Wild: A Machine Vision Tutorial
FIW: LABEL PROPOSAL GENERATION

Unlabeled Faces

Labeled Faces

Cluster

Label Proposals

\[
\min \sum_{k=1}^{K} \sum_{x_i \in C_k} f_{cos}(x_i, m_k) + \lambda U_c(S, H \otimes S)
\]

To Recognize Families In the Wild: A Machine Vision Tutorial
To Recognize Families In the Wild: A Machine Vision Tutorial

**FIW: ANNOTATION**

Input:
Proposed Labels (Clusters)

Cluster Validation GUI

Output: Labeled samples added to FIW
Clustering 1,000 Families

\[
\min \sum_{k=1}^{K} \sum_{x_i \in C_k} f_{cos}(x_i, m_k) + \lambda U_c(S, H \otimes S)
\]
### Table: Time and Clicks Comparison

<table>
<thead>
<tr>
<th>FID</th>
<th>Images</th>
<th>No. of Clicks</th>
<th>Time (h:m:s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F0378</td>
<td>5</td>
<td>97</td>
<td>0:5:31</td>
</tr>
<tr>
<td>F0425</td>
<td>35</td>
<td>551</td>
<td>0:15:08</td>
</tr>
<tr>
<td>F0458</td>
<td>10</td>
<td>153</td>
<td>0:5:18</td>
</tr>
<tr>
<td>F0480</td>
<td>13</td>
<td>178</td>
<td>0:6:16</td>
</tr>
<tr>
<td>F0601</td>
<td>128</td>
<td>1,838</td>
<td>1:25:23</td>
</tr>
<tr>
<td>F0624</td>
<td>7</td>
<td>35</td>
<td>0:4:24</td>
</tr>
<tr>
<td>F0995</td>
<td>76</td>
<td>1,272</td>
<td>0:44:52</td>
</tr>
<tr>
<td>Total</td>
<td>274</td>
<td>4,124</td>
<td>2:46:52</td>
</tr>
</tbody>
</table>

- Time per Image decreased by **89.11%**
- Clicks per Image decreased by **97.67%**
To Recognize Families In the Wild: A Machine Vision Tutorial

FIW: SAMPLE FAMILIES

Abdullah of Jordan  Prince of Monaco  King Letsie of Lesotho  Michael Jackson

Margaret Thatcher  Harald V of Norway  Guo Tao  Gronkowski
FIW: SAMPLE FAMILIES (2)
FIW: FAMILY TREES

- Over 13,000 family photos of 1,000 unique family trees
- **Depth, Breadth, and Diverse**
# Families In the Wild (FIW)

Database Overview: Family-Level Stats

<table>
<thead>
<tr>
<th>Dataset</th>
<th>No. Family</th>
<th>No. People</th>
<th>No. Faces</th>
<th>Age Varies</th>
<th>Family Trees</th>
</tr>
</thead>
<tbody>
<tr>
<td>CornellKin [1]</td>
<td>150</td>
<td>300</td>
<td>300</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>UBKinFace [16], [25]</td>
<td>200</td>
<td>400</td>
<td>600</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>KFW-I [26]</td>
<td>×</td>
<td>533</td>
<td>1,066</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>KFW-II [26]</td>
<td>×</td>
<td>1,000</td>
<td>2,000</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>TSKinFace [12]</td>
<td>787</td>
<td>2,589</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Family101 [7]</td>
<td>101</td>
<td>607</td>
<td>14,816</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>FIW [23]</td>
<td>1,000</td>
<td>10,676</td>
<td>30,725</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Margaret Thatcher

Harald V of Norway

Guo Tao

Gronkowski
Families In the Wild (FIW)
Database Overview: Family-Level Stats

<table>
<thead>
<tr>
<th>Name</th>
<th>No. photo</th>
<th>No. faces</th>
<th>Avg. Faces per Member</th>
</tr>
</thead>
<tbody>
<tr>
<td>King Letsie II of Lesotho</td>
<td>12</td>
<td>40</td>
<td>3.33</td>
</tr>
<tr>
<td>Abdullah II of Jordan</td>
<td>10</td>
<td>20</td>
<td>2.00</td>
</tr>
<tr>
<td>Prince of Monaco</td>
<td>8</td>
<td>15</td>
<td>1.88</td>
</tr>
<tr>
<td>Margaret Thatcher</td>
<td>16</td>
<td>14</td>
<td>0.88</td>
</tr>
<tr>
<td>Michael Jackson</td>
<td>6</td>
<td>8</td>
<td>1.33</td>
</tr>
<tr>
<td>Gronkowski</td>
<td>4</td>
<td>10</td>
<td>2.50</td>
</tr>
<tr>
<td>Guo Tao</td>
<td>15</td>
<td>35</td>
<td>2.33</td>
</tr>
<tr>
<td>Harald V of Norway</td>
<td>40</td>
<td>16</td>
<td>0.40</td>
</tr>
</tbody>
</table>
What can we do with FIW?

- Define protocols & benchmarks for database
- Transfer learning: from FIW to other analytics
- Understanding the key factors in visual kinship problem
- Design new problems, e.g., family retrieval, image set based verification/recognition
- Bridge large-scale learning algorithms and social media
Mode 1: Kinship Verification
Task 1: Kinship Verification

- 1-to-1 Boolean (kin/non-kin)
- Different kin types share different hereditary features
  - Types handled separately
- 11 pairwise types
TASK 1: PAIR TYPES

F-D

GF-GD

B-B

F-S

GF-GS

S-S

M-D

GM-GD

SIBS

M-S

GM-GS

Increasing Age
<table>
<thead>
<tr>
<th>Type</th>
<th>KFW-II</th>
<th>Sibling Face</th>
<th>Group Face</th>
<th>Family 101</th>
<th>FIW (Ours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-B</td>
<td>--</td>
<td>232</td>
<td>40</td>
<td>--</td>
<td>103,574</td>
</tr>
<tr>
<td>S-S</td>
<td>--</td>
<td>211</td>
<td>32</td>
<td>--</td>
<td>39,998</td>
</tr>
<tr>
<td>SIBS</td>
<td>--</td>
<td>277</td>
<td>53</td>
<td>--</td>
<td>73,716</td>
</tr>
<tr>
<td>F-D</td>
<td>500</td>
<td>--</td>
<td>69</td>
<td>147</td>
<td>127,876</td>
</tr>
<tr>
<td>F-S</td>
<td>500</td>
<td>--</td>
<td>69</td>
<td>213</td>
<td>91,884</td>
</tr>
<tr>
<td>M-D</td>
<td>500</td>
<td>--</td>
<td>--</td>
<td>148</td>
<td>82,518</td>
</tr>
<tr>
<td>M-S</td>
<td>500</td>
<td>--</td>
<td>70</td>
<td>184</td>
<td>112,378</td>
</tr>
<tr>
<td>GF-GD</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>7,164</td>
</tr>
<tr>
<td>GF-GS</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>4,904</td>
</tr>
<tr>
<td>GM-GD</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>6,302</td>
</tr>
<tr>
<td>GM-GS</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>4,600</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>2,000</strong></td>
<td><strong>720</strong></td>
<td><strong>395</strong></td>
<td><strong>607</strong></td>
<td><strong>654,368</strong></td>
</tr>
</tbody>
</table>
TASK 1: BENCHMARK

- Replaced top layers as **SphereFace**

\[
L_{ang} = \frac{1}{N} \sum_{i=1}^{N} - \log \left( \frac{e^{\|x_i\|\psi(\theta_{y_i,i})}}{e^{\|x_i\|\psi(\theta_{y_i,i})} + \sum_{j \neq y_i} e^{\|x_i\|\cos(\theta_{j,i})}} \right)
\]

Sample code [https://github.com/visionjo/FIW_KRT/tree/master/sphereface_rfiw_baseline](https://github.com/visionjo/FIW_KRT/tree/master/sphereface_rfiw_baseline)

*SphereFace: Deep Hypersphere Embedding for Face Recognition [C] Liu, Weiyang and Wen, Yandong and Yu, Zhiding and Li, Ming and Raj, Bhiksha and Song, Le. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017*
Deep methods outperform hand-crafted features by nearly 15%
Mode 2:
Family Classification
**TASK 2: FAMILY CLASSIFICATION**

**Family Classification**

- One-to-many classification problem
- At least 6 members (565 families used)
- Leave 1 member out train, test on left-out

To Recognize Families In the Wild: A Machine Vision Tutorial
**TASK 2: BENCHMARK**

- Replaced top layers as **Center-loss**

\[
P(y = j \mid x) = \frac{e^{x^T w_j}}{\sum_{k=1}^{K} e^{x^T w_k}}
\]

\[
L = L_S + \lambda L_C
\]

\[
L_C = \frac{1}{2} \sum_{i=1}^{m} \| x_i - c_{y_i} \|_2^2
\]

- Batch
- Pre-trained CNN

A Discriminative Feature Learning Approach for Deep Face Recognition [C] Yandong Wen, Kaipeng Zhang, Zhifeng Li*, Yu Qiao
# Results

Deep methods outperform hand-crafted features by nearly 15%

<table>
<thead>
<tr>
<th>Run ID</th>
<th>Network(s)</th>
<th>Acc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Run-1</td>
<td>VGG-Face, ( f_{c7} ) (4,096D)+one-( vs )-rest SVMs</td>
<td>3.04</td>
</tr>
<tr>
<td>Run-2</td>
<td>VGG-Face, replaced softmax (564D)+fine-tuned</td>
<td>10.42</td>
</tr>
<tr>
<td>Run-3</td>
<td>ResNet-22 + softmax (564D)</td>
<td>14.17</td>
</tr>
<tr>
<td>Run-4</td>
<td>SphereFace, ( f_{c5} ) (512D)</td>
<td>13.86</td>
</tr>
<tr>
<td>Run-5</td>
<td>ResNet-22 + CF, ( f_{c5} ) (512D)</td>
<td>16.18</td>
</tr>
</tbody>
</table>
Mode 3:
Human Performance on Kinship Verification
Deep methods outperform humans by nearly 13%
What can we do with FIW?

- Define protocols & benchmarks for database
- Design new problems, e.g., family retrieval, image set based verification/recognition
- Transfer learning: from FIW to other analytics
- Bridge large-scale learning algorithms and social media
- Understanding the key factors in visual kinship problem
Experimental Study
What layer of face classifier CNN best describes kinship?
Where in the network is kin information mostly represented?

**Tendencies of CNN**
- Each layer closer to top of network describes input with a higher abstraction than those below (i.e., combination of layers passed by)
- Bottom layers learn simple features, like edges, lines, blobs
- Top layers learn complex features like shapes, parts, labels (top-most softmax layer)
DEEP FEATURE EXPERIMENT

- Where in the network is kin information mostly represented?
## Deep Feature Experiment

### KinWild-I

<table>
<thead>
<tr>
<th></th>
<th>FD</th>
<th>FS</th>
<th>MD</th>
<th>MS</th>
<th>Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPB</td>
<td>0.716</td>
<td>0.753</td>
<td>0.678</td>
<td>0.638</td>
<td>0.696</td>
</tr>
<tr>
<td>SIFT</td>
<td>0.736</td>
<td>0.814</td>
<td>0.764</td>
<td>0.711</td>
<td>0.757</td>
</tr>
<tr>
<td>pool5</td>
<td>0.682</td>
<td>0.801</td>
<td>0.785</td>
<td>0.766</td>
<td>0.758</td>
</tr>
<tr>
<td>conv5_2</td>
<td>0.684</td>
<td>0.794</td>
<td>0.755</td>
<td>0.700</td>
<td>0.733</td>
</tr>
<tr>
<td>conv5_3</td>
<td><strong>0.733</strong></td>
<td>0.747</td>
<td>0.804</td>
<td>0.667</td>
<td>0.738</td>
</tr>
<tr>
<td>fc6</td>
<td>0.723</td>
<td><strong>0.838</strong></td>
<td><strong>0.857</strong></td>
<td><strong>0.807</strong></td>
<td><strong>0.806</strong></td>
</tr>
<tr>
<td>fc7</td>
<td>0.710</td>
<td>0.792</td>
<td>0.766</td>
<td>0.759</td>
<td>0.757</td>
</tr>
</tbody>
</table>

### KinWild-II

<table>
<thead>
<tr>
<th></th>
<th>FD</th>
<th>FS</th>
<th>MD</th>
<th>MS</th>
<th>Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.708</td>
<td>0.784</td>
<td>0.690</td>
<td>0.732</td>
<td>0.695</td>
</tr>
<tr>
<td></td>
<td>0.730</td>
<td>0.788</td>
<td>0.822</td>
<td>0.796</td>
<td>0.782</td>
</tr>
<tr>
<td></td>
<td>0.688</td>
<td>0.819</td>
<td>0.762</td>
<td>0.775</td>
<td>0.761</td>
</tr>
<tr>
<td></td>
<td>0.682</td>
<td>0.793</td>
<td>0.730</td>
<td>0.765</td>
<td>0.743</td>
</tr>
<tr>
<td></td>
<td>0.664</td>
<td>0.789</td>
<td>0.748</td>
<td>0.741</td>
<td>0.735</td>
</tr>
<tr>
<td></td>
<td><strong>0.740</strong></td>
<td><strong>0.825</strong></td>
<td><strong>0.845</strong></td>
<td><strong>0.802</strong></td>
<td><strong>0.803</strong></td>
</tr>
<tr>
<td></td>
<td>0.724</td>
<td>0.764</td>
<td>0.804</td>
<td>0.778</td>
<td>0.767</td>
</tr>
</tbody>
</table>

To Recognize Families In the Wild: A Machine Vision Tutorial
## Deep Feature Experiment

<table>
<thead>
<tr>
<th></th>
<th>KinWild-I</th>
<th>KinWild-II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FD</td>
<td>FS</td>
</tr>
<tr>
<td><strong>LPB</strong></td>
<td>0.716</td>
<td>0.753</td>
</tr>
<tr>
<td><strong>SIFT</strong></td>
<td>0.736</td>
<td>0.814</td>
</tr>
<tr>
<td><strong>pool5</strong></td>
<td>0.682</td>
<td>0.801</td>
</tr>
<tr>
<td><strong>conv5_2</strong></td>
<td>0.684</td>
<td>0.794</td>
</tr>
<tr>
<td><strong>conv5_3</strong></td>
<td>0.733</td>
<td>0.747</td>
</tr>
<tr>
<td><strong>fc6</strong></td>
<td>0.723</td>
<td><strong>0.838</strong></td>
</tr>
<tr>
<td><strong>fc7</strong></td>
<td>0.710</td>
<td>0.792</td>
</tr>
<tr>
<td><strong>pool5</strong></td>
<td>0.697</td>
<td><strong>0.835</strong></td>
</tr>
<tr>
<td><strong>conv5_2</strong></td>
<td>0.701</td>
<td>0.784</td>
</tr>
<tr>
<td><strong>conv5_3</strong></td>
<td>0.715</td>
<td>0.754</td>
</tr>
<tr>
<td><strong>fc6</strong></td>
<td><strong>0.746</strong></td>
<td>0.812</td>
</tr>
<tr>
<td><strong>fc7</strong></td>
<td>0.673</td>
<td>0.785</td>
</tr>
</tbody>
</table>
• Replaced topmost layer w triplet-loss
• Fine-tune on FIW dataset
To Recognize Families In the Wild: A Machine Vision Tutorial
- Replaced topmost layer with triplet-loss
- Fine-tune on FIW dataset

Pre-trained VGG-Face

\[ L = \sum_i \left[ \|f(x_i^a) - f(x_i^p)\|^2_2 - \|f(x_i^a) - f(x_i^n)\|^2_2 + \alpha \right]_+ \]
To Recognize Families In the Wild: A Machine Vision Tutorial

Pre-trained VGG-Face

Triplet loss

Input Batch

FIW Pairs

<table>
<thead>
<tr>
<th>B-B</th>
<th>S-S</th>
<th>SIBS</th>
<th>F-D</th>
<th>F-S</th>
<th>M-D</th>
<th>M-S</th>
<th>GFGD</th>
<th>GFGS</th>
<th>GMGD</th>
<th>GMGS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>130,000</td>
<td>130,000</td>
<td>105,000</td>
<td>75,000</td>
<td>75,000</td>
<td>64,000</td>
<td>60,000</td>
<td>1,000</td>
<td>1,350</td>
<td>950</td>
<td>2,000</td>
<td>644,300</td>
</tr>
</tbody>
</table>
## RESULTS

### Accuracy Scores (%)

<table>
<thead>
<tr>
<th></th>
<th>KinWild-I</th>
<th></th>
<th>KinWild-II</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FD</td>
<td>FS</td>
<td>MD</td>
<td>MS</td>
</tr>
<tr>
<td>LBP</td>
<td>71.6</td>
<td>75.3</td>
<td>67.8</td>
<td>63.8</td>
</tr>
<tr>
<td>LBP + PCA</td>
<td>72.8</td>
<td>79.5</td>
<td>71.7</td>
<td>68.1</td>
</tr>
<tr>
<td>SIFT</td>
<td>71.0</td>
<td>78.9</td>
<td>71.8</td>
<td>65.9</td>
</tr>
<tr>
<td>SIFT + PCA</td>
<td>73.9</td>
<td>81.4</td>
<td>76.4</td>
<td>71.1</td>
</tr>
<tr>
<td>NRML (LBP)</td>
<td>81.4</td>
<td>69.8</td>
<td>67.2</td>
<td>72.9</td>
</tr>
<tr>
<td>NRML (HOG)</td>
<td>83.7</td>
<td>74.6</td>
<td>71.6</td>
<td>80.0</td>
</tr>
<tr>
<td>BIU (LBP)</td>
<td>85.5</td>
<td>76.5</td>
<td>69.9</td>
<td>74.4</td>
</tr>
<tr>
<td>BIU (HOG)</td>
<td><strong>86.9</strong></td>
<td>76.5</td>
<td>70.6</td>
<td>79.8</td>
</tr>
<tr>
<td>VGG-Face</td>
<td>67.6</td>
<td>73.7</td>
<td>73.6</td>
<td>73.7</td>
</tr>
<tr>
<td>VGG-Face + PCA</td>
<td>72.0</td>
<td>77.6</td>
<td>78.3</td>
<td>80.6</td>
</tr>
<tr>
<td>Fine-Tuned ($f_{c5}$)</td>
<td>78.0</td>
<td><strong>83.7</strong></td>
<td><strong>87.0</strong></td>
<td><strong>80.8</strong></td>
</tr>
</tbody>
</table>
What can we do with FIW?

- Transfer learning: from FIW to other analytics
- Define protocols & benchmarks for database
- Understand the key factors in visual kinship problem
- Design new problems, e.g., family retrieval, image set based verification/recognition
- Bridge large-scale learning algorithms and social media
DEEP FEATURES FOR SIBLINGS VS REST

![Graph showing distribution of classification scores for different family relationships: Brother-Brother, Sister-Sister, Sister-Brother. The graph compares PDFs for self, sibling, and imposter scores.](image-url)
What can we do with FIW?

- Define protocols & benchmarks for database
- Design new problems, e.g., family retrieval, image set based verification/recognition
- Transfer learning: from FIW to other analytics
- Understanding the key factors in visual kinship problem
- Bridge large-scale learning algorithms and social media

To Recognize Families In the Wild: A Machine Vision Tutorial
Recent Study
From the same photo:
Cheating on Visual Kinship Challenges
FROM SAME PHOTO: CHEATING ON VISUAL KINSHIP CHALLENGES

Mitchell Dawson
Visual Geometry Group
Dept. of Engineering Science
University of Oxford
mdawson@robots.ox.ac.uk

Andrew Zisserman
Visual Geometry Group
Dept. of Engineering Science
University of Oxford
az@robots.ox.ac.uk

Christoffer Nellåker
Nuffield Dept. of Women’s
& Reproductive Health
Big Data Institute, IBME
University of Oxford
christoffer.nellaker@bdi.ox.ac.uk

September 18, 2018
FROM SAME PHOTO: CHEATING ON VISUAL KINSHIP CHALLENGES
FROM SAME PHOTO: CHEATING ON VISUAL KINSHIP CHALLENGES

INPUT IMAGE PAIRS
224 x 224 x 3

SAME PHOTO

DIFFERENT PHOTO

FROZEN WEIGHTS

TRAINED WEIGHTS
FROM SAME PHOTO: CHEATING ON VISUAL KINSHIP CHALLENGES

<table>
<thead>
<tr>
<th>Data set</th>
<th>MD</th>
<th>MS</th>
<th>FD</th>
<th>FS</th>
<th>FMD</th>
<th>FMS</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>KinFaceW-I</td>
<td>86.0</td>
<td>78.3</td>
<td>74.1</td>
<td>74.6</td>
<td></td>
<td></td>
<td>76.8</td>
</tr>
<tr>
<td>KinFaceW-II</td>
<td>94.8</td>
<td>90.3</td>
<td>84.5</td>
<td>92.3</td>
<td></td>
<td></td>
<td>90.2</td>
</tr>
<tr>
<td>FIW</td>
<td>60.3</td>
<td>59.3</td>
<td>59.0</td>
<td>57.5</td>
<td></td>
<td></td>
<td>58.6</td>
</tr>
<tr>
<td>TSKinFace</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>88.6</td>
<td>89.4</td>
<td>88.6</td>
</tr>
<tr>
<td>Cornell KinFace</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>76.7</td>
</tr>
</tbody>
</table>
From Same Photo: Cheating on Visual Kinship Challenges
### From Same Photo: Cheating on Visual Kinship Challenges

<table>
<thead>
<tr>
<th>Paper</th>
<th>Year</th>
<th>KFW-I</th>
<th>KFW-II</th>
<th>Cornell KF</th>
<th>FLW</th>
<th>TSKinFace</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSP classifier (ours)</td>
<td>2018</td>
<td>76.8</td>
<td>90.2</td>
<td>76.7</td>
<td>58.6</td>
<td>88.6</td>
</tr>
<tr>
<td>Mean Accuracy</td>
<td></td>
<td>78.3</td>
<td>82.0</td>
<td>79.2</td>
<td>67.2</td>
<td>85.7</td>
</tr>
<tr>
<td>Median Accuracy</td>
<td></td>
<td>78.8</td>
<td>82.8</td>
<td>79.0</td>
<td>68.8</td>
<td>87.2</td>
</tr>
<tr>
<td>Max Accuracy</td>
<td></td>
<td>96.9</td>
<td>97.1</td>
<td>94.4</td>
<td>74.9</td>
<td>93.4</td>
</tr>
<tr>
<td>Human [44, 11, 28]</td>
<td></td>
<td>71.0</td>
<td>74.0</td>
<td>67.2</td>
<td>79.5</td>
<td></td>
</tr>
<tr>
<td>Aliardi et al. [1]</td>
<td>2018</td>
<td>80.6</td>
<td>88.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moujahid et al. [25]</td>
<td>2018</td>
<td>88.2</td>
<td>88.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lopez et al. [21]</td>
<td>2018</td>
<td>68.4</td>
<td>66.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yan et al. [38]</td>
<td>2018</td>
<td>77.6</td>
<td>78.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robinson et al. [30]</td>
<td>2018</td>
<td>82.4</td>
<td>86.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kohli et al. [15]</td>
<td>2018</td>
<td>96.9</td>
<td>97.1</td>
<td>94.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhou et al. [46]</td>
<td>2018</td>
<td>82.8</td>
<td>85.7</td>
<td>81.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mahpod et al. [24]</td>
<td>2018</td>
<td>79.8</td>
<td>87.0</td>
<td>76.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhao et al. [44]</td>
<td>2018</td>
<td>81.5</td>
<td>82.5</td>
<td>81.7</td>
<td>84.5</td>
<td></td>
</tr>
<tr>
<td>Wang et al. [32]</td>
<td>2018</td>
<td>69.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xia et al. [35]</td>
<td>2018</td>
<td></td>
<td></td>
<td></td>
<td>90.7</td>
<td></td>
</tr>
<tr>
<td>Yang et al. [40]</td>
<td>2017</td>
<td>88.6</td>
<td>90.3</td>
<td></td>
<td></td>
<td>93.4</td>
</tr>
<tr>
<td>Chen et al. [4]</td>
<td>2017</td>
<td>83.3</td>
<td>84.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lu et al. [23]</td>
<td>2017</td>
<td>83.5</td>
<td>84.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patel et al. [27]</td>
<td>2017</td>
<td>78.7</td>
<td>80.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kohli et al. [14]</td>
<td>2017</td>
<td>96.1</td>
<td>96.2</td>
<td>89.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laiadi et al. [17]</td>
<td>2017</td>
<td></td>
<td>83.2</td>
<td>54.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duan et al. [9]</td>
<td>2017</td>
<td></td>
<td></td>
<td>66.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dahan et al. [6]</td>
<td>2017</td>
<td></td>
<td></td>
<td>65.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li et al. [18]</td>
<td>2017</td>
<td></td>
<td></td>
<td>74.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang et al. [31]</td>
<td>2017</td>
<td></td>
<td></td>
<td>68.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhou et al. [45]</td>
<td>2016</td>
<td>78.8</td>
<td>75.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xu et al. [37]</td>
<td>2016</td>
<td>77.9</td>
<td>77.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liu et al. [19]</td>
<td>2016</td>
<td>77.9</td>
<td>81.4</td>
<td>75.9</td>
<td></td>
<td>89.7</td>
</tr>
<tr>
<td>Zhang et al. [41]</td>
<td>2016</td>
<td></td>
<td></td>
<td>71.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robinson et al. [29]</td>
<td>2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duan et al. [10]</td>
<td>2015</td>
<td>73.3</td>
<td>75.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bottinik et al. [2]</td>
<td>2015</td>
<td>86.3</td>
<td>83.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhang et al. [42]</td>
<td>2015</td>
<td>77.5</td>
<td>88.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kou et al. [16]</td>
<td>2015</td>
<td>63.5</td>
<td>68.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yan et al. [39]</td>
<td>2015</td>
<td>70.1</td>
<td>77.0</td>
<td>71.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhang et al. [43]</td>
<td>2015</td>
<td>67.0</td>
<td>86.0</td>
<td></td>
<td>89.0</td>
<td></td>
</tr>
<tr>
<td>Qin et al. [28]</td>
<td>2015</td>
<td></td>
<td></td>
<td></td>
<td>85.4</td>
<td></td>
</tr>
<tr>
<td>Lu et al. [22]</td>
<td>2014</td>
<td>69.9</td>
<td>76.5</td>
<td>66.5</td>
<td>72.1</td>
<td></td>
</tr>
<tr>
<td>Dehghan et al. [7]</td>
<td>2014</td>
<td></td>
<td></td>
<td></td>
<td>80.8</td>
<td></td>
</tr>
<tr>
<td>Nguyen et al. [26]</td>
<td>2011</td>
<td>62.5</td>
<td>71.9</td>
<td></td>
<td></td>
<td>70.7</td>
</tr>
<tr>
<td>Fang et al. [11]</td>
<td>2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weinberger et al. [33]</td>
<td>2009</td>
<td>63.3</td>
<td>74.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Mode 3: Missing Child Search & Retrieval

PROBLEM DEFINITION

- Kinship search & retrieval: images of **missing** children

Parents

Children

Missing

recognize

To Recognize Families In the Wild: A Machine Vision Tutorial
PROBLEM DEFINITION

- Source database: complete parents & children pairs
OUR APPROACH (I)

- Cross-family knowledge transfer

Source Database

Target Database

Recover

Missing

To Recognize Families In the Wild: A Machine Vision Tutorial
Latent Subspace Transfer

- Missing modality problem*

Low-rank Transfer

\[ Y_{S,p} + Y_{S,c} \rightarrow Y_{t,p} + [Y_{t,c}] \quad \xrightarrow{\text{\(Y_S = Y_TZ\)}} \quad Z \text{ to be low-rank} \]

Latent Subspace Transfer

- Missing modality problem*

Low-rank Transfer

\[ Y_{s,p} + Y_{s,c} \rightarrow Y_{t,p} + [Y_{t,c}] \]

\[ Y_S = Y_T Z \]

\[ Z \& L \text{ low-rank} \]

\[ Y_S = Y_T Z + LY_S \]

\[ Y_S = Y_T Z^* = [Y_{T,A}, Y_{T,B}] Z_c^* \]

\[ = [Y_{T,A}, Y_{T,B}] \begin{bmatrix} V_{T,A} \\ V_{T,B} \end{bmatrix} V_S^T \]

\[ = Y_{T,A} (V_{T,A} V_S^T) + Y_{T,B} V_{T,B} V_S^T \]

\[ = Y_{T,A} (V_{T,A} V_S^T) + U \Sigma V_{T,B}^T V_{T,B} V_S^T \]

\[ = Y_T Z + (U \Sigma V_{T,B}^T V_{T,B} \Sigma^{-1} U^T) Y_S \]

\[ = Y_T Z + LY_S, \]

\[ Z_c^* = V_T V_S^T = \begin{bmatrix} V_{T,A} \\ V_{T,B} \end{bmatrix} V_S^T \]

**Latent Subspace Transfer**

- **Missing modality problem**

\[ Y_{s,p} + Y_{s,c} \rightarrow Y_{t,p} + [Y_{t,c}] \]

\[ Y_S = Y_T Z \]

\[ Y_s = Y_t Z + L Y_s \]

- **Latent Low-rank Constraint**

\[ Y_s = P^T X_t Z + L Y_s \]

- **Missing Modality**

Our Approach (II)

- Person and Family-wise constraints

Person-wise Constraint

\[
\mu_m = \frac{1}{n_m} \sum_{i=1}^{n_m} P^T x_i
\]

\[
\Omega_m(P) = \frac{1}{n_m} \sum_{i=1}^{n_m} \| P^T x_i - \mu_m \|^2_2
\]

\[
\Omega(P) = \sum_{m=1}^{M} \Omega_m(P)
\]

Our Approach (II)

- Person and Family-wise constraints

\[ c_n = \frac{1}{K_n} \sum_{k=1}^{K_n} u_k \]

\[ \Psi_n(P) = \frac{1}{K_n} \sum_{k=1}^{K_n} \| \mu_k - c_n \|_2^2 \]

\[ \Phi(P) = \sum_{n=1}^{N} \Psi_n(P) \]

Our Approach (II)

- Person and Family-wise constraints

Our Approach (II)

In summary:

- **Family-person** structured hierarchy guided with latent subspace learning

\[
\min_{Z,L,P} \|Z\|_* + \|L\|_* + \beta \Omega(P) + \gamma \Psi(P),
\]

\[
s.t. \quad Y_s = P^T X_t Z + LY_s, \quad P^T P = I_p.
\]

Our Approach (II)

- **Person and Family-wise constraints**

<table>
<thead>
<tr>
<th>Family 1</th>
<th>Father</th>
<th>Mother</th>
<th>Child</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="image1.png" alt="Father" /></td>
<td><img src="image2.png" alt="Mother" /></td>
<td><img src="image3.png" alt="Child" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Family 2</th>
<th>Father</th>
<th>Mother</th>
<th>Child</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><img src="image4.png" alt="Father" /></td>
<td><img src="image5.png" alt="Mother" /></td>
<td><img src="image6.png" alt="Child" /></td>
</tr>
</tbody>
</table>

Latent Transfer Subspace

Low-rank Subspace

\[ P^T X_{t1} \]

\[ P^T X_{t2} \]
### Datasets and Evaluation

- **Datasets**
  - Families In the Wild (FIW)\(^1\)
  - Family 101 (FM101)\(^2\)

- **Evaluation**
  - Measurement: Cumulative Match Characteristic (CMC)\(^2\)

<table>
<thead>
<tr>
<th></th>
<th>No. Families</th>
<th>#Family</th>
<th>Parent #Images</th>
<th>Children #Images</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FIW</strong></td>
<td>276 of 1,000</td>
<td>Source</td>
<td>138</td>
<td>2,612</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Target</td>
<td>138</td>
<td>2,480</td>
</tr>
<tr>
<td><strong>FM101</strong></td>
<td>101</td>
<td>Source</td>
<td>50</td>
<td>2,308</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Target</td>
<td>51</td>
<td>3,396</td>
</tr>
</tbody>
</table>


CMC CURVE (FAMILY101)

![CMC Curve Graph]

<table>
<thead>
<tr>
<th>Rank</th>
<th>Centerface</th>
<th>VGGFace</th>
<th>DSDA</th>
<th>LAC</th>
<th>SRRS</th>
<th>NRML</th>
<th>DML</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.6401</td>
<td>0.6031</td>
<td>0.5904</td>
<td>0.6787</td>
<td>0.6393</td>
<td>0.6361</td>
<td>0.6267</td>
<td>0.7006</td>
</tr>
<tr>
<td>20</td>
<td>0.7222</td>
<td>0.6881</td>
<td>0.6635</td>
<td>0.7624</td>
<td>0.7224</td>
<td>0.7206</td>
<td>0.7111</td>
<td>0.7876</td>
</tr>
<tr>
<td>30</td>
<td>0.7746</td>
<td>0.7404</td>
<td>0.7015</td>
<td>0.8106</td>
<td>0.7741</td>
<td>0.7722</td>
<td>0.7591</td>
<td>0.8319</td>
</tr>
<tr>
<td>40</td>
<td>0.8077</td>
<td>0.7815</td>
<td>0.7289</td>
<td>0.8393</td>
<td>0.8070</td>
<td>0.8080</td>
<td>0.7939</td>
<td>0.8585</td>
</tr>
<tr>
<td>50</td>
<td>0.8301</td>
<td>0.8089</td>
<td>0.7507</td>
<td>0.8548</td>
<td>0.8313</td>
<td>0.8298</td>
<td>0.8193</td>
<td>0.8794</td>
</tr>
</tbody>
</table>
RESULTS (FIW)

<table>
<thead>
<tr>
<th>FIW</th>
<th>Rank</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centerface</td>
<td>0.2285</td>
<td>0.3305</td>
<td>0.4008</td>
<td>0.4610</td>
<td>0.5162</td>
<td></td>
</tr>
<tr>
<td>VGGFace</td>
<td>0.2313</td>
<td>0.3400</td>
<td>0.4119</td>
<td>0.4660</td>
<td>0.5111</td>
<td></td>
</tr>
<tr>
<td>DSDA</td>
<td>0.2324</td>
<td>0.3344</td>
<td>0.4069</td>
<td>0.4660</td>
<td>0.5106</td>
<td></td>
</tr>
<tr>
<td>LAC</td>
<td>0.2586</td>
<td>0.3629</td>
<td>0.4404</td>
<td>0.5056</td>
<td>0.5563</td>
<td></td>
</tr>
<tr>
<td>SRRS</td>
<td>0.2285</td>
<td>0.3305</td>
<td>0.3974</td>
<td>0.4615</td>
<td>0.5173</td>
<td></td>
</tr>
<tr>
<td>NRML</td>
<td>0.2252</td>
<td>0.3300</td>
<td>0.3974</td>
<td>0.4615</td>
<td>0.5151</td>
<td></td>
</tr>
<tr>
<td>DML</td>
<td>0.2514</td>
<td>0.3690</td>
<td>0.4504</td>
<td>0.5006</td>
<td>0.5475</td>
<td></td>
</tr>
<tr>
<td>Ours</td>
<td><strong>0.3166</strong></td>
<td><strong>0.4353</strong></td>
<td><strong>0.5256</strong></td>
<td><strong>0.5920</strong></td>
<td><strong>0.6371</strong></td>
<td></td>
</tr>
<tr>
<td>Queries</td>
<td>Rank</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family101</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To Recognize Families In the Wild: A Machine Vision Tutorial
PARAMETERS ANALYSIS

- FIW dataset with Center Face Features

\[
\min_{Z, L, P} \|Z\|_* + \|L\|_* + \beta \Omega(P) + \gamma \Psi(P)
\]

\( \beta \) : intra-person variations

\( \gamma \) : intra-family variations
Questions?
Then Let’s Break
About Next Speaker: **Ming Shao**

- Research interests:
  - Applied machine learning
  - Social media analytics
  - Large-scale mining and learning

http://www.cis.umassd.edu/~mshao/
RECENT VISUAL KINSHIP ADVANCES
OUTLINE

- **Spatial**: Visual Kinship and Graph
- **Temporal**: Kinship in Video
- **Learning**: Kinship Metric Learning
- **Deep** Visual Kinship Modeling

Jointly modeling kinship relation in a PHOHO by:
- Demographics
- Spatial context
- Age difference
- Transfer learning based kinship verification

\[
\begin{align*}
    h(X, Y) &= \arg \max_R f(X, Y, R) \\
    f(X, Y, R) &= \sum_{i \neq j} \beta_{r_{ij}} \phi_{r_{ij}}(r_{ij}|y_i, y_j, X)
\end{align*}
\]

\( \phi_{r_{ij}} \) is a linear model that considers:
- gender relation,
- age difference,
- relative
- distance, and kinship score
VISUAL KINSHIP AND GRAPH

- Guo et al. Graph-based Kinship Recognition, ICPR 2014.
- Visual kinship context in family photo is the major motivation

- a multi-class kinship classifier is jointly trained on different kinship relations.
- All valid kinship graphs are generated
- Classifier scores are summed for candidate graph to obtain an overall score
- The kinship graph with the highest overall score is selected as the prediction.
FAMILY PHOTO DETECTION

- Zhang et al. Family Photo Recognition via Multiple Instance Learning, ICMR 2017.

- Recognizing family photos and non-family photos
  - Urquhart Graph is generated to model spatial relation
  - Attribute graph is explored for 1\textsuperscript{st} and 2\textsuperscript{nd} order features
  - Features will be fed to Multiple Instance Learning
  - Solved by Gradient Boosting Decision Tree (GBDT)
**GRAPH AND FAMILY PARSING IN A PHOTO**

- Xia et al., Graph Based Family Relationship Recognition from a Single Image, PRICAI, 2018.

- Parsing the family relation in a photo, one step further from “Family Photo Detection”
GRAPh AND FAMILY PARSING IN ALBUM

- Xia et al., Album to Family Tree: A Graph based Method for Family Relationship Recognition, ACCV, 2018.

- Parsing family in an album (more than one photos)

- Clustering is conducted first to identify different people

A Deformable model based on structure

$$E(v_1, ..., v_n) = \sum_{i=1}^{n} P_i(l_i) + \sum_{i,j} C_{ij}(l_i, l_j)$$

Table 1. Kinship constraint rules

- A child has only one pair of parents. 
- Parents of husband and wife are different (A:B Parents-Son) ∧ (A:C Parents-Daughter) ⇒ ~ (B:C Husband-Wife)
- Siblings of husband and wife are different (A:B Siblings) ∧ (A:C Siblings) ⇒ ~ (B:C Husband-Wife)
- Siblings have the same parents (A:B Parents-Child) ∧ (A:C Parents-Child) ⇒ (B:C Siblings)
- Siblings have the same sibling (A:B Sibling) ∧ (A:C Sibling) ⇒ (B:C Sibling)

Family tree construction and whole family relationship recognition

To Recognize Families In the Wild: A Machine Vision Tutorial
**KINSHIP IN VIDEO: EXPRESSION DYNAMICS**

- Dibeklioglu et al., Like Father, Like Son: Facial Expression Dynamics for Kinship Verification, ICCV 2013

- May explore temporal information to capture the dynamics

- Facial expression dynamics have been demonstrated useful in kinship verification

- Contribute a new kinship expression dataset: **UvA-NEMO Smile Database**

---

**Table 1. Definitions of the extracted features.**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration</td>
<td>$\left[ \frac{\eta(D^+)}{\eta(D)}, \frac{\eta(D^-)}{\eta(D)}, \frac{\eta(D)}{\eta(D)} \right]$</td>
</tr>
<tr>
<td>Duration Ratio</td>
<td>$\left[ \frac{\eta(D^+)}{\eta(D)}, \frac{\eta(D^-)}{\eta(D)} \right]$</td>
</tr>
<tr>
<td>Maximum Amplitude</td>
<td>$\max(D)$</td>
</tr>
<tr>
<td>Mean Amplitude</td>
<td>$\frac{\sum D}{\eta(D)}$</td>
</tr>
<tr>
<td>Maximum Speed</td>
<td>$\left[ \max(V^+), \max(</td>
</tr>
<tr>
<td>Mean Speed</td>
<td>$\left[ \frac{\sum V^+}{\eta(V^+)} \frac{\sum</td>
</tr>
<tr>
<td>Maximum Acceleration</td>
<td>$\left[ \max(A^+), \max(</td>
</tr>
<tr>
<td>Mean Acceleration</td>
<td>$\left[ \frac{\sum A^+}{\eta(A^+)} \frac{\sum</td>
</tr>
</tbody>
</table>
KINSHIP IN VIDEO: METRIC LEARNING

- Haibin Yan and Junlin Hu, Video-based kinship verification using distance metric learning, Pattern Recognition, 2018
- Explore multiple frames from videos for verification
- More frames add up evidence
- Contribute a new dataset Kinship Face Videos in the Wild (KFVW)

Information theoretic metric learning (ITML)
Side-information based linear discriminant analysis (SILD)
KISS metric learning (KISSME)
Cosine similarity metric learning (CSML)
KINSHIP IN VIDEO: VIDLET PAIR AND AUTOENCODER

- Kohli et al., Supervised Mixed Norm Autoencoder for Kinship Verification in Unconstrained Videos, IEEE Trans IP, 2018

- Propose a novel Supervised Mixed Norm regularization Autoencoder (SMNAE) based on sparsity prior

- Vidlet enables frames based kinship verification

- Contribute one video based on kinship dataset: KIVI
Hamdi Dibeklioglu, Visual Transformation Aided Contrastive Learning for Video-based Kinship Verification, ICCV, 2017

“2m+1”-frame sequences (neighborhood of m frames) of given videos are matched based on expression similarity.

A pair of Autoencoders are learned with inverse input-output pairs

Decoder and encoder are shared

Both maximize similarity (kin) and minimize similarity (non-kin)
KINSHIP METRIC LEARNING: NRML AND MNRML

- Lu et al. Neighborhood Repulsed Metric Learning for Kinship Verification, IEEE Trans PAMI

- Similar to conventional metric/subspace learning

- It models the between/within class relation of the data using: $H_1, H_2, H_3$

\[
\max_W J(W) = \text{tr}[W^T (H_1 + H_2 - H_3) W]
\]

subject to $W^T W = I$,

- Can incorporate different features into the learning: neighborhood repulsed metric learning (NRML)

- $\beta$ (account for different features) ensures the maximum

\[
\max_{W, \beta} \sum_{p=1}^{K} \beta_p \text{tr}[W^T (H^p_1 + H^p_2 - H^p_3) W]
\]

subject to $W^T W = I, \sum_{p=1}^{K} \beta_p = 1, \beta_p \geq 0$. 
**Kinship Metric Learning: Multiple Kernels**

- Zhao et al., Learning a Multiple Kernel Similarity Metric for kinship verification, Information Sciences, 2018

- The similarity computation is essentially based on an implicit nonlinear feature transformation

- MKSM is a weighted combination of basic similarities and therefore possesses the capacity for feature fusion

- Essentially solving a constrained linear programming (LP) problem that originates from a Large-margin (LM) criterion

\[
\min_{\{\alpha_{pm}\},\{\xi_i\},b} \sum_{i=1}^{N} w_i \xi_i + \lambda \sum_{p=1}^{P} \sum_{m=1}^{M} \alpha_{pm} \\
\left\{ \begin{array}{l}
z_i(\text{sim}(x_i, y_i; \alpha) + b) \geq \tau - \xi_i, \ i = 1, \ldots, N \\
\alpha_{pm} \geq 0, \ m = 1, \ldots, M, \ p = 1, \ldots, P \\
\xi_i \geq 0, \ i = 1, \ldots, N \\
\end{array} \right.
\]

where \( \sum_{i=1}^{N} w_i = 1 \), \( w_i = \begin{cases} 1/(2N_p), & z_i = 1 \\ 1/(2N_n), & z_i = -1 \end{cases} \)

- \( \alpha \) is sparse weight vector for different kernels
- \( z_i \) is the similarity score for the i-th pair
To Recognize Families In the Wild: A Machine Vision Tutorial

- Lu et al., Discriminative Deep Metric Learning for Face and Kinship Verification, IEEE Trans IP, 2017

- Visual descriptions and knowledge will be distilled through deep networks

- Linear combination of difference distance metrics on top level
**DEEP VISUAL KINSHIP MODELING: CNN**

- Zhang et al., Kinship Verification with Deep Convolutional Neural Networks, BMVC, 2015
- Use the concatenated face images as input and naturally model a binary classification using CNN
- Combine 10 different regions to improve the performance
DEEP VISUAL KINSHIP MODELING: DBN

- Kohli et al., Hierarchical Representation Learning for Kinship Verification, IEEE Trans IP, 2017
- Explore HUMAN PERFORMANCE inspired facial components
- Use unsupervised filtered contractive Deep Belief Networks (fcDBN) for facial feature extraction
- Verification is done by Neural Network and score fusion
DEEP VISUAL KINSHIP MODELING: GAN

- Savas Ozkan and Akin Ozkan, KinshipGAN: Synthesizing Of Kinship Faces From Family Photos By Regularizing A Deep Face Network, ICIP, 2018

- The model takes parents images and generates children’s face images

- Built on conventional GAN but has two Gs and Ds
  - G: parent and child generators
  - D: based on GAN loss, and gender classification

- The model employs a cycle-domain transformation to achieve more stable results

\[ \mathcal{L}_D = \lambda_{gan}\mathcal{L}_{gan}^D + \lambda_{aux}\mathcal{L}_{aux}^D, \]

Father-daughter

\[ \mathcal{L}_G = \lambda_c\mathcal{L}_{con}^C + \lambda_p\mathcal{L}_{con}^P + \mathcal{L}_{gan}^G + \lambda_{aux}\mathcal{L}_{aux}^G. \]
Rapid DNA: DNA in the Field

Reduces Multi-million Dollar Laboratory Processes to One Field Device

- Integration of five forensic lab processes with disposable microfluidic technology.
- Automation allows DHS officers to process samples and receive final results.
- Two U.S. small businesses have commercial devices ready for purchase.

10 hours processing reduced to 90 min.
40% sample cost reduction.
System $250K vs. $1M lab.

Courtesy of Chris Miles from his talk given during RFIW at ACM MM 2017:
Miles, Christopher. "Rapid DNA Performance Results on Family Relationship Verification." Proceedings of the 2017 Workshop on Recognizing Families In the Wild. ACM, 2017
Half of Child’s DNA from Each Parent

Mother (10-1)

Child (10-2)

DNA Profile: 16, 16; 17, 19; 11, 11; 10,10; 8, 8; -; X, X; 13,15; 28, 30; 15, 18; -; -; 11, 11; 12, 12; 6, 9.3; 20, 25

DNA Profile: 16, 16; 17,19; 10, 11; 9,10; 8, 8; 2, 2; X, Y; 13,15; 28, 28; 15, 17; 11, 11; 11, 13; 12, 15; 9.3, 9.3; 21, 25

Probability of Maternity 99.99999996%

Courtesy of Chris Miles from his talk given during RFIW at ACM MM 2017:
Miles, Christopher. "Rapid DNA Performance Results on Family Relationship Verification." Proceedings of the 2017 Workshop on Recognizing Families In the Wild. ACM, 2017
Aunt Can’t Falsely Claim to be Mother

Aunt (11-3)

Child (11-6)

Excluded Match (Six locations don’t match):
- D3
- vWA
- D16
- D21
- TH01
- FGA

DNA Profile: 14,10; 18,18; 9,12; 11,12; 10,10;
-, -, X,X; 13,14; 29,30; 12,13; -, -;
10,11; 12,14; 7,8; 21,21

DNA Profile: 17,18; 16,19; 13,14; 11,11; 10,10;
2,2; X,Y; 11,14; 30,2,31,2; 13,18; 10,10;
11,11; 12,15; 6,9,3; 20,24

Courtesy of Chris Miles from his talk given during RFIW at ACM MM 2017:
Miles, Christopher. "Rapid DNA Performance Results on Family Relationship Verification." Proceedings of the 2017 Workshop on Recognizing Families In the Wild. ACM, 2017
S&T Kinship Validation Samples

Rapid DNA Performance Evaluation based on 31 Known Persons

- 16 Mother to Child Relationship
- 14 Father to Child Relationship
- 11 Mother/Father/Child Relationships
- 12 Full Siblings, 5 Half Siblings

Courtesy of Chris Miles from his talk given during RFIW at ACM MM 2017:
Miles, Christopher. "Rapid DNA Performance Results on Family Relationship Verification." Proceedings of the 2017 Workshop on Recognizing Families In the Wild. ACM, 2017
## DHS Kinship - True Claims

<table>
<thead>
<tr>
<th>PTC Kinship Samples</th>
<th># of tests</th>
<th>Average Relationship Result</th>
<th>Minimum Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mother – Child</td>
<td>16</td>
<td>99.999999997%</td>
<td>99.99995%</td>
</tr>
<tr>
<td>Father – Child</td>
<td>14</td>
<td>99.9999998%</td>
<td>99.99998%</td>
</tr>
<tr>
<td><strong>Family Trio (Paternity)</strong></td>
<td>11</td>
<td>99.999999998%</td>
<td>99.99998%</td>
</tr>
<tr>
<td><strong>Full Siblings</strong></td>
<td>12</td>
<td>99.999999997%</td>
<td>94.84%</td>
</tr>
<tr>
<td><strong>Half Siblings</strong></td>
<td>5</td>
<td>99.78%</td>
<td>95.53%</td>
</tr>
<tr>
<td>Grandmother – Grandchild</td>
<td>4</td>
<td>99.93%*</td>
<td>91.66%</td>
</tr>
<tr>
<td>Grandfather – Grandchild</td>
<td>8</td>
<td>98.93%*</td>
<td>8.12%</td>
</tr>
<tr>
<td>Aunt/Uncle – Niece/Nephew</td>
<td>5</td>
<td>99.21%</td>
<td>30.02%</td>
</tr>
</tbody>
</table>

*When additional family is added for grandparent-grandchild relationships in 8 available tests, the average probability of relationship is 99.99997% with a low of 99.4%*
## DHS Kinship - False Claims

<table>
<thead>
<tr>
<th>PTC Kinship Samples</th>
<th># of tests</th>
<th>Average Relationship Result</th>
<th>Maximum Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>False Claims</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aunt – Child (Mother Claim)</td>
<td>3</td>
<td>Exclusion</td>
<td>NA</td>
</tr>
<tr>
<td>Uncle – Child (Father Claim)</td>
<td>2</td>
<td>Exclusion</td>
<td>NA</td>
</tr>
<tr>
<td>Aunt – Uncle – Child (Mother/Father/Child Claim)</td>
<td>3</td>
<td>Exclusion</td>
<td>NA</td>
</tr>
<tr>
<td>Half Siblings (Full Siblings Claim)</td>
<td>5</td>
<td>98.60%*</td>
<td>99.65%</td>
</tr>
<tr>
<td>Aunt – Niece (Grandmother – Grandchild Claim)</td>
<td>3</td>
<td>98.34%**</td>
<td>98.48%</td>
</tr>
<tr>
<td>Uncle – Nephew (Grandfather – Grandchild Claim)</td>
<td>2</td>
<td>99.52%</td>
<td>98.98%</td>
</tr>
</tbody>
</table>

* False Half Siblings averaged 0.019% when additional siblings are added across 2 tests
** False Grandmother-Grandchild with the mother included averaged 36.44% across 3 tests

Courtesy of Chris Miles from his talk given during RFIW at ACM MM 2017:
Miles, Christopher. "Rapid DNA Performance Results on Family Relationship Verification." Proceedings of the 2017 Workshop on Recognizing Families In the Wild. ACM, 2017
### New Technology is More Robust

<table>
<thead>
<tr>
<th>Relationship</th>
<th># of tests</th>
<th>13 Loci</th>
<th>20 Loci</th>
<th>24 Loci</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent-Child</td>
<td>10</td>
<td>99.98%</td>
<td>99.99996%</td>
<td>99.999997%</td>
</tr>
<tr>
<td>Full Siblings</td>
<td>5</td>
<td>99.97%</td>
<td>99.99996%</td>
<td>99.999996%</td>
</tr>
<tr>
<td>Half Siblings</td>
<td>3</td>
<td>73.64%</td>
<td>73.1%</td>
<td>96.88%</td>
</tr>
<tr>
<td>Grandparent-Grandchild</td>
<td>7</td>
<td>74.10%</td>
<td>88.72%</td>
<td>95.88%</td>
</tr>
<tr>
<td>Aunt/Uncle-Niece/Nephew</td>
<td>3</td>
<td>23.18%</td>
<td>36.53%</td>
<td>74.82%</td>
</tr>
</tbody>
</table>
Mass Casualty Applications

Courtesy of Chris Miles from his talk given during RFIW at ACM MM 2017:
Miles, Christopher. "Rapid DNA Performance Results on Family Relationship Verification." Proceedings of the 2017 Workshop on Recognizing Families In the Wild. ACM, 2017
Data Challenges

• Data challenge, Recognizing Families In the Wild (RFIW)
  – In conjunction with 2017 ACM MM
  – In conjunction with 2018 FG in China
  – In conjunction with 2019 FG in France (currently running, links below)

https://web.northeastern.edu/smilelab/RFIW2019/
Data Challenges

• Data challenge, *Recognizing Families In the Wild* (RFIW)
  – In conjunction with 2017 ACM MM
  – In conjunction with 2018 FG in China
  – In conjunction with 2019 FG in France (currently running, links below)

https://competitions.codalab.org/competitions/20196

https://competitions.codalab.org/competitions/20180
Demo Code on GitHub

https://github.com/visionjo/FIW_KRT

FIW API

Families In the Wild: A Kinship Recognition Toolbox. Visit FiW project page to download and learn more!

https://web.northeastern.edu/smilelab/fiw/

Version 1.0
Joseph P. Robinson, Ming Shao, Handong Zhao, Yue Wu, Timothy Gillis, Yun Fu Recognizing Families In the Wild (RFIW): Data Challenge Workshop in Conjunction with ACM MM 2017 Proceedings of the 2017 Workshop on Recognizing Families In the Wild, 2017
CONCLUSION

In Summary

- FIW:
  - More families, faces, and pairwise types
  - High quality family photos with text captions
  - Diverse families from around the globe
  - Supports multiple tasks

- Deep Learning can now be explored for kinship problems

What’s Next?

- Continue hosting RFIW as annual effort.
  - Participate in this year’s FG Challenge
- Improve models (ie push SOA), more data exploration, nature-based studies
- Design new problems with practical significance
Check all that’s happening at SMILE

https://web.northeastern.edu/smilelab/

https://twitter.com/SmileLabNEU
ABOUT YUN FU

- General Tutorial Chair
- SMILE Lab Founder and Director
  - Forefront in research on visual representation for applications of face & gesture recognition, vision-based HCI, human action/activity understanding.
  - >250 publications (including 7 books) and peer recognition, with >7,000 citations
  - Serves as associate editor, chairs, PC member and reviewer of many top journals and international conferences/workshops.
  - Recent Recognition:
    - 7 Prestigious Young Investigator Awards from NAE, ONR, ARO, IEEE, INNS, UIUC, Grainger Foundation;
    - 7 Best Paper Awards from IEEE, IAPR, SPIE, SIAM;
    - 3 major Industrial Research Awards from Google, Samsung, and Adobe, etc.
  - Currently Associate Editor of the IEEE Transactions on Neural Networks Systems (TNNLS).
  - He is fellow of IAPR, a Lifetime Senior Member of ACM and SPIE, Lifetime Member of AAAI, OSA, and Institute of Mathematical Statistics, member of Global Young Academy (GYA), INNS and Beckman Graduate Fellow during 2007-2008.

Top students are encouraged to apply to SMILE (PhD, Post-Doc)!

http://www1.ece.neu.edu/~yunfu/  http://www.northeastern.edu/smilelab/
SOME RESOURCES
(MANY OTHERS PROVIDED THROUGHOUT SLIDES)

➢ Organizers
  ▪ Joseph Robinson, Ming Shao, Yun Fu, SMILE Lab

➢ Datasets
  ▪ Cornell UB Kin KinWild I & II Family101 TS Kin SiblingDB Families In the Wild (FIW) UvA-NEMO Smile (videos)

➢ FIW Resources and current challenge
  ▪ RFIW 2019 at IEEE FG (https://web.northeastern.edu/smilelab/RFIW2019/)
    ▪ Track 1 Portal: https://competitions.codalab.org/competitions/20180
    ▪ Track 2 Portal: https://competitions.codalab.org/competitions/20196
  ▪ GitHub: https://github.com/visionjo/FIW_KRT/tree/master/sphereface_rfiw_baseline
  ▪ Works Published: https://web.northeastern.edu/smilelab/fiw/publications.html
Many Thanks!

To Recognize Families In the Wild: A Machine Vision Tutorial