

# Visual Kinship Recognition of Families in the Wild (FIW)

Joseph P. Robinson, Ming Shao, Yue Wu, Hongfu Liu, Timothy Gillis, and Yun Fu

Best viewed as PP slideshow (includes animations).

# Families in the Wild

## Motivation

- Automatic kinship recognition is a challenging feat
- Pre-existing datasets do not represent true data distributions
  - Poor image quality
  - Image quantity is lacking
  - Pair-wise samples do not properly reflect family tree structure
- Many factors are still undiscovered by the machine vision community
- Research has not yet reached reality



Search Investigations



Social Media



Genealogy Services/Research



Missing Children

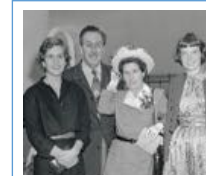
# Families in the Wild

## Database Overview

Type	KFW-I	KFW-II	Sibling Face	Group Face	FIW (Ours)
B-B	--	--	232	40	103,724
S-S	--	--	211	32	39,978
SIB	--	--	277	53	73,506
F-D	134	250	--	69	92,088
F-S	156	250	--	69	129,846
M-D	127	787	--	--	82,160
M-S	116	101	--	70	112,618
GF-GD	--	--	--	--	7,078
GF-GS	--	--	--	--	4,830
GM-GD	--	--	--	--	6,512
GM-GS	--	--	--	--	4,614
Total	533	1,000	720	395	656,954

Amount of pairwise types several orders of magnitude larger than predecessors, including four new grandparent – grandchild pairs

Dataset	No. Family	No. People	No. Faces	Age Varies	Family Structure	Highlights
CornellKin [5]	150	300	300	No	No	Parent-child pairs.
UB KinFace-I [8]	90	180	270	Yes	No	Parent-child pairs. Parents' 139 images at various ages.
UB KinFace-II [8]	200	400	600	Yes	No	Parent-child pairs. Parents' 139 images at various ages.
KFW-I [6]	--	533	1,066	No	No	Parent-child pairs.
KFW-II [6]	--	1,000	2,000	No	No	Parent-child pairs.
TSKinFace [9]	787	2,589	--	Yes	Yes	Two parents-child pairs for tri-verification.
Family101 [7]	101	607	14,816	Yes	Yes	Family structured, variations in age and ethnicity.
FIW(Ours)	1,000	10,676	30,725	Yes	Yes	Corpus of 1,000 family trees that provides both depth and breadth, along with multi-task evaluation offerings.



Disney



Freeman



Zardari



British Royal



Lee



Spanish Royal



Einstein



Japanese Royal



McGwire



Jackson 5



Affleck



King of Thailand



Gronkowski



Nicholas II



Roosevelt

Outperforms all predecessors in number of families, people, and faces

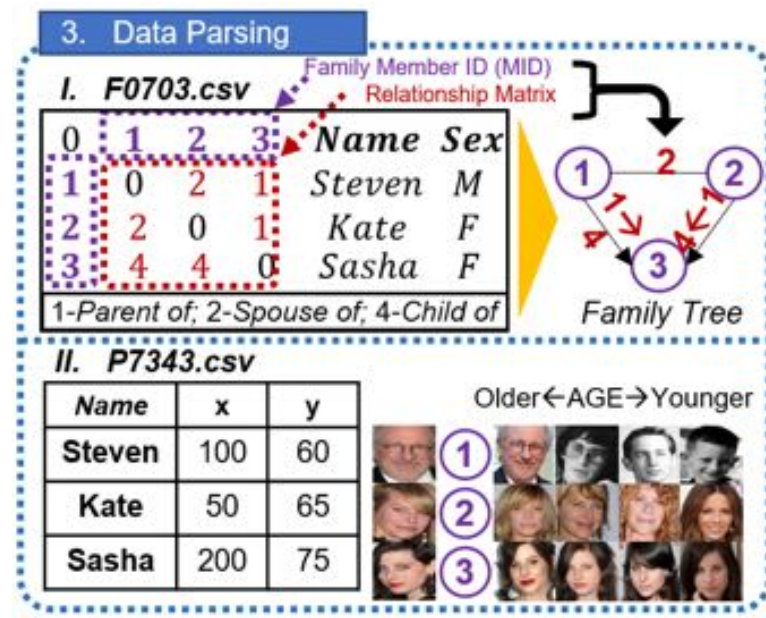
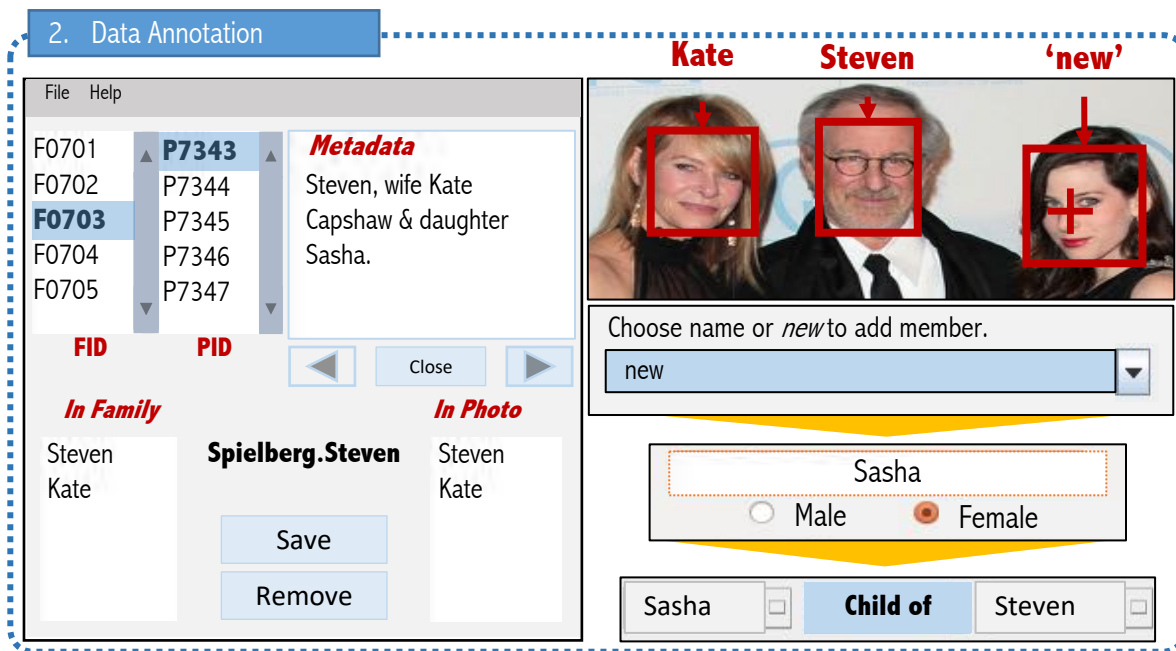
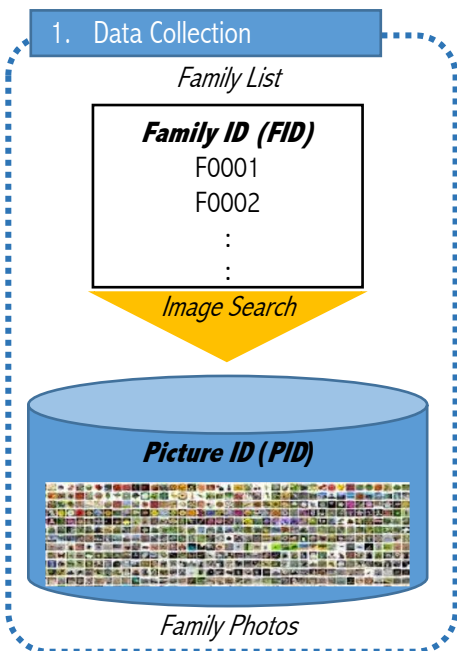
# Families in the Wild

## FIW 1.0 Annotation Summary

1. Collect family images from web

2. Annotate images using Java tool

3. Parse MID matrix for final incorporation into database.



Labeling process before our improved semi-automatic approach.



# Families in the Wild

## Objectives

1. Finalize incomplete families
  - Low resolution images
  - Less than 3 kin relationships
  - Small sample size
2. Add to families with larger online presence than represented in FIW 1.0
  - Royal Families
  - Politicians
3. Include very rich metadata
  - “Left to right”
  - Kinship keywords next to name

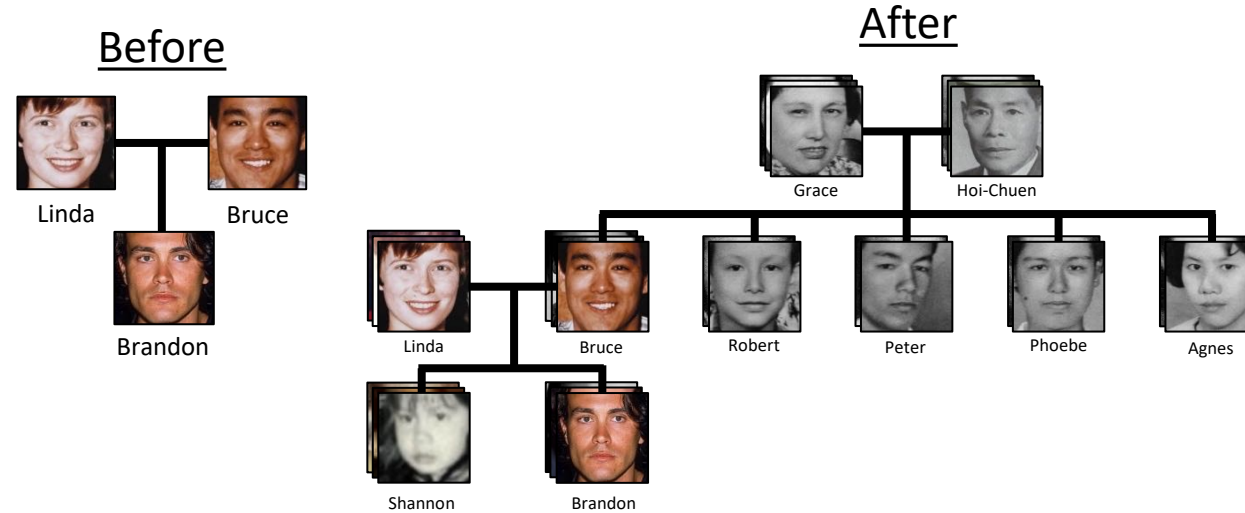


Figure 1: Bruce Lee family before and after, extending X in “depth” (number of faces) and X in “breadth” (number of relationships)

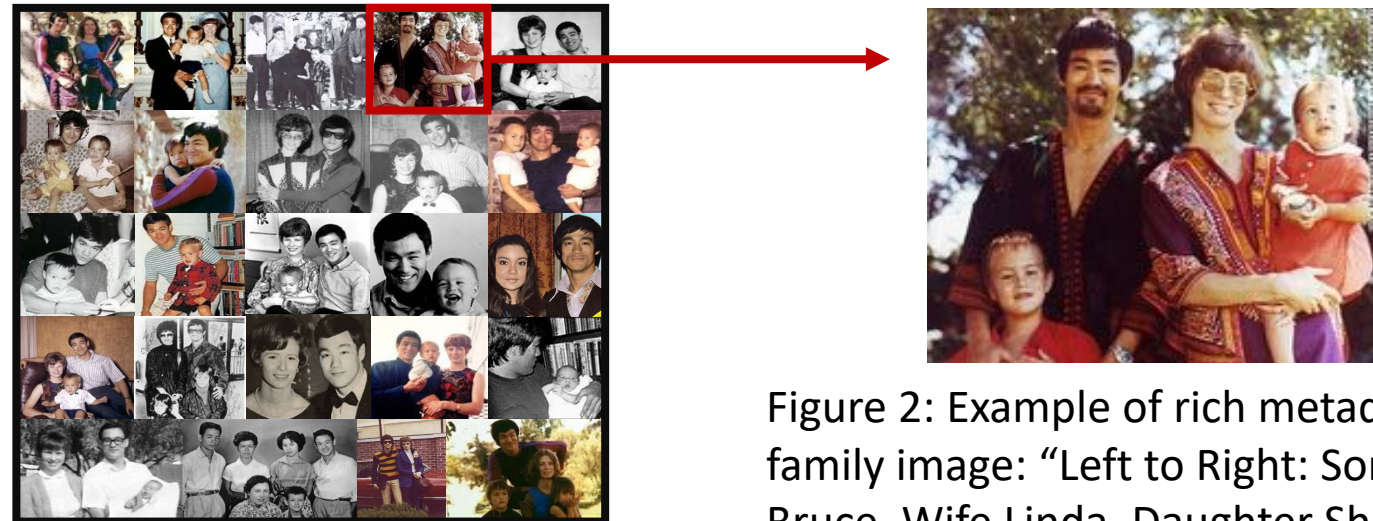
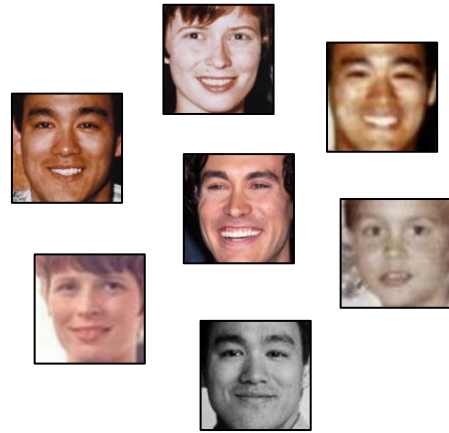


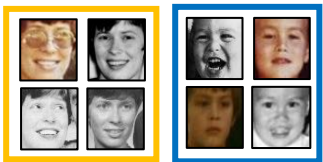
Figure 2: Example of rich metadata for family image: “Left to Right: Son Brandon, Bruce, Wife Linda, Daughter Shannon”

# Families in the Wild

Semi-Atomic Clustering



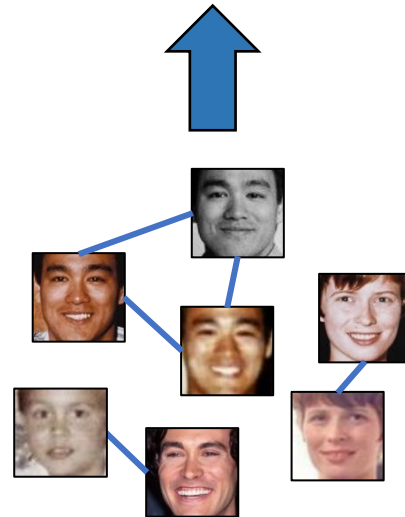
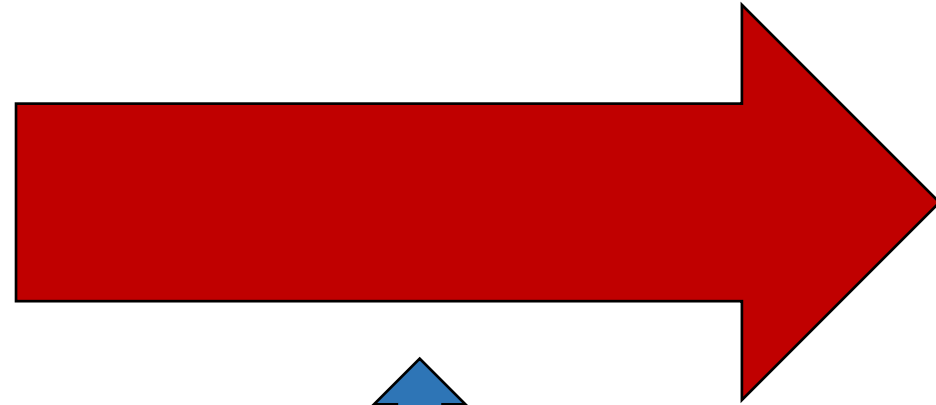
Unlabeled Faces



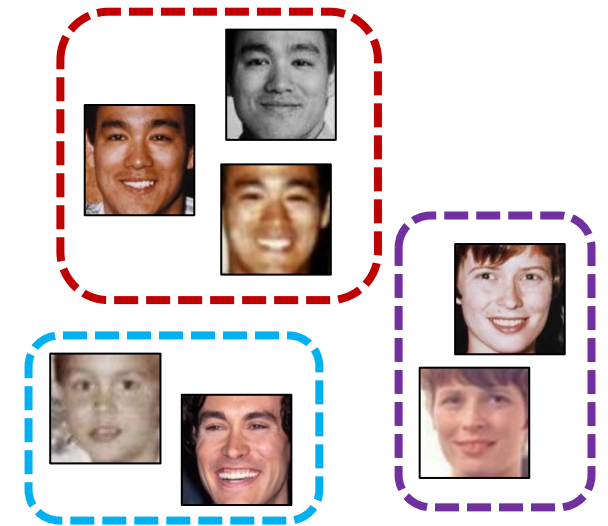
True Labels



Portrait Photos



Clustering



Proposed Labels

# Families in the Wild

Semi-Atomic Clustering





# Families in the Wild

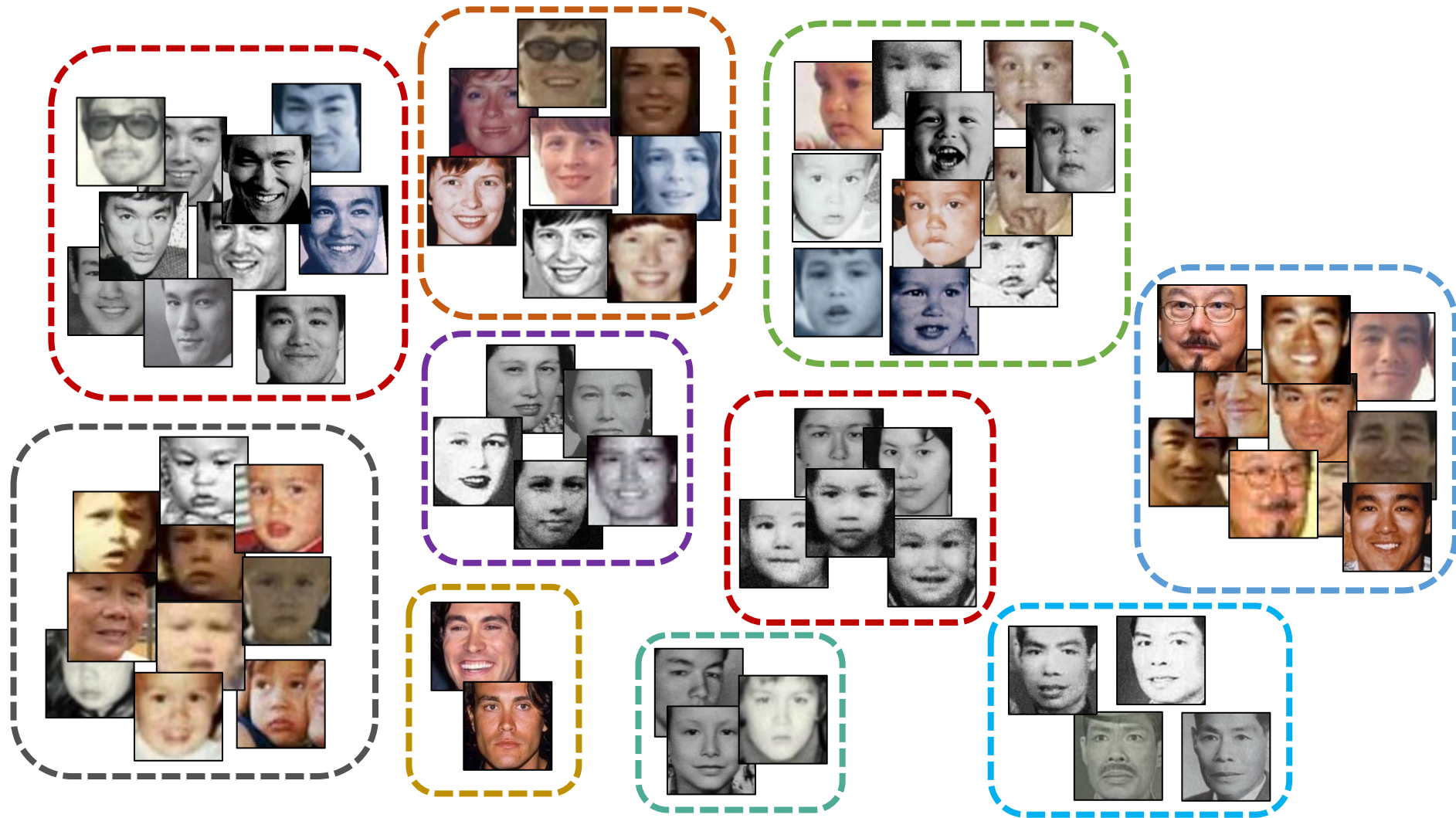
Semi-Atomic Clustering





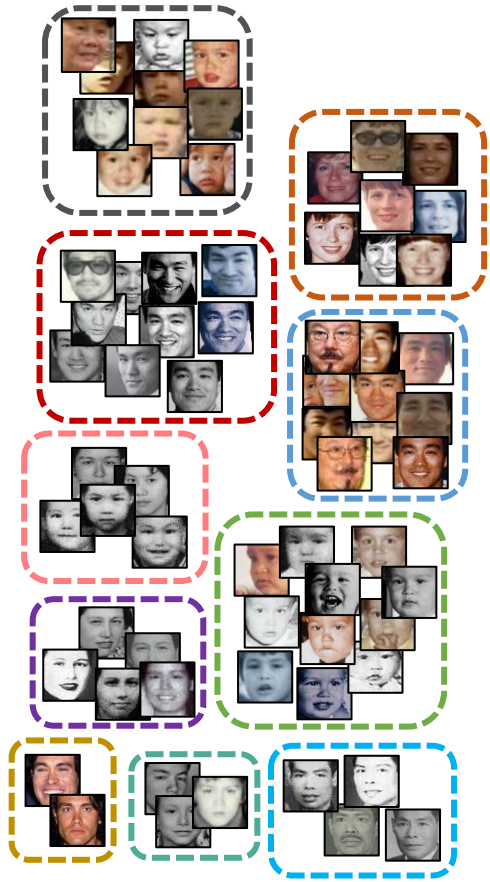
# Families in the Wild

Semi-Atomic Clustering

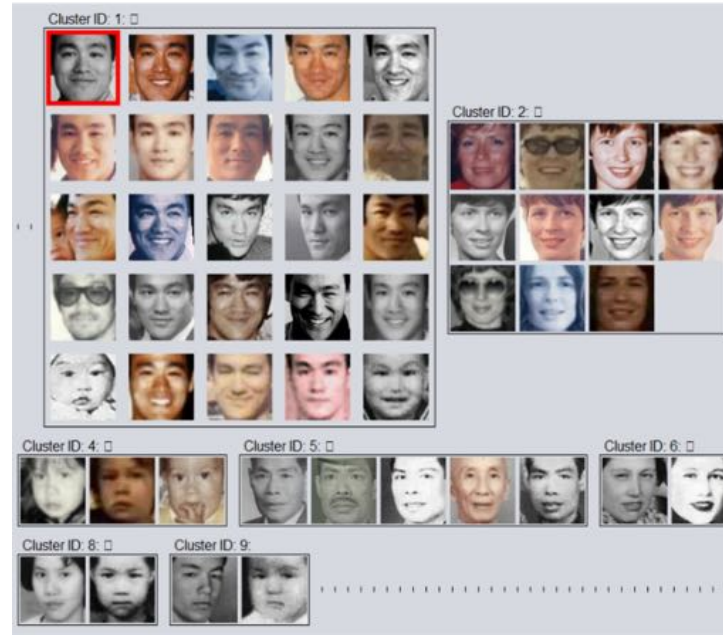
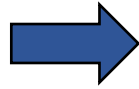


# Families in the Wild

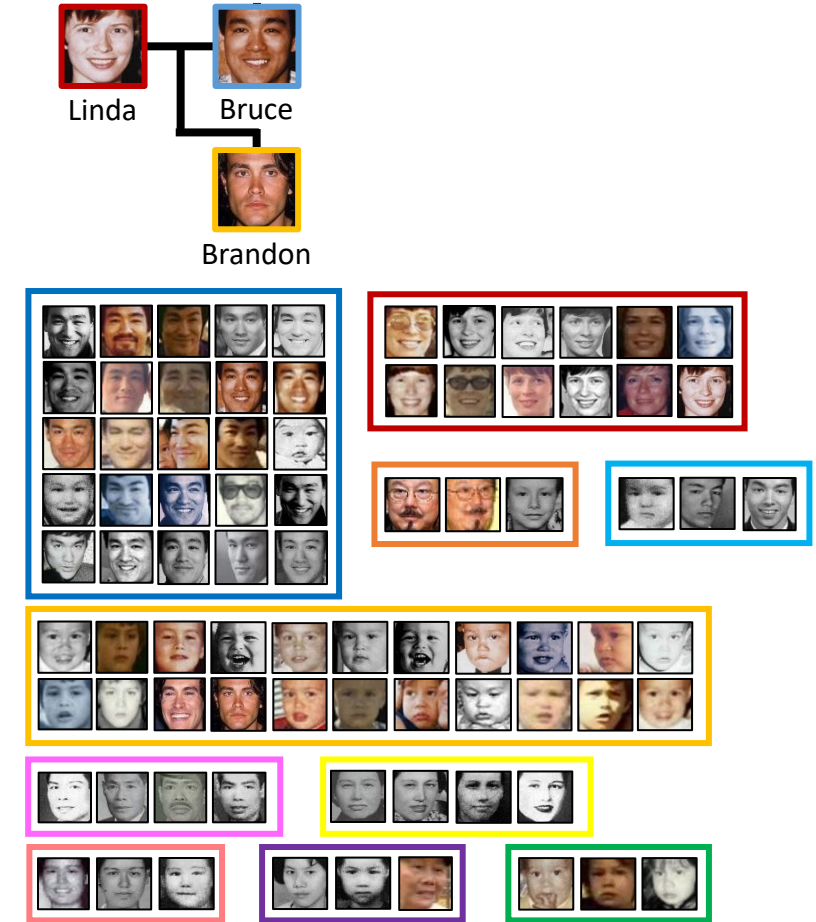
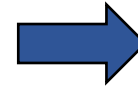
Annotation



Input: Proposed Labels (Clusters)



Cluster Validation GUI



Output: Labeled samples added to Families in the Wild

# Families in the Wild

Annotation



See video demo included as supplemental material.



# Families in the Wild

## Annotation Performance

### Old Performance [Table 1]:

Average Time per Image: **10min 13sec**

Average Clicks per Image: **40.67 clicks**

### New Performance [Table 2]:

Average Time per Image: **8.68s**

Average Clicks per Image: **4.01 clicks**

### Conclusions:

Time per Image decreased by **~601%**

Clicks per Image decreased by **~10%**

FID	Images	Total Clicks	Total Time
F0378	5	97	0:5:31
F0425	35	551	0:15:08
F0458	10	153	0:5:18
F0480	13	178	0:6:16
F0624	128	1,838	1:25:23

**Table 1:** Annotation statistics of the 5 family subset using FIW 1.0 labeling tool (time in h:m:s)

FID	Images	Total Clicks	Total Time (s)
F0378	5	37	0:0:31
F0425	35	12	0:1:11
F0458	10	10	0:1:05
F0480	13	15	0:0:56
F0624	128	21	0:6:44

**Table 2:** Annotation statistics of the 5 family subset using Clustering Verification GUI (time in h:m:s)

# Families in the Wild

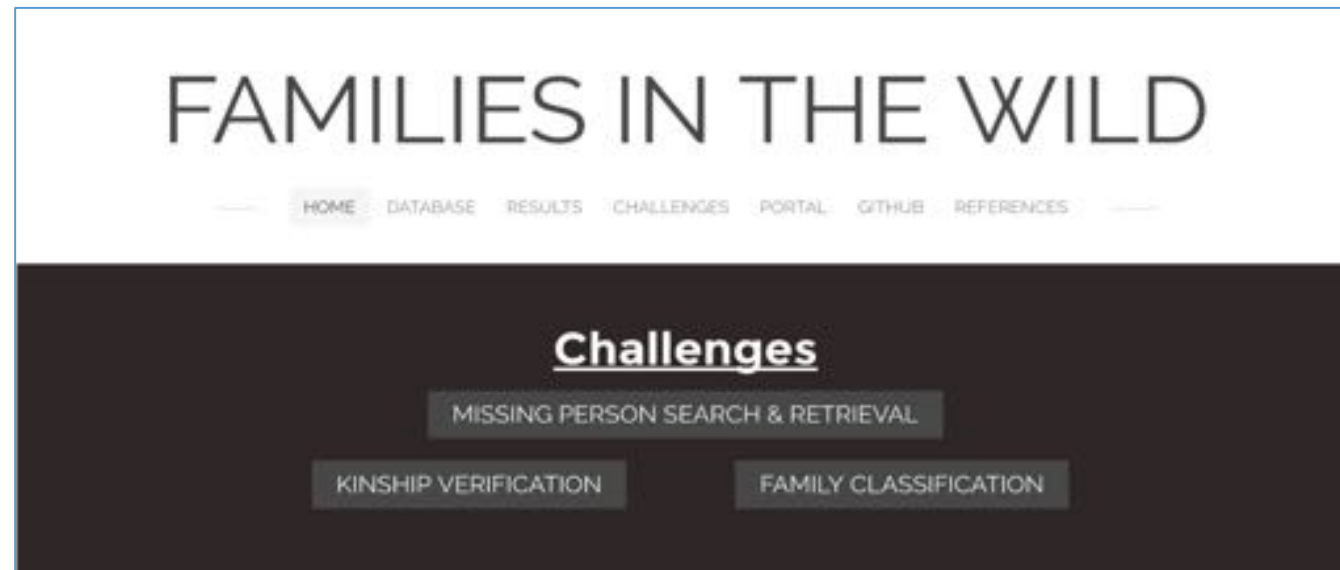
## Discussion

### In Summary

- Pre-existing kinship datasets do not represent true data distributions
- Families in the Wild greatly outperforms in number of families, faces, and pairwise types
- Extending the database

### What's Next

- Organize kinship recognition into challenge; support additional tasks.
- Improve models and perform more intensive analysis.



<http://smile-fiw.weebly.com/>