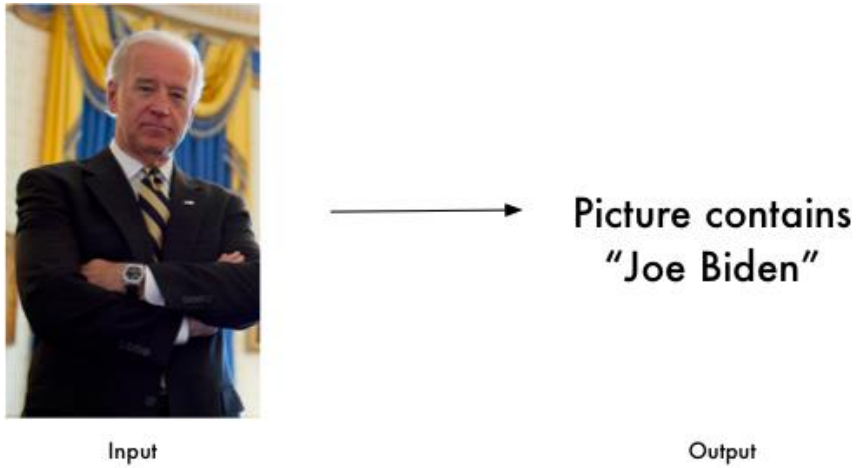
Face Verification in blurred and profile faces

Amir mohammad Salehoof

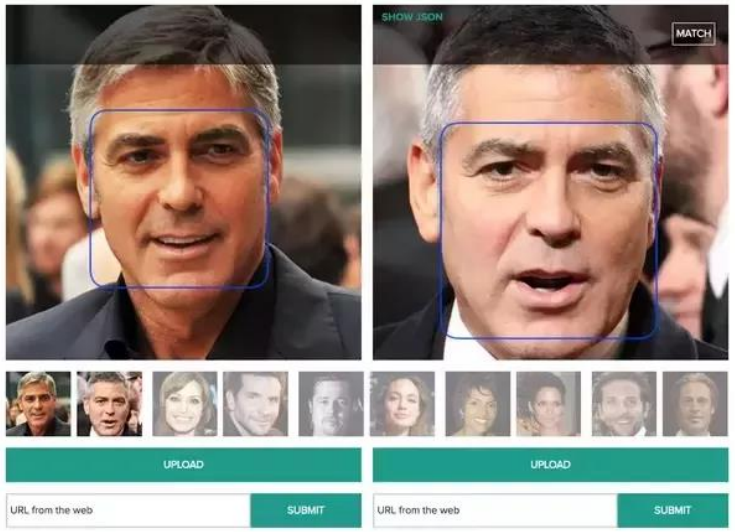
Mohammad hasan Shamgholi

Face verification vs face recognition

- Face Recognition



- Face Verification



Problem

- The achieved accuracy of face verification is
  - high on simple and non-noisy datasets
  - Low on harder detection cases.

- In most case pictures in datasets gathered:
  - From Google
  - By high resolution camera

  In stable condition with no blurring (of any kinds)

Problem (Cont'd)

- In real-world applications we have default blurriness

- default blurriness caused by many reasons:
  - motion blur
  - image/video compression
  - face profile changes

Solution

- Gather dataset with faces in hard situations.
- Find a model that trained on large face dataset.
- Fine-tune pre-trained model with gathered data.

Dataset (VoxCeleb2)

- VoxCeleb2 contains over 1 million utterances for 6,112 celebrities, extracted from videos uploaded to YouTube.

Speech Separation

Emotion Recognition

VoxCeleb2 (CONT'd)

- Laplacian Operator for blurriness
- 25 faces from low values and 25 faces from high values for each person

High value

Low value

Dataset(LFW)

- 13233 images
- 5749 people
- 1680 people with two or more images
- Use for validation

Dataset(YTF)

- YouTube Faces Database: a database of face videos designed for studying the problem of unconstrained face recognition in videos.
- 3,425 videos
- 1,595 different people.
- The shortest clip duration is 48 frames,
- The longest clip is 6,070 frames
- The average length of a video clip is 181.3 frames.

Pre-trained model

- MobileNet
- Architecture:

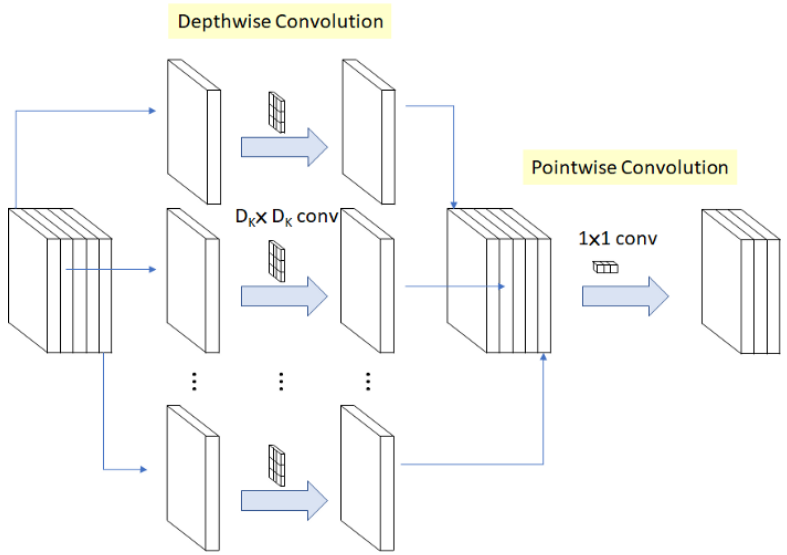


Table 1. MobileNet Body Architecture

Type / Stride	Filter Shape	Input Size
Conv / s2	$3 \times 3 \times 3 \times 32$	$224 \times 224 \times 3$
Conv dw / s1	$3 \times 3 \times 32$ dw	$112 \times 112 \times 32$
Conv / s1	$1 \times 1 \times 32 \times 64$	$112 \times 112 \times 32$
Conv dw / s2	$3 \times 3 \times 64$ dw	$112 \times 112 \times 64$
Conv / s1	$1 \times 1 \times 64 \times 128$	$56 \times 56 \times 64$
Conv dw / s1	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 128$	$56 \times 56 \times 128$
Conv dw / s2	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 256$	$28 \times 28 \times 128$
Conv dw / s1	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 256$	$28 \times 28 \times 256$
Conv dw / s2	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 512$	$14 \times 14 \times 256$
$5\times$ Conv dw / s1	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
$5\times$ Conv / s1	$1 \times 1 \times 512 \times 512$	$14 \times 14 \times 512$
Conv dw / s2	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
Conv / s1	$1 \times 1 \times 512 \times 1024$	$7 \times 7 \times 512$
Conv dw / s2	$3 \times 3 \times 1024$ dw	$7 \times 7 \times 1024$
Conv / s1	$1 \times 1 \times 1024 \times 1024$	$7 \times 7 \times 1024$
Avg Pool / s1	Pool 7×7	$7 \times 7 \times 1024$
FC / s1	1024×1000	$1 \times 1 \times 1024$
Softmax / s1	Classifier	$1 \times 1 \times 1000$

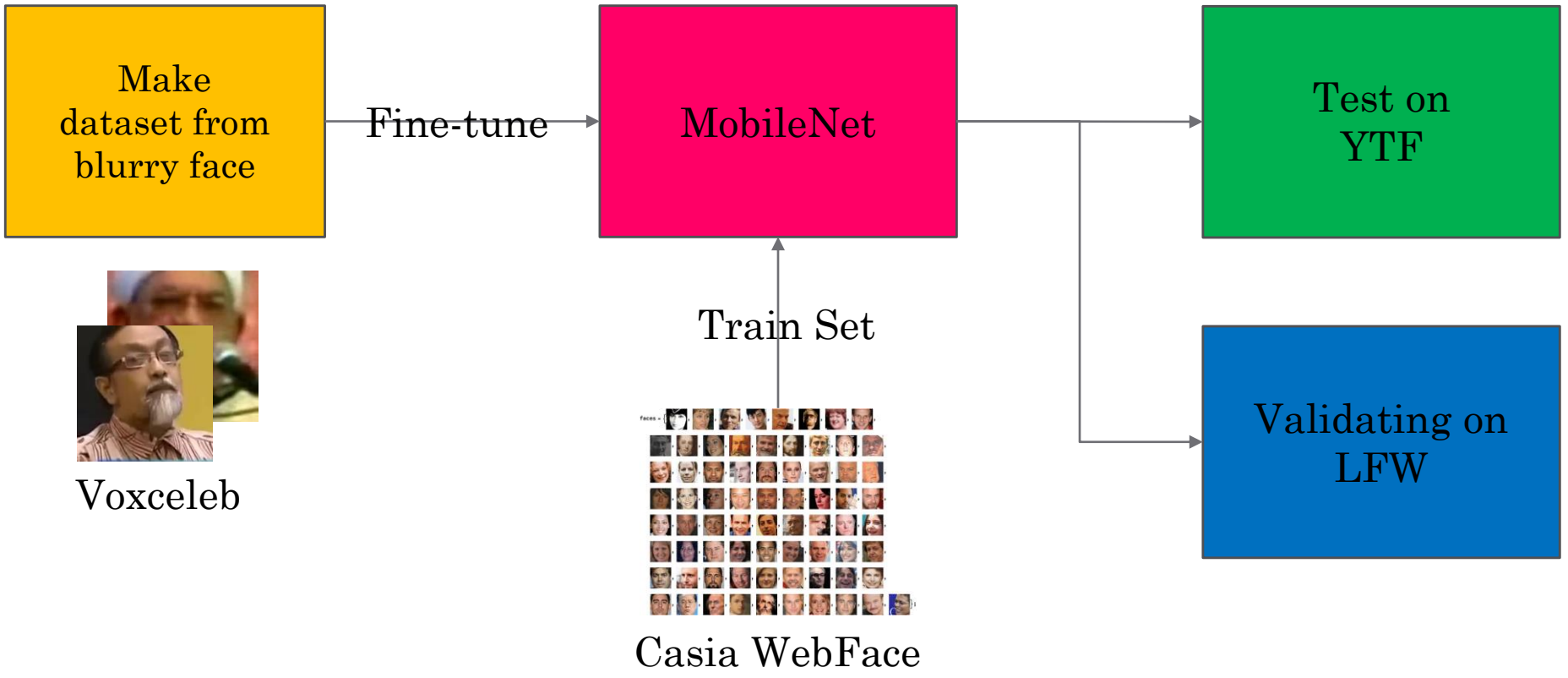
- Trained on Casia WebFace (dataset of 453,453 images over 10,575 identities after face detection)

Damaging pre-trained MobileNet

- After fine-tuning new data, we may damage model's accuracy on its test set
- Validate model on LFW

LFW

Work Overview



Results

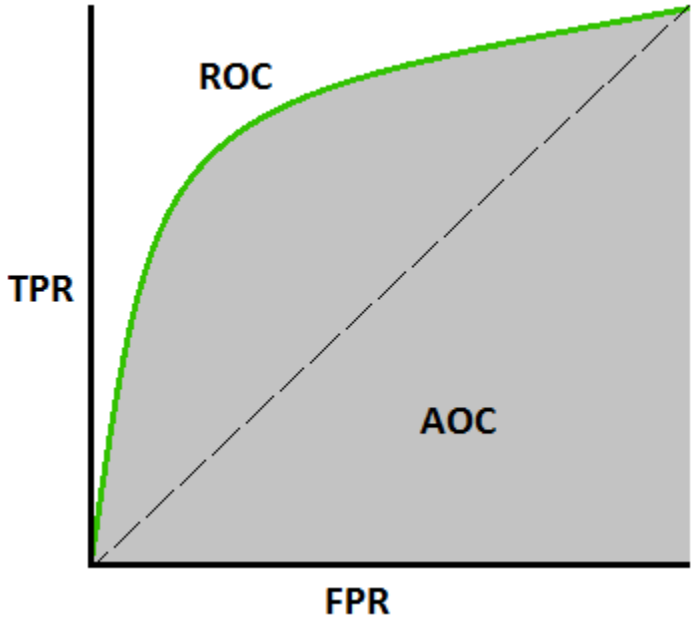
- Metric: ROC (Receiver Operating Characteristic)

$$\text{TPR /Recall / Sensitivity} = \frac{TP}{TP + FN}$$

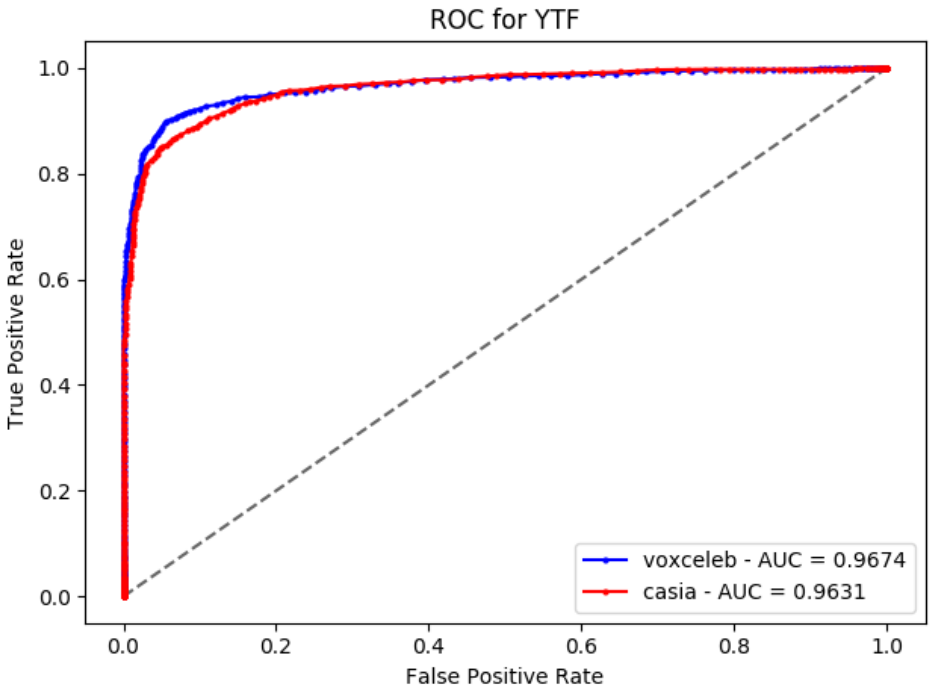
$$\text{FPR} = 1 - \text{Specificity}$$

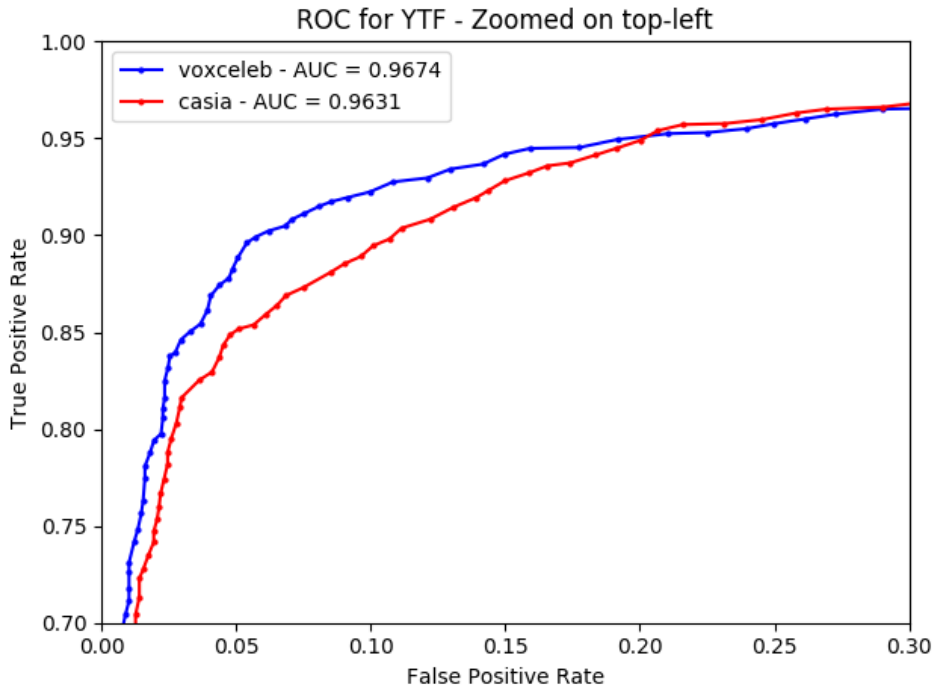
$$= \frac{FP}{TN + FP}$$

$$\text{Specificity} = \frac{TN}{TN + FP}$$

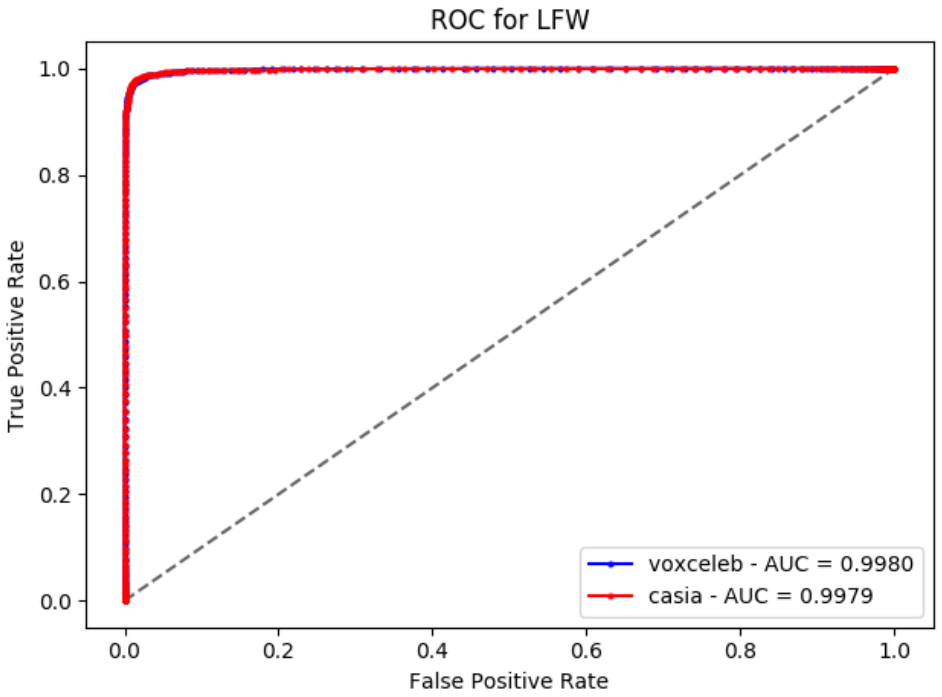


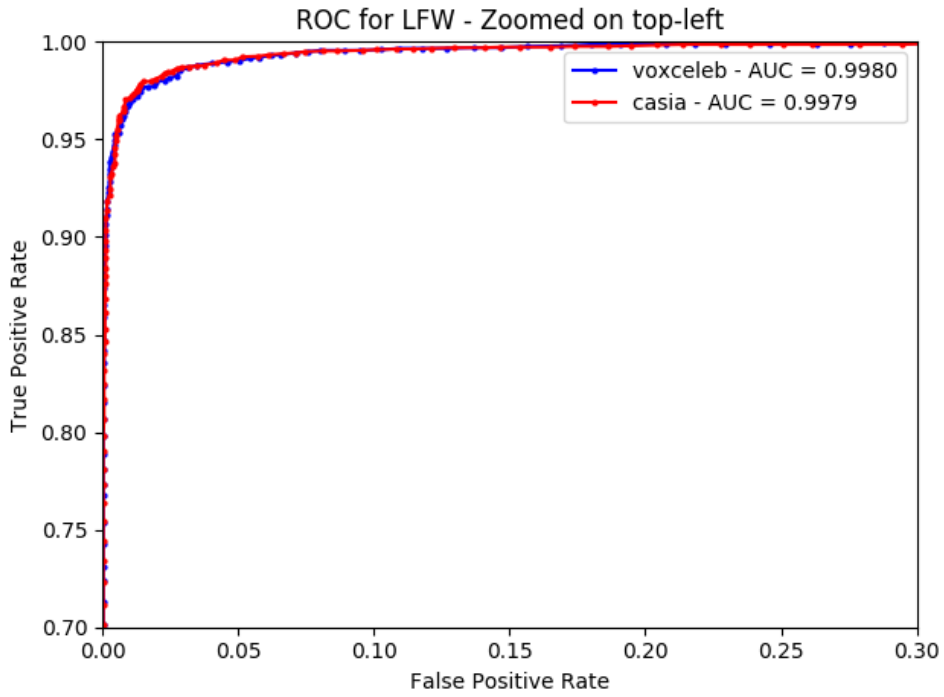
Result (Cont'd)



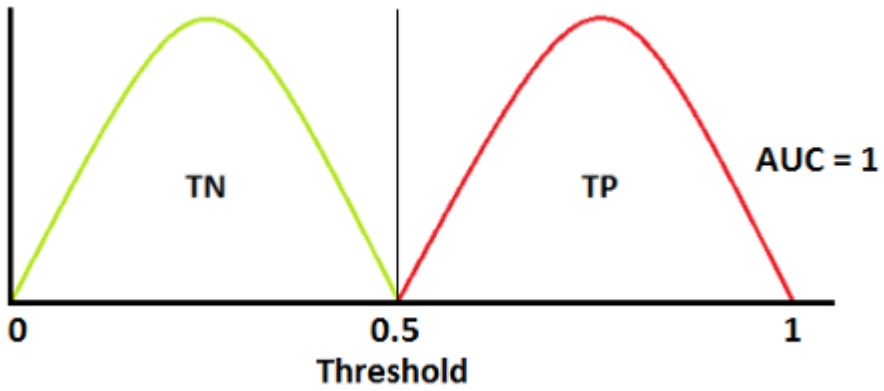


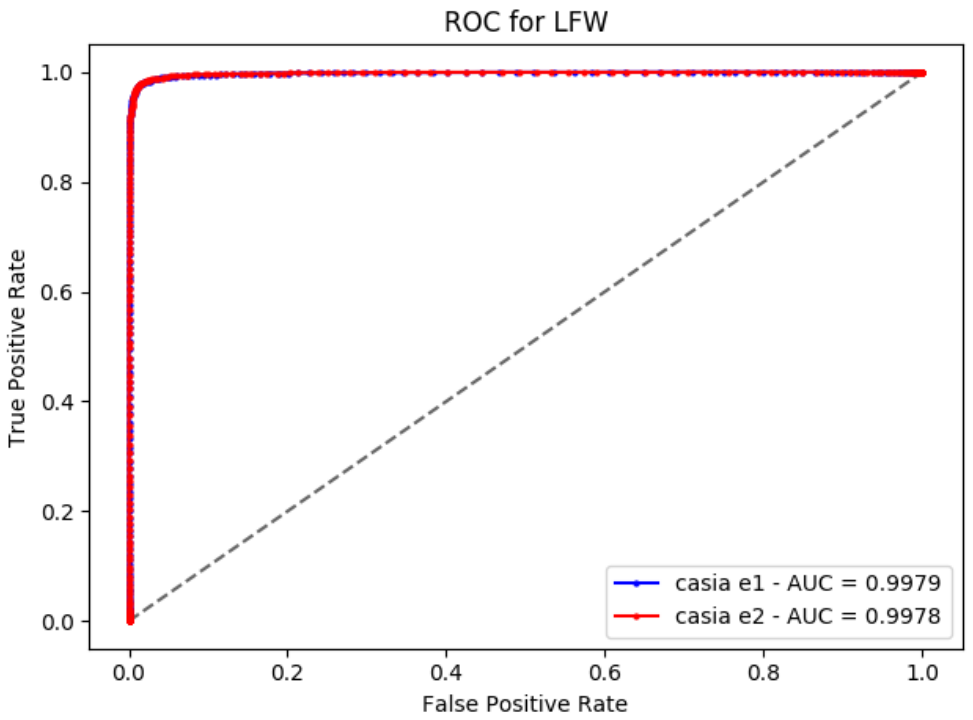
Result (Cont'd)

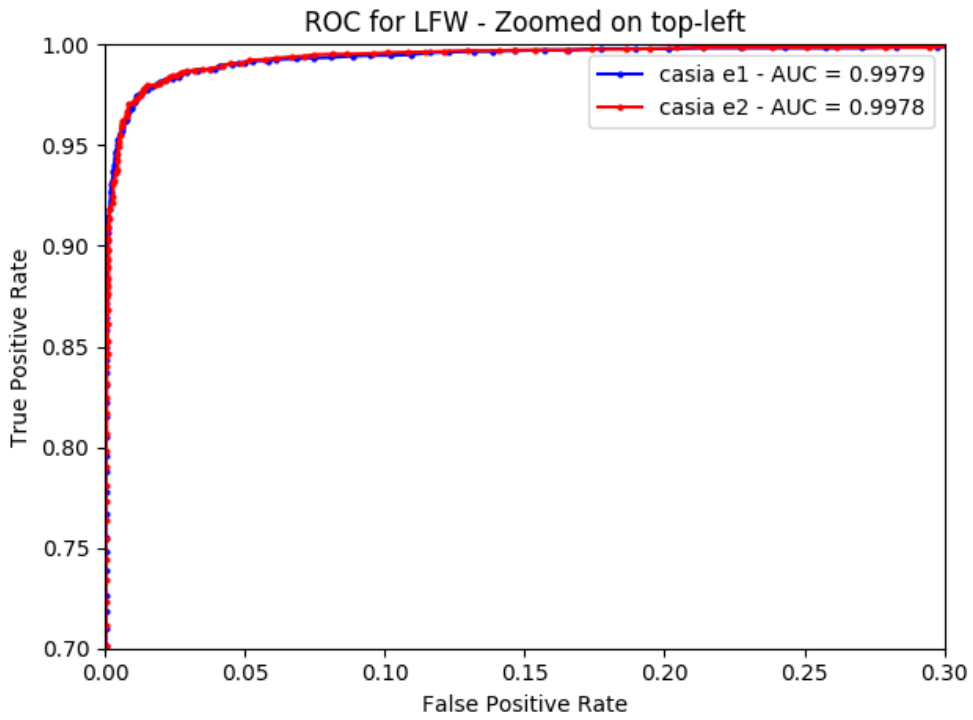


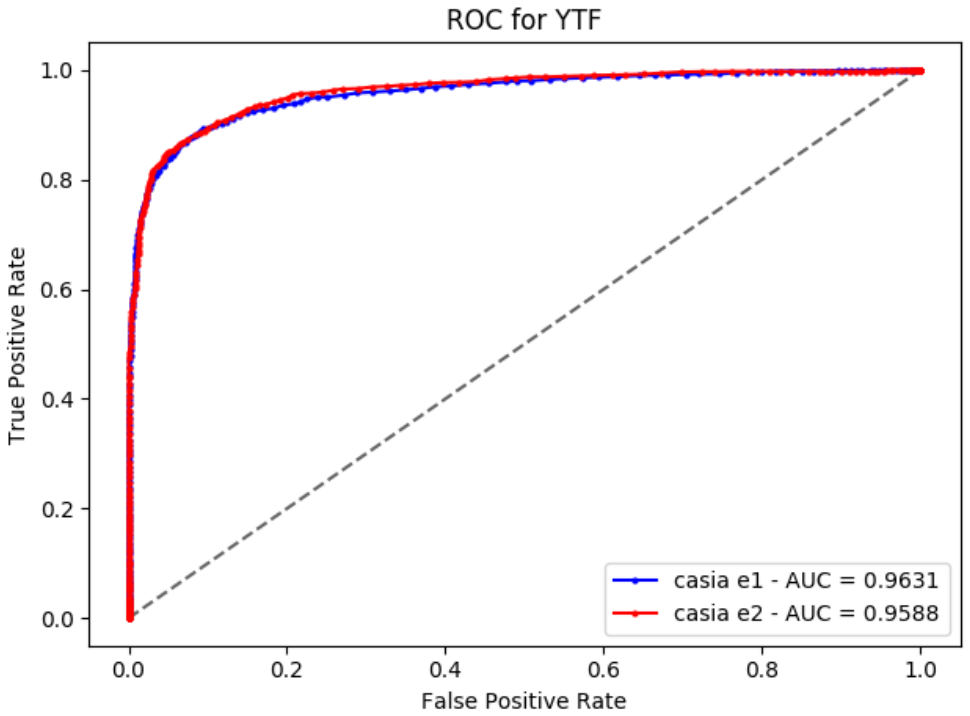


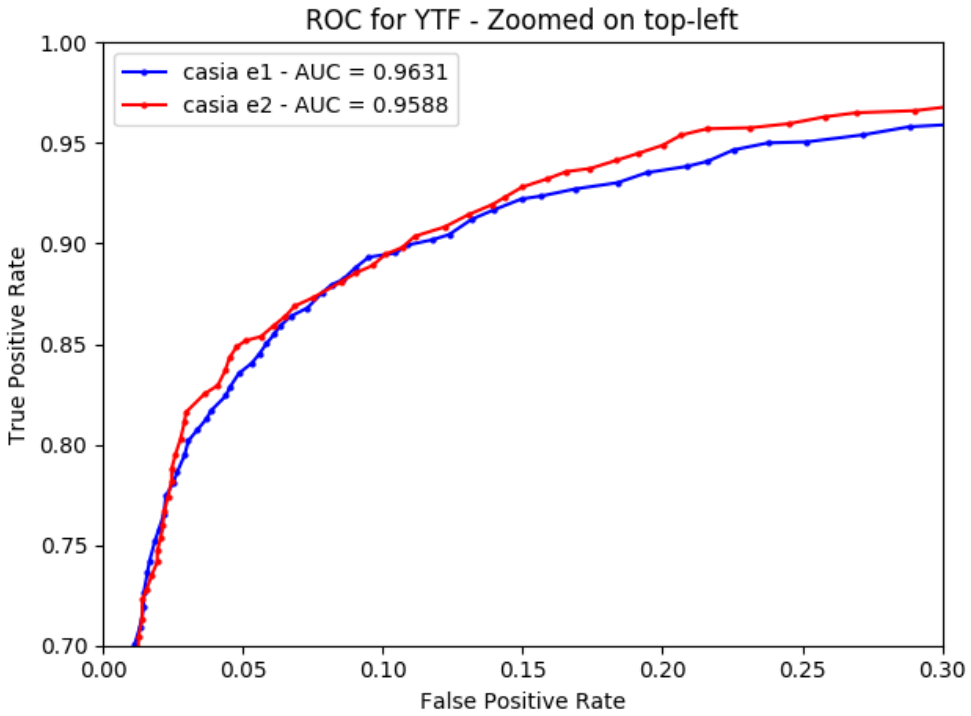
THANK YOU


TN
TP
AUC = 1
0
0.5
1
Threshold


ROC for LFW
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
casia e1 - AUC = 0.9979
casia e2 - AUC = 0.9978


ROC for LFW - Zoomed on top-left
True Positive Rate
False Positive Rate
0.70
0.75
0.80
0.85
0.90
0.95
1.00
0.00
0.05
0.10
0.15
0.20
0.25
0.30
casia e1 - AUC = 0.9979
casia e2 - AUC = 0.9978


ROC for YTF
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
casia e1 - AUC = 0.9631
casia e2 - AUC = 0.9588


ROC for YTF - Zoomed on top-left
True Positive Rate
False Positive Rate
0.70
0.75
0.80
0.85
0.90
0.95
1.00
0.00
0.05
0.10
0.15
0.20
0.25
0.30
casia e1 - AUC = 0.9631
casia e2 - AUC = 0.9588