Creating a Computationally Efficient Embedded CNN Face Recognizer

G.B.Praveen

23-May-2018

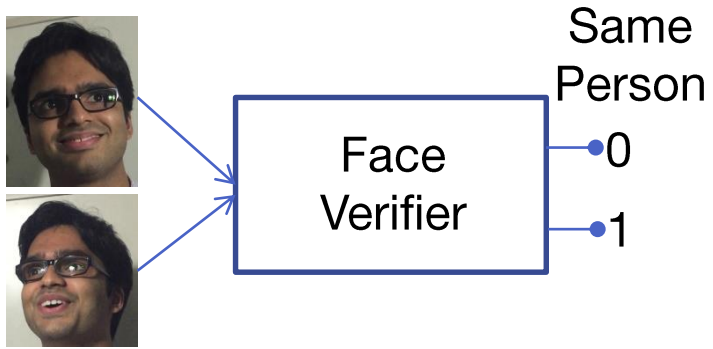
PATHPARTNER

Agenda

- Face Verification
  - Introduction
  - Challenges Involved
  - Existing Literature
  - Proposed Method & Results
- Face Recognition
  - Introduction
  - Limitation in Using Non-CNN Approaches
  - PathPartner FaceRecognition Dataset
  - Face Recognition Tracker & Results

Introduction

- Face verification: Prediction of two images belonging to same person



Challenges involved in Face Verification

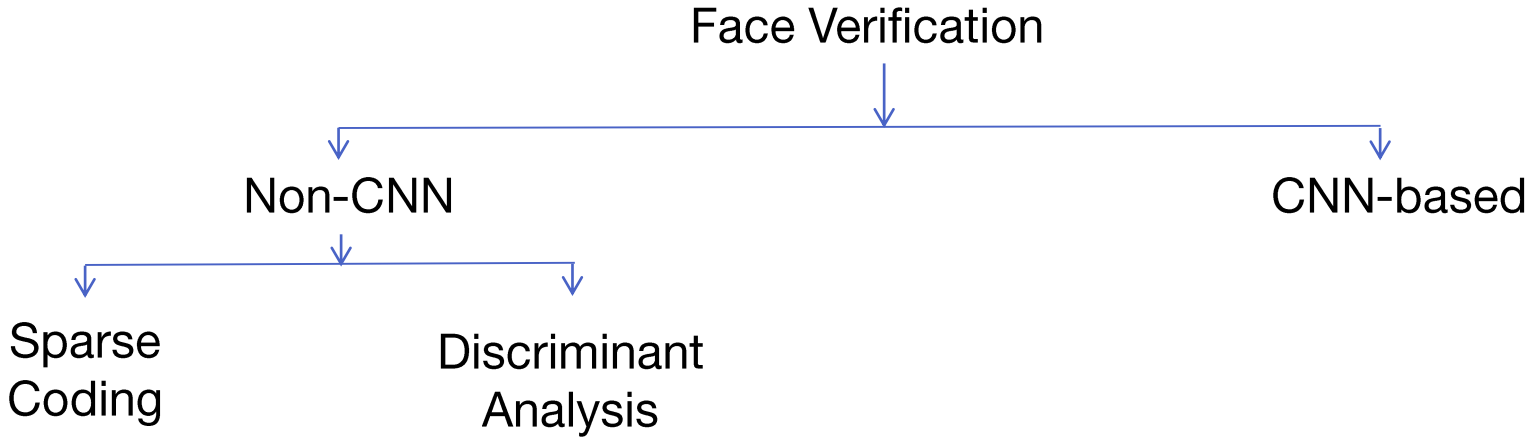
- Illumination Variations
- Pose Variations
- Expression Variations
- Occlusion
- Appearance Variations

Expression Variations

Illumination Variations

Pose Variations

Approaches for Face Verification



Limitation in using non-CNN approach:

- Need additional methods to handle illumination, occlusion & pose variations

CNN for Face Verification

- Objective: Should be discriminative across subjects and invariant enough to intra-class variations, i.e., invariant to pose, expression, illumination and occlusion of same subject
- Feature space should be invariant to pose, expression, illumination and occlusions of same person

Introduction to Loss Function

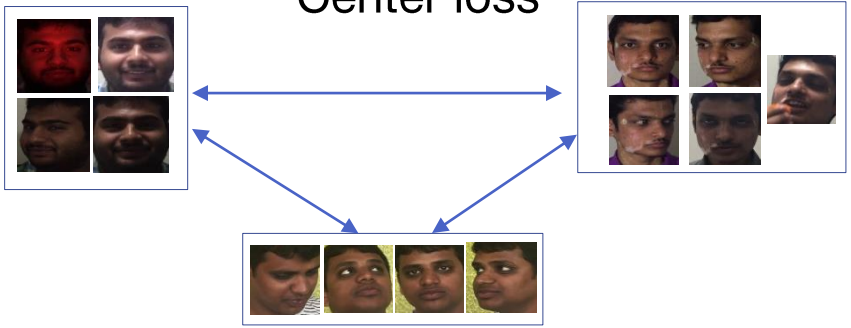
SoftMax loss

- Limitation :
  - Less effective for intra-class variations

Triplet loss

$$Loss_{Triplet} = D_{ap} - D_{an} + Margin$$

Center loss

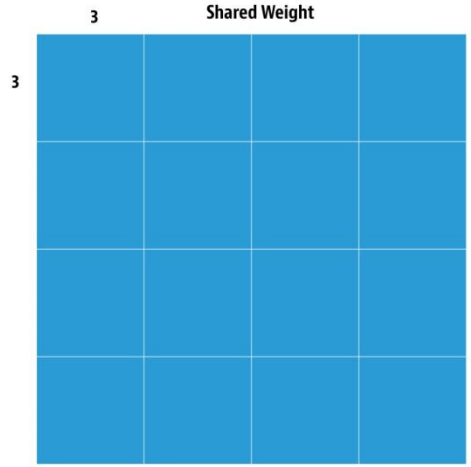


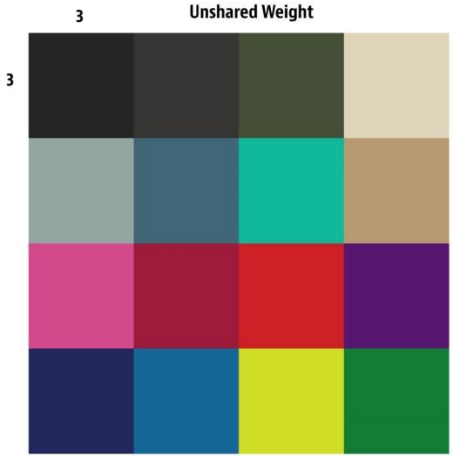
$$Loss_{Center} = Distance(FV_{inp}, FV_{c(inp)})$$

$$Loss = Loss_{SM} + (\lambda * Loss_{Center})$$

Locally connected layer with unshared weights

Filter : 3x3
Input Channels : 10
Output Channels : 20
Total Params = 3x3x10x20





Filter : 3x3
Input Channels : 10
Output Channels : 20
Number of filtering ops : 16
Total Params = 3x3x10x20x16

- Limitations of using locally connected (LC) layer:
  - Higher model size
  - Higher memory bandwidth
  - Using deeper networks with combination of unshared weights

Test Dataset : LFW (Labelled Faces in Wild)

- LFW is used as test set for Face Verification
- 5,749 people (1,680 people with 2 or more images)
- 13,233 images
- 6,000 pairs used in test set

Existing Literature – CNN-based approaches

Unshared Weights based approaches:

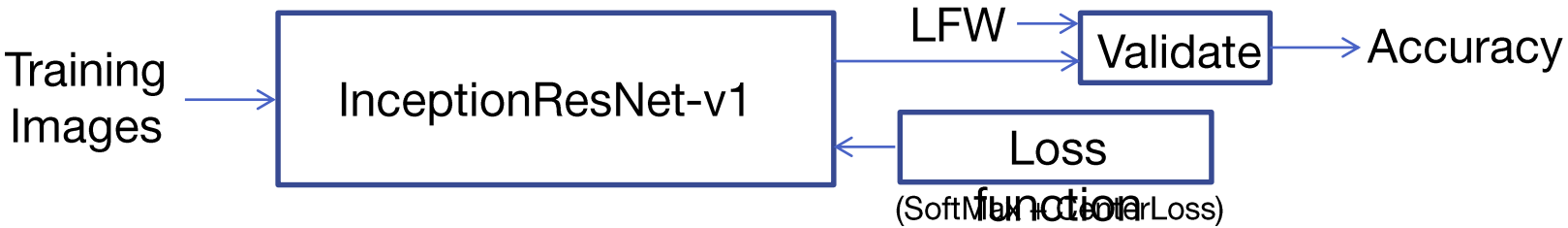
Literature Title	DeepFace (Taigman *et al.*)	Identity Preserving Face Space (Zhu *et al.*)	Discriminative Feature Learning (Yandong *et al.*)
CNN used	2Conv+1P+3LC+1FC (Frontalized input face image)	3LC+1P+1FC	3Conv+4P+3LC+ 1FC
Training info	SoftMax loss	L2 loss(input, canonical of same subject)	SoftMax + Center Loss
Inference info	Exp-1: FC layer Exp-2: $\chi 2$ method		L2-distance
Accuracy	LFW: 97.35%	Multi-PIE: 94-98%	LFW: 99.28%

Existing Literature – CNN-based approaches

Too-Deep network based approaches :

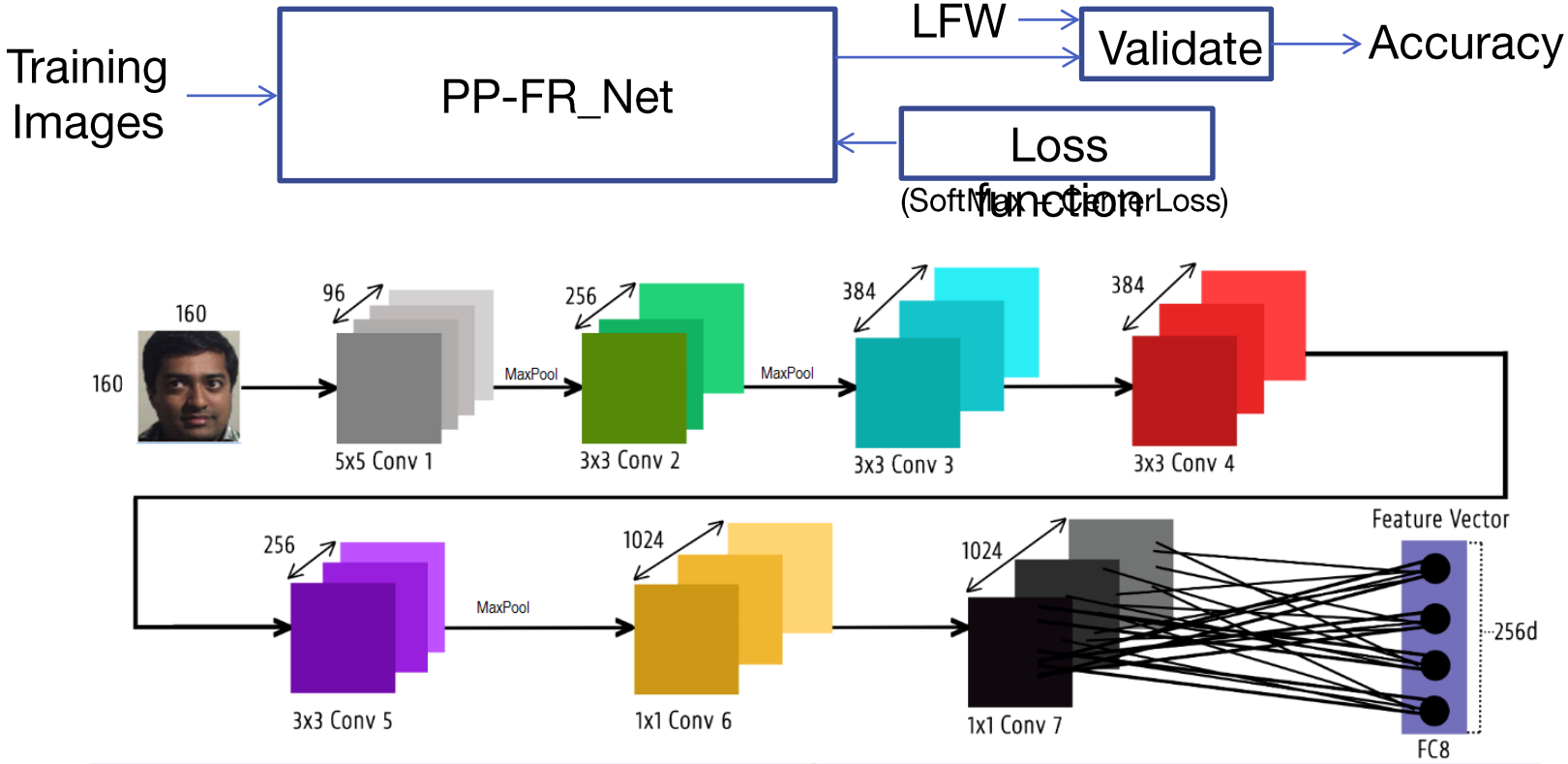
Literature Title	Deep Attributes (Aishwarya *et al.*)	DeepFace (Parkhi *et al.*)	FaceNet (Schroff *et al.*)	FaceNet (David Sandberg)
CNN used	VGG Face (13Conv+5P+3FC)	13Conv+5P +3FC	11Conv+3P+3FC	**InceptionResNet-v1**
Training	SoftMax loss on 8-attributes (5local + 3global)	Pre-train: SoftMax loss Fine-tune: Triplet loss	Triplet loss	**SoftMax+ CenterLoss**
Inference	Generate 8x4096 Feature Vector	L2-distance	K-means clustering on feature vector	**L2-distance**
Accuracy	LFW: 93%	LFW: 98.95%	LFW: 99.63%	**LFW: 99.05%**

Benchmark Model

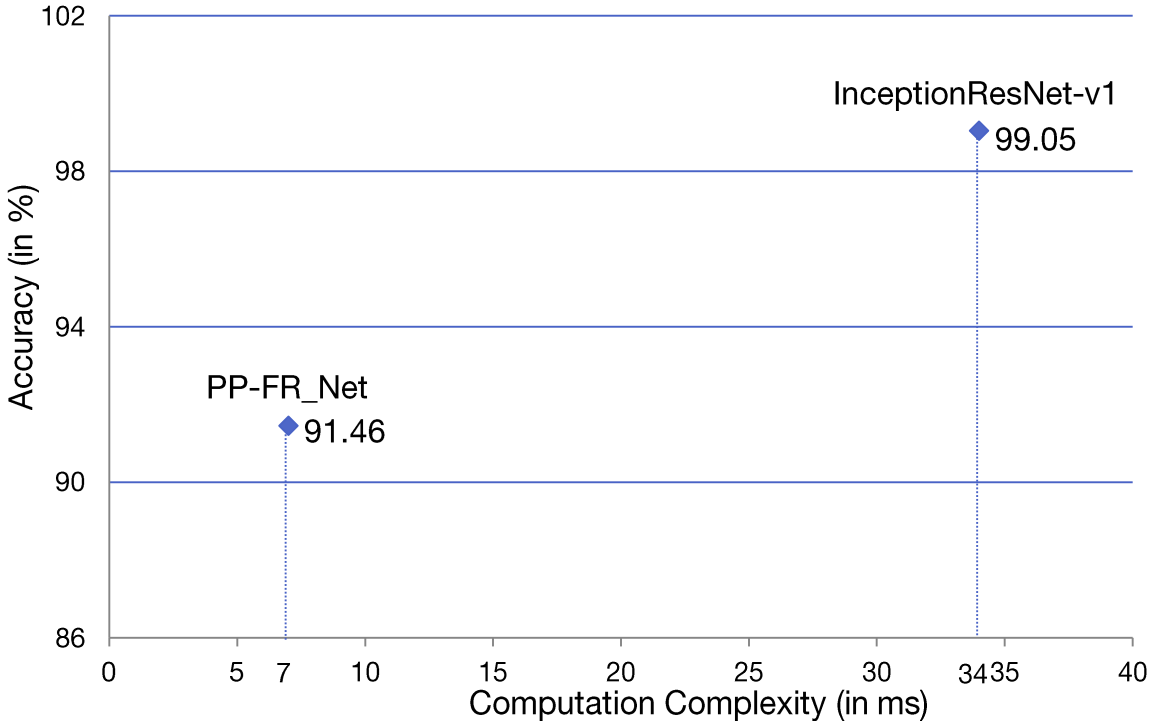


- Accuracy on LFW: 99.05%
- Complexity: 34ms (on TitanX Pascal)

Proposed Network

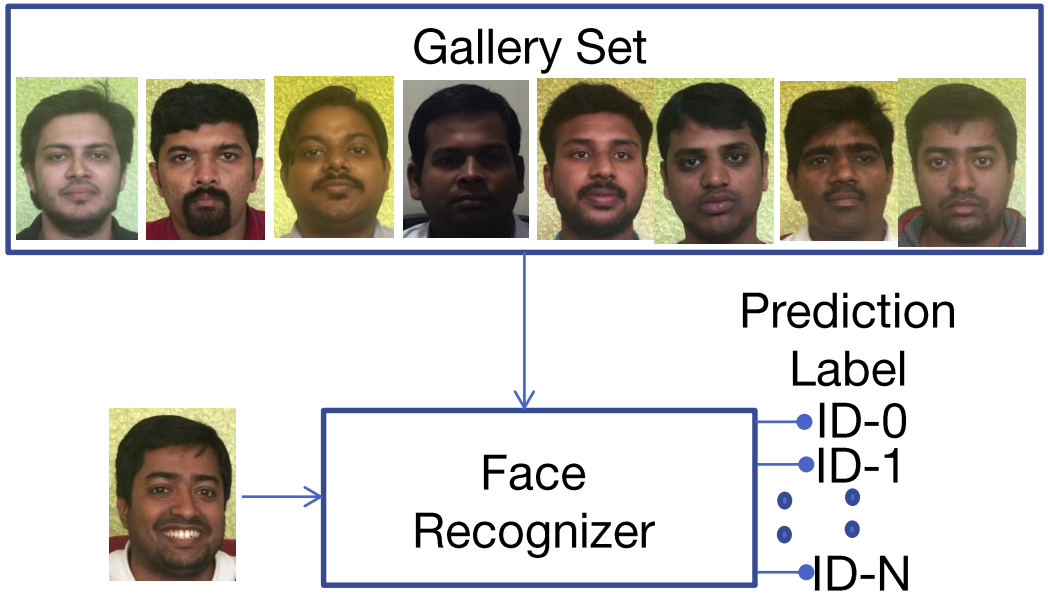


InceptionResNet-v1 Vs PP-FR_Net



Introduction – Face Recognition

- Face identification/recognition/retrieval: Identifying given test sample from “N” registered images

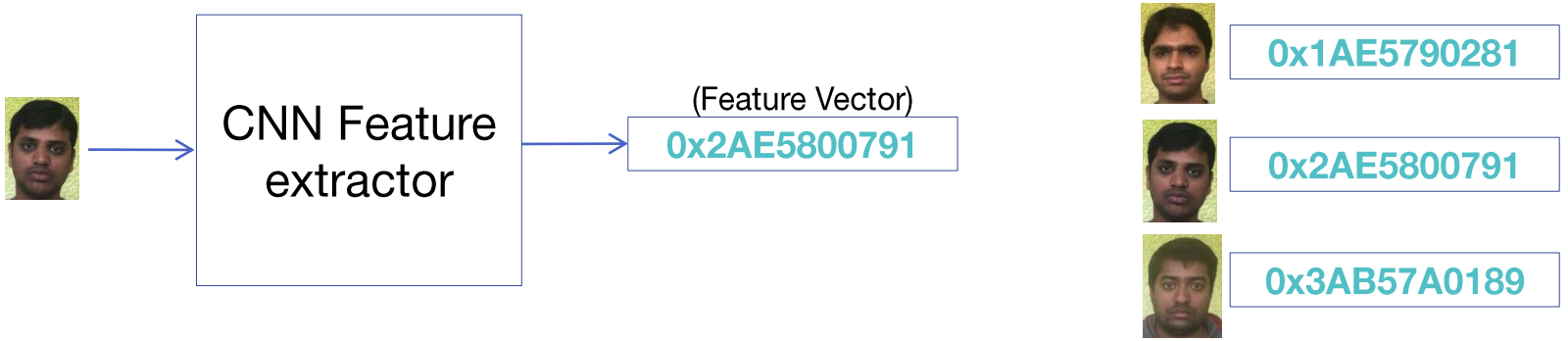


Limitations of non-CNN approaches

- Registration: Multiple images per subject are required
- Larger memory to store representation of registered subjects
- Limited scalability on number of registered users
- Need additional methods to handle illumination, occlusion & pose variations

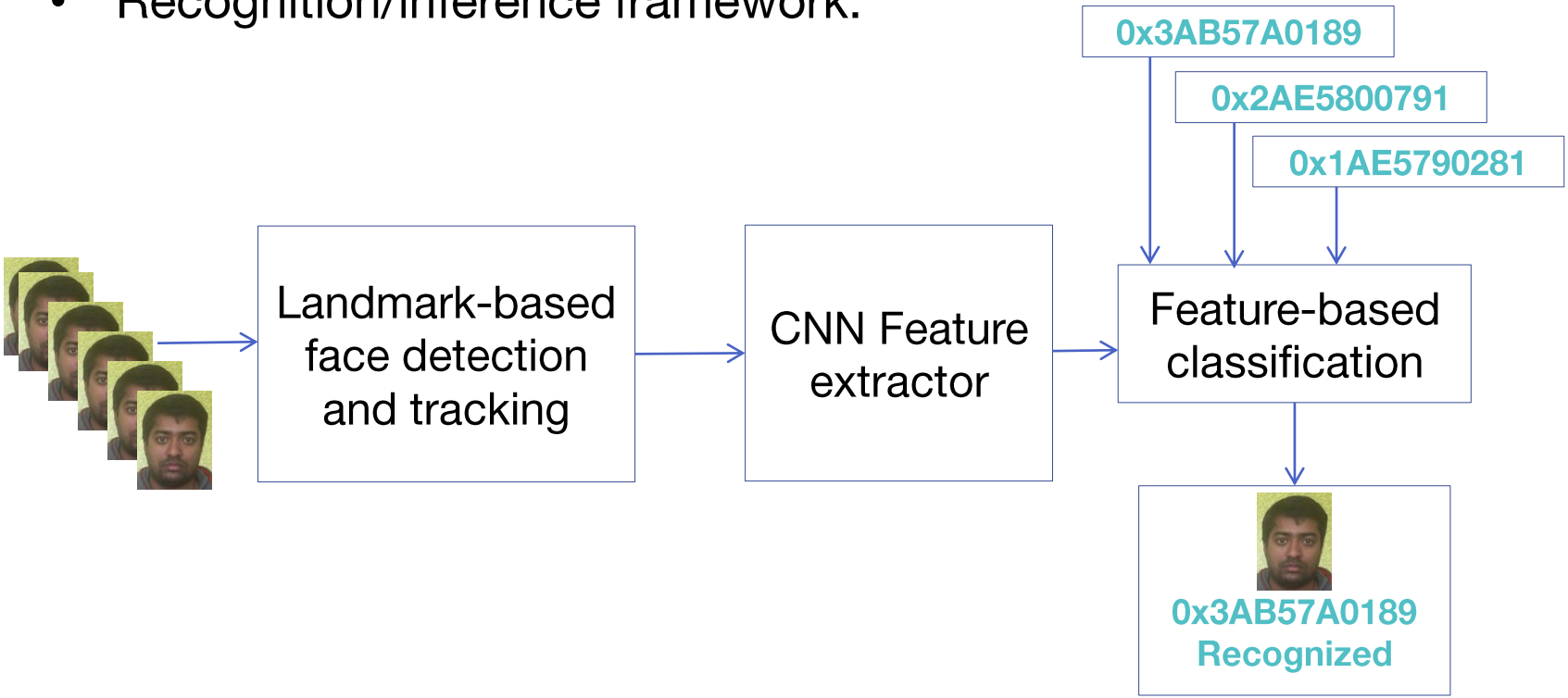
Face Recognition using CNN - Framework

- Registration framework:



Face Recognition using CNN – Framework

- Recognition/inference framework:



Dataset used for Evaluation: PP-FR Dataset

Parameter	Remarks
Pose Variation	Yaw: (-75,75) Pitch: (-60,60) Roll: (-30,30)
Illumination	Moving light sources, switching b/w sources and constant source
Expression	Sleep, partially closed eyes, laugh, talk, smile, yawn, neutral
Occlusion	Clear glasses
Eye Variations	Wide open, open, partially closed and completely closed
Backgrounds	Different environments
Registration	Single image with neutral expression and uniform illumination
Gender	15 Male and 7 Female
Duration of Capture	10 days (41,717 frames in total)

What I need to register a person?

- Registration Data
  - Uniform illumination, neutral expression.
  - Zero pose, varying background.

Testset used for Evaluation

	TestSet-1	TestSet-2	TestSet-3
Illumination	Moving source (left to right)	Switching source	Constant
Pose	Controlled variation	Controlled variation	Random
Expression & Occlusion	NIL	NIL	YES
Images from 13 registered subjects	8,759	7,326	7,742
Images from 9 non-registered subjects	6,611	5,385	5,894
Total	15,370	12,711	13,636

Dataset used for Evaluation

- TestSet-1

- TestSet-2

- TestSet-3

- TestSet-3

Experimental Results

Network	Recall Accuracy (in %)					
	TestSet-1		TestSet-2		TestSet-3	
	+ve	-ve	+ve	-ve	+ve	-ve
InceptionResNetv1	91.34	98.12	84.14	98.90	83.62	98.23
PP-FR_Net	86.97	98.36	79.56	98.30	78.18	98.06

		Predicted	
		Neg	Pos
Ground Truth	Neg	TN	FP
	Pos	FN	TP

- +ve Re-call accuracy = $\frac{TP}{(TP+FN)}$
- -ve Re-call accuracy = $\frac{TN}{(TN+FP)}$

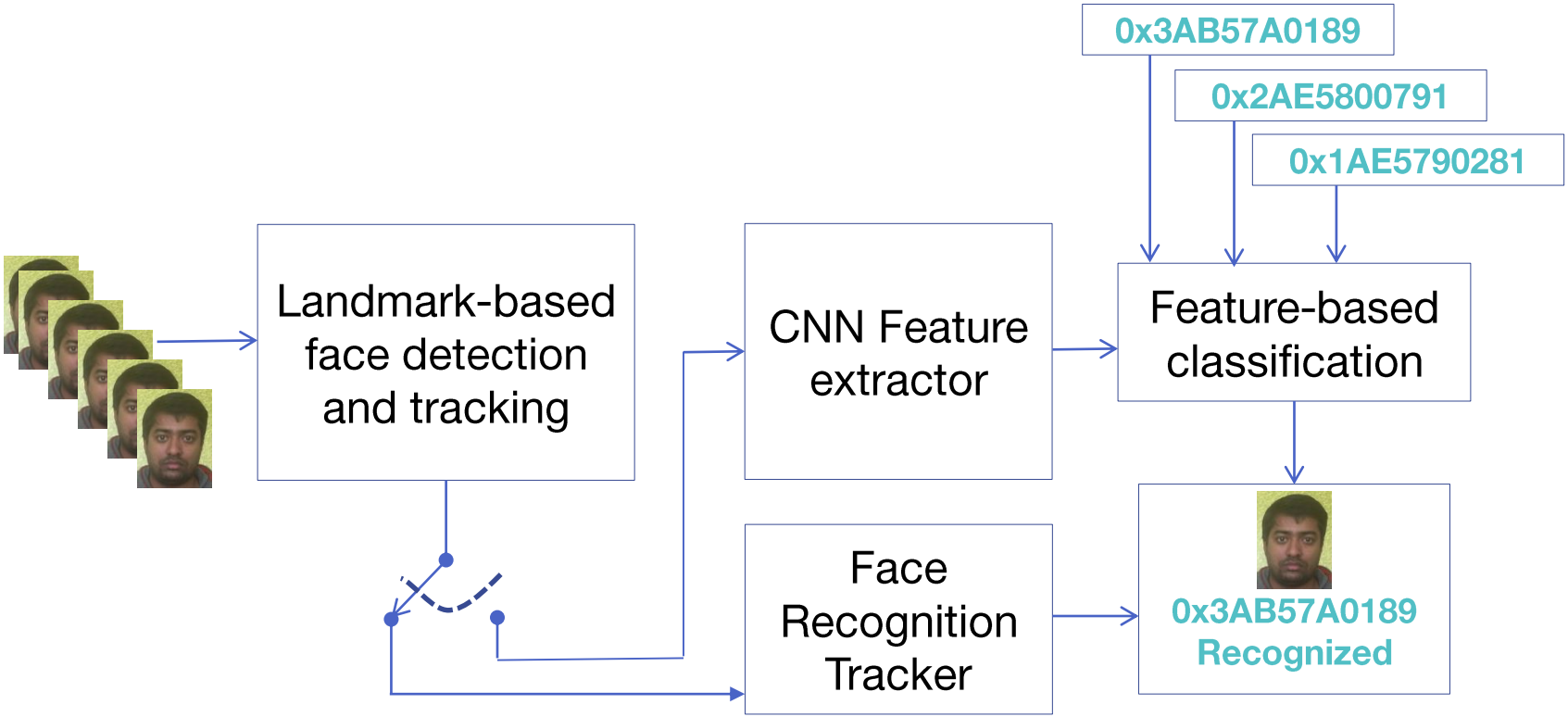
Computation time on Embedded platforms

Architecture	Computational time of PP-FR_Net
Nvidia Tegra TX2	6.6ms
Qualcomm Snapdragon 820	22ms

System Level Optimization

- Limitations of every frame face recognition system:
  - Unstable recognition results
  - Computationally complex
- Proposed approaches:
  - Face Recognition Tracker
  - Prediction of results based on output of “N” frames

Face Recognition Tracker - Framework



Face Recognition Tracker

- Input to Face Recognition Tracker: Landmark tracker parameters
- Tracker resets recognition result on following conditions:
  - Face is not detected in previous input frame
  - Change in face position compared to previous frame
  - Change of face scale compared to previous frame larger than permissible limits

Experimental Results

Network	Recall Accuracy (in %)					
	TestSet-1		TestSet-2		TestSet-3	
	+ve	-ve	+ve	-ve	+ve	-ve
InceptionResNetv1	91.34	98.12	84.14	98.90	83.62	98.23
PP-FR_Net	86.97	98.36	79.56	98.30	78.18	98.06
PP-FR_Net (tracker enabled)	88.83	98.28	81.35	98.17	80.07	98.27

Experimental Results (contd)

Architecture	Computational time (for 30 frames)	
	PP-FR_Net	PP-FR_Net (with tracker)
Nvidia Tegra TX2	198ms	66-99ms
Qualcomm Snapdragon 820	660ms	220-330ms

Summary

- With right combination of training data and cost function, low-complex CNN models can be used for Face Recognition without much impact on accuracy
- Use Face Recognition tracker to achieve stable results & to reduce computational complexity

References

1) Aishwarya Jadhav et al., “Deep Attributes for One-Shot Face Recognition”
2) Omkar M. Parkhi et al., “Deep Face Recognition”
3) Schroff et al., “FaceNet: A Unified Embedding for Face Recognition and Clustering”
4) Taigman et al., “DeepFace: Closing the Gap to Human-level Performance in Face Verification”
5) Zhu et al., “Deep Learning Identity Preserving Face Space”
6) Yandong et al., “A Discriminative Feature Learning Approach for Deep Face Recognition”
7) DavidSandberg, “github.com/davidsandberg/facenet”

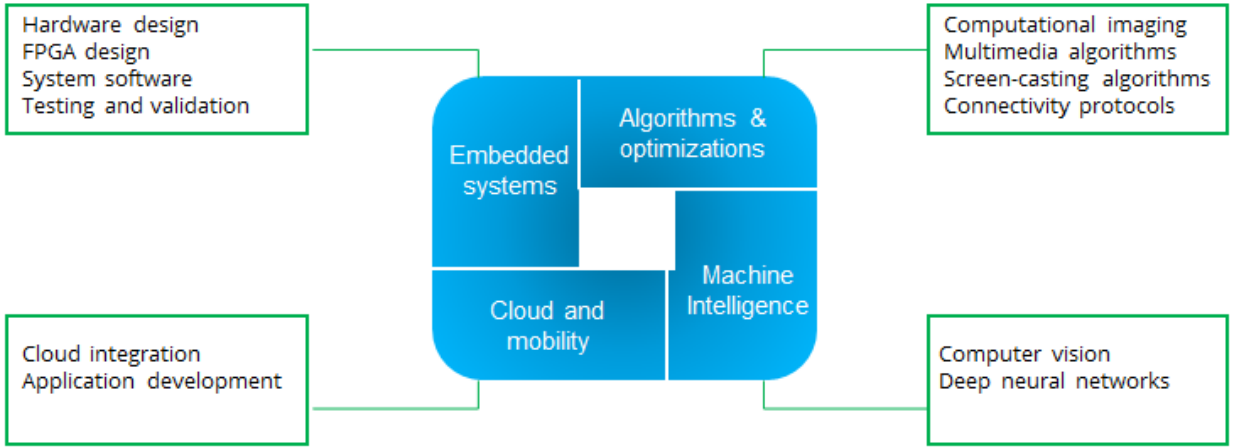
About PathPartner

PathPartner is a top notch design engineering services company and works with several semiconductor companies, OEMs and ODMs in embedded media-centric devices and intelligent systems.

- Incorporated in July 2006; HQ in Bangalore, India
- R&D Centers: Bangalore, India and California, USA
- Marketing representatives in USA, Europe, Japan and India
- PathPartner is a member of Embedded Vision Alliance and partner of various semiconductor companies

- Present company strength is ~280
- Quality : ISO 9001:2015, 27001:2013
- R&D Workforce : >10%

- Semiconductor Companies
- OEMs and ODMs



+91 80 6772 2000 |
+1 408 242 7411 |
+81 9055466819

sales.india@pathpartnertech.com |
sales.usa@pathpartnertech.com |
sales.japan@pathpartnertech.com

Thanks!!!

- Questions?