

# CSE802 Final Project

## CNN on Unconstrained Face Recognition

Mengying Sun

Deliang Yang

04/26/2017

# Outline

- Network Structure
- Model 1
- Model 2
- Model 3
- Evaluation & Comparison
- Summary & Future improvement
- Conclusion

# Work Flow

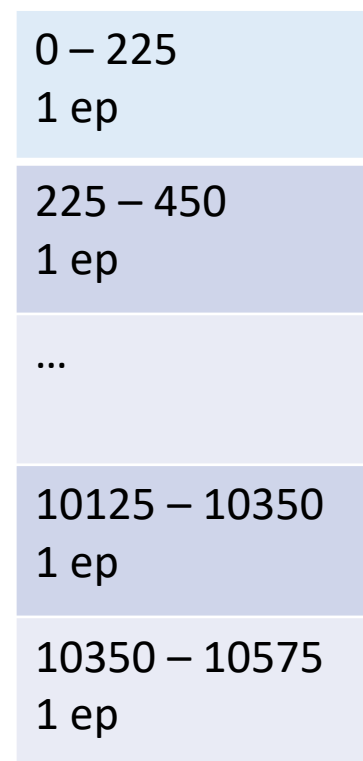
- Use 100 classes' samples as developmental evaluation
- 80% train, 20% test for each class
- Compare loss and accuracy for training set and test set
- Parameters tuning
- Apply the final model structure and hyperparameters to the large scale dataset (CASIA)
- Validate the model with LFW + BLUFR

Layer	Output dim	Parameters #
conv2d_1	110 x 110 x 32	896
conv2d_2	108 x 108 x 64	18496
activation_1 ReLU		
max_pooling2d_1	54 x 54 x 64	
conv2d_3	52 x 52 x 64	36928
conv2d_4	50 x 50x 128	73856
activation_2 ReLU		
max_pooling2d_2	25 x 25 x 128	
conv2d_5	23 x 23 x 128	147584
conv2d_6	21 x 21 x 256	295168
activation_3 ReLU		
max_pooling2d_3	10 x 10 x 256	
conv2d_7	8 x 8 x 256	590080
conv2d_8	6 x 6 x 384	885120
activation_4 ReLU		
average_pooling2d_1	3 x 3 x 384	
flatten_1	3456	
dropout_1	3456	
dense_1	10575	36557775
activation_5 Softmax		

# Model 1

- Loop 20 times:
  - Loop 47 times:
    - Fetch data from 225 new classes
    - Fit the CNN model for 1 epoch
- (Equivalent to 20 epochs, but...)
- Low efficiency (I/O) and overfitted, 95% train accuracy, but
- Bad results:

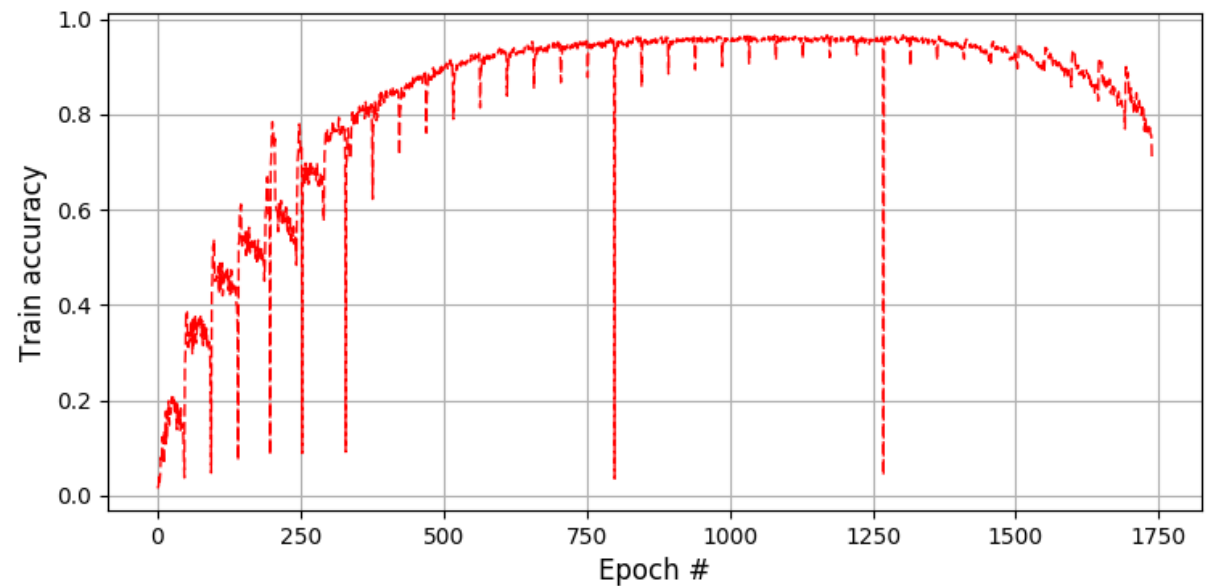
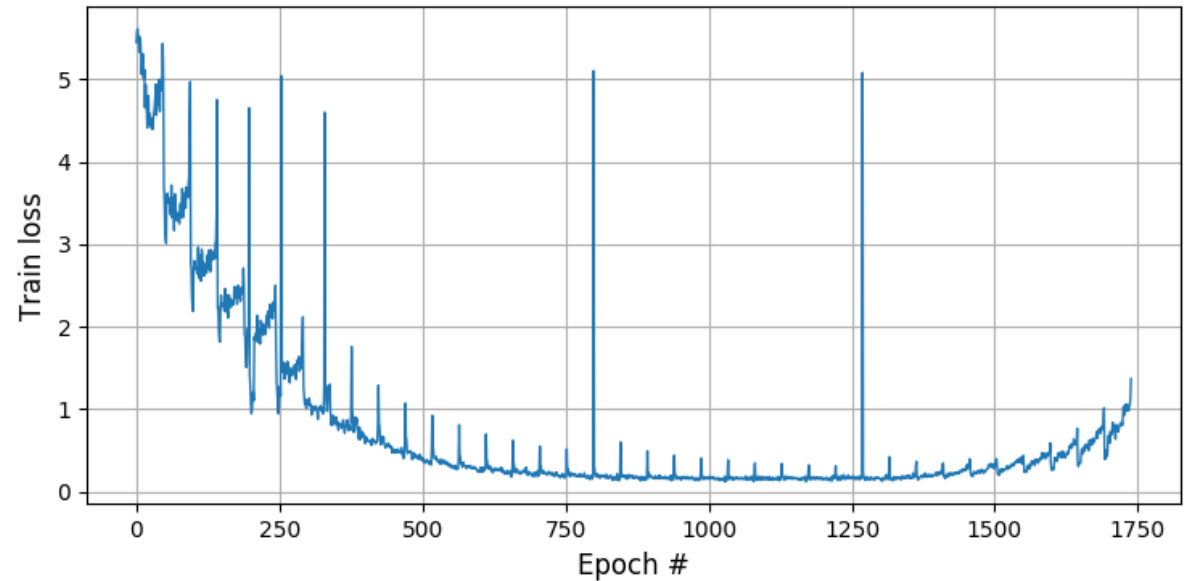
Loop 20  
times



	PCA @ 500 dim	LDA
Verification @ FAR = 0.1%:	<b>0.22%</b>	<b>0.10%</b>
Open-set Identification @ Rank = 1	<b>0.00%</b>	<b>-0.01%</b>

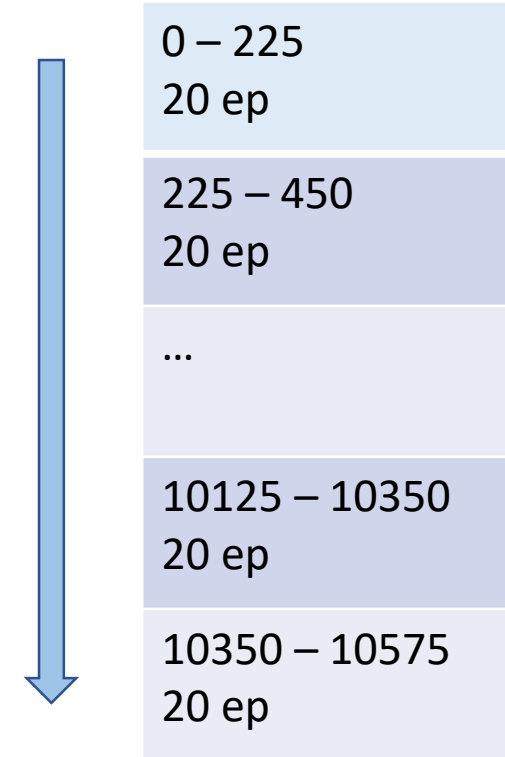
# Model 1

- 225 classes per epoch
- Overfitted
  - High training accuracy
  - Low VR and DIR on LFW



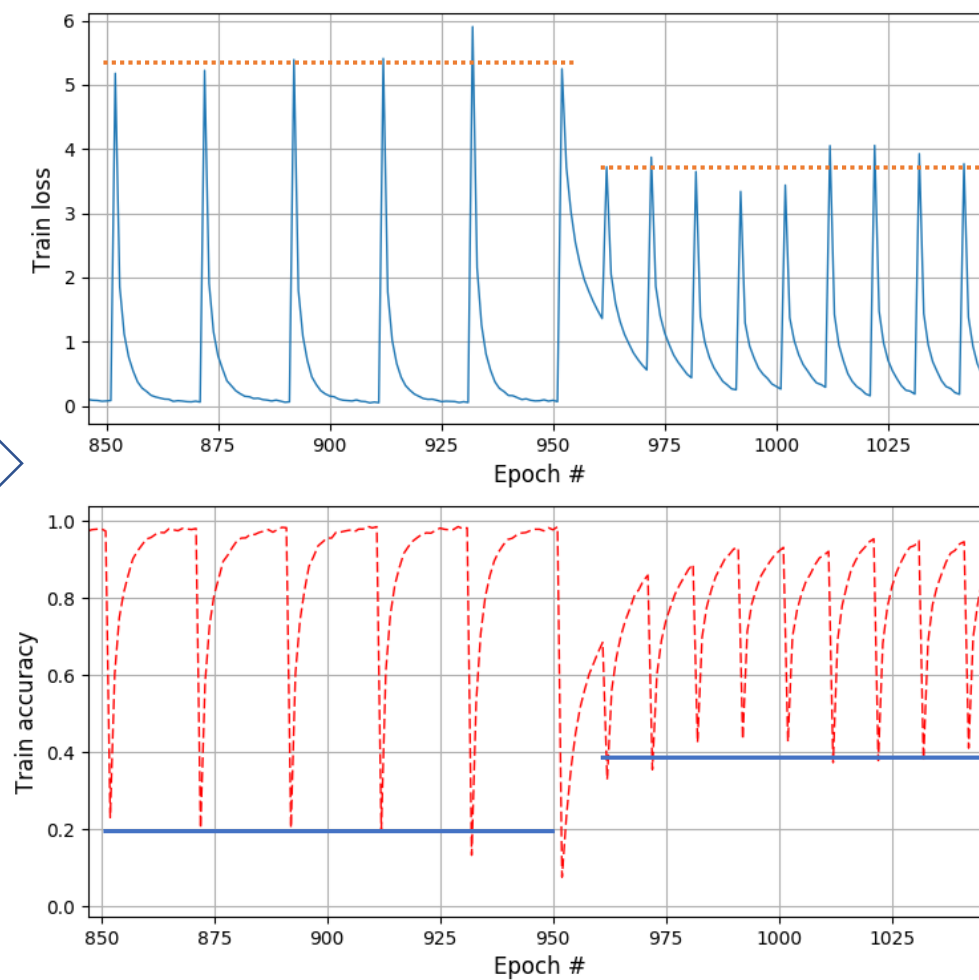
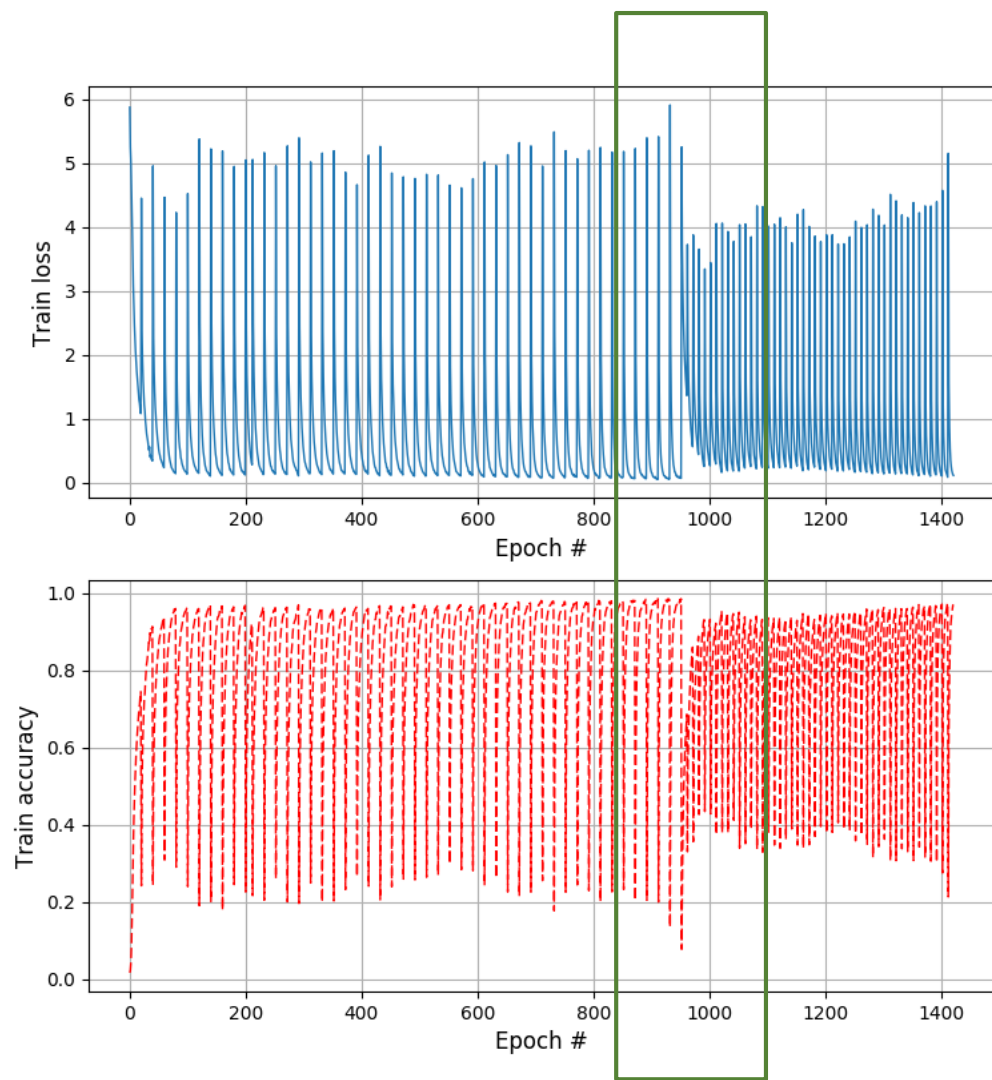
# Model 2

- Loop 47 times:
  - Fetch data from 225 new classes
  - Fit the CNN model for 20 epochs
- (Hopefully equivalent to 20 epochs, but...)
- Wrong again.
- Results are better:



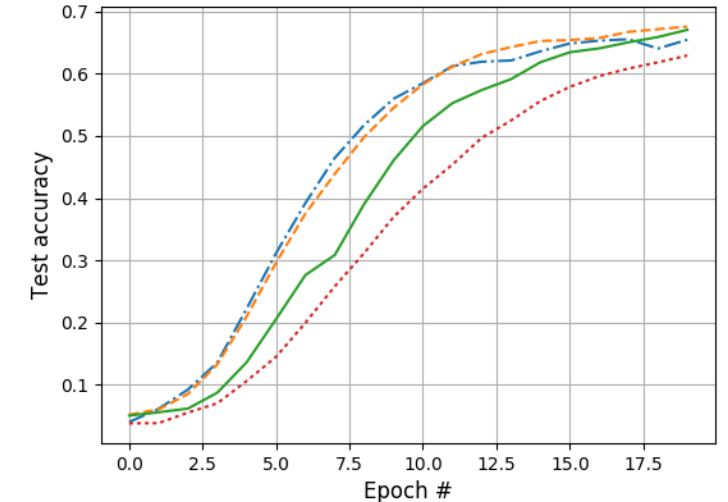
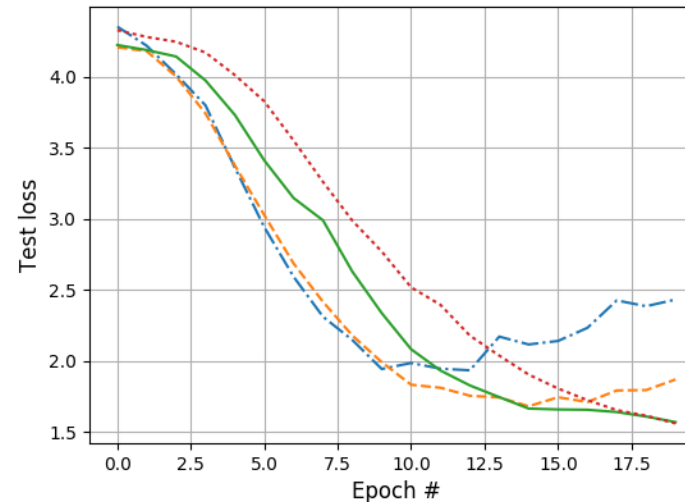
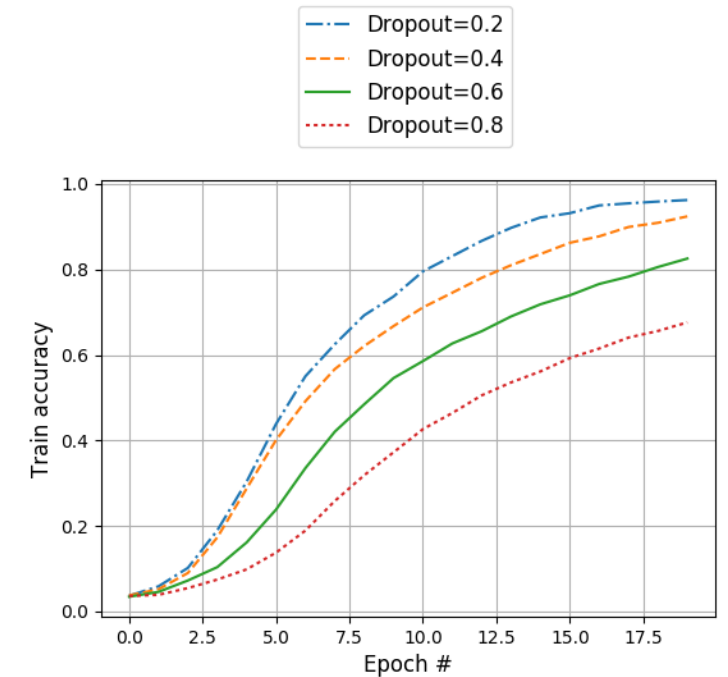
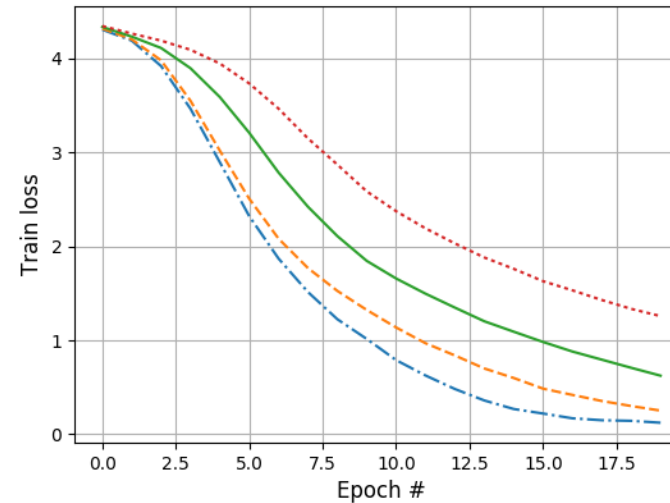
	PCA @ 500 dim	LDA
Verification @ FAR = 0.1%:	<b>9.67%</b>	<b>8.02%</b>
Open-set Identification @ Rank = 1	<b>1.54%</b>	<b>2.52%</b>

# Model 2



# Dropout Rate Exploration

- In 100-class dev mode
- Higher dropout rate:
  - Higher train loss
  - Lower train accuracy
  - Lower test loss
  - Higher test accuracy
- Tradeoff
- Finally, dropout rate is chosen to be 0.60





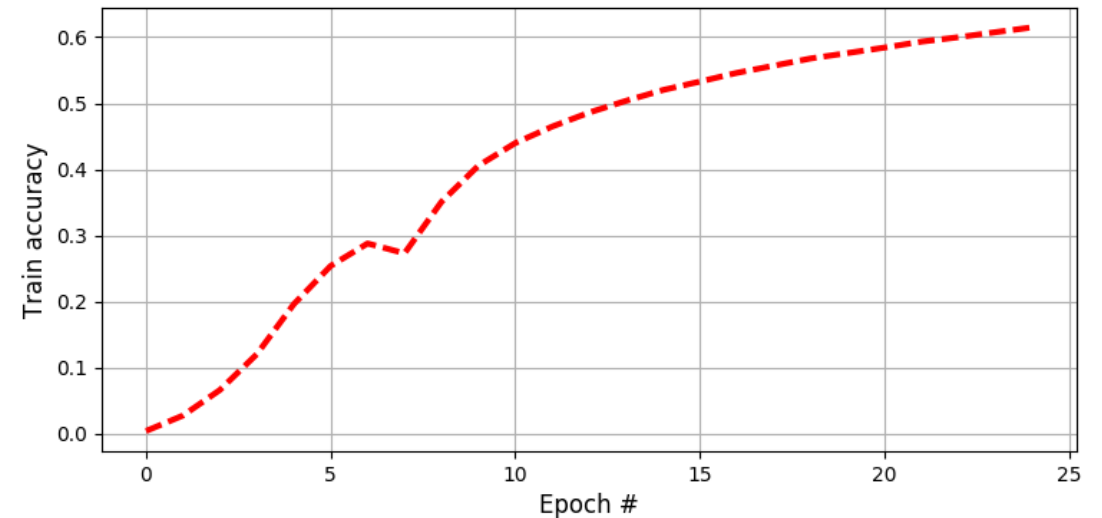
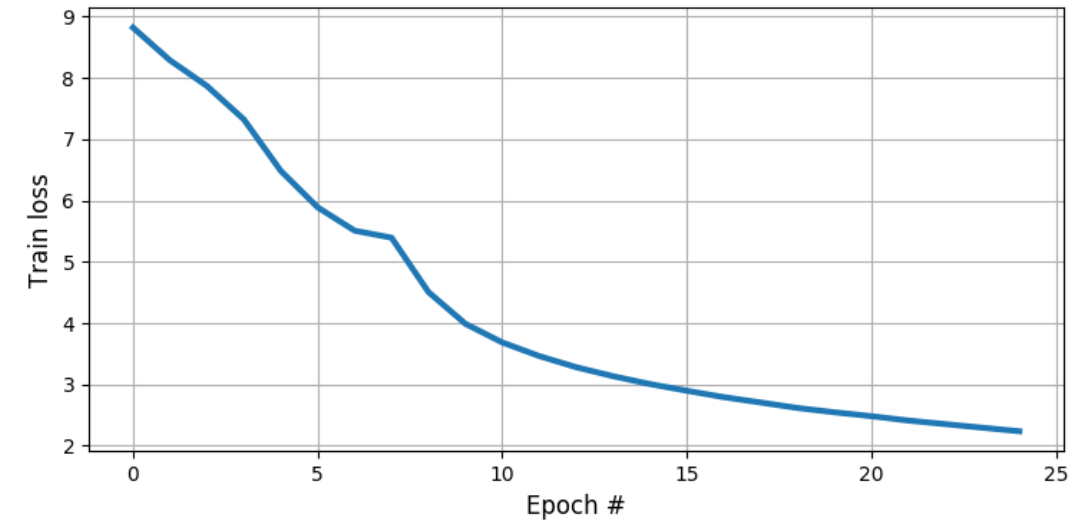
# Model 3 (Latest)

- Use `ImageDataGenerator` in Keras
- Function `flow_from_directory()`
- Rescale, shuffle, transformation supported
- Samples in CASIA are chosen randomly to fit the model
- Longer time for each epoch

# Model 3

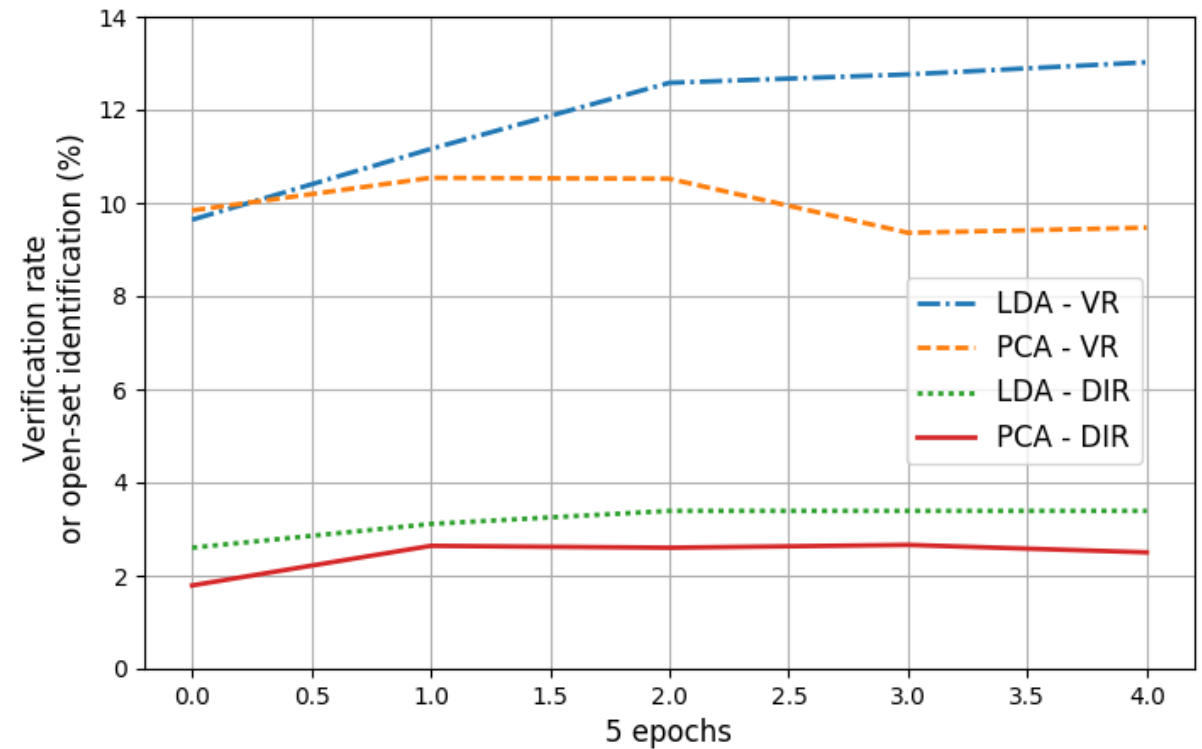
- Converge slowly
- Training accuracy is not high
- Results:

	PCA @ 500 dim	LDA
Verification @ FAR = 0.1%:	<b>10.54%</b>	<b>13.02%</b>
Open-set Identification @ Rank = 1	<b>2.66%</b>	<b>3.39%</b>



# Model 3

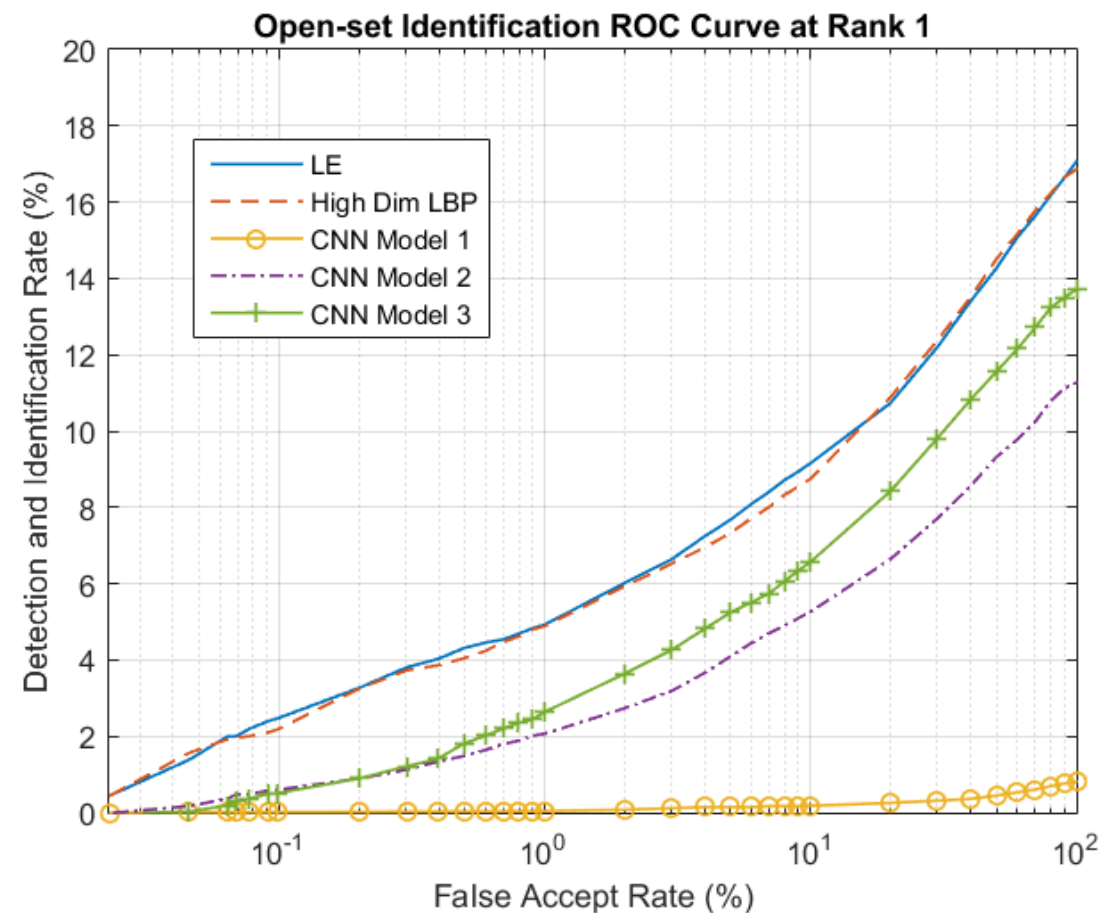
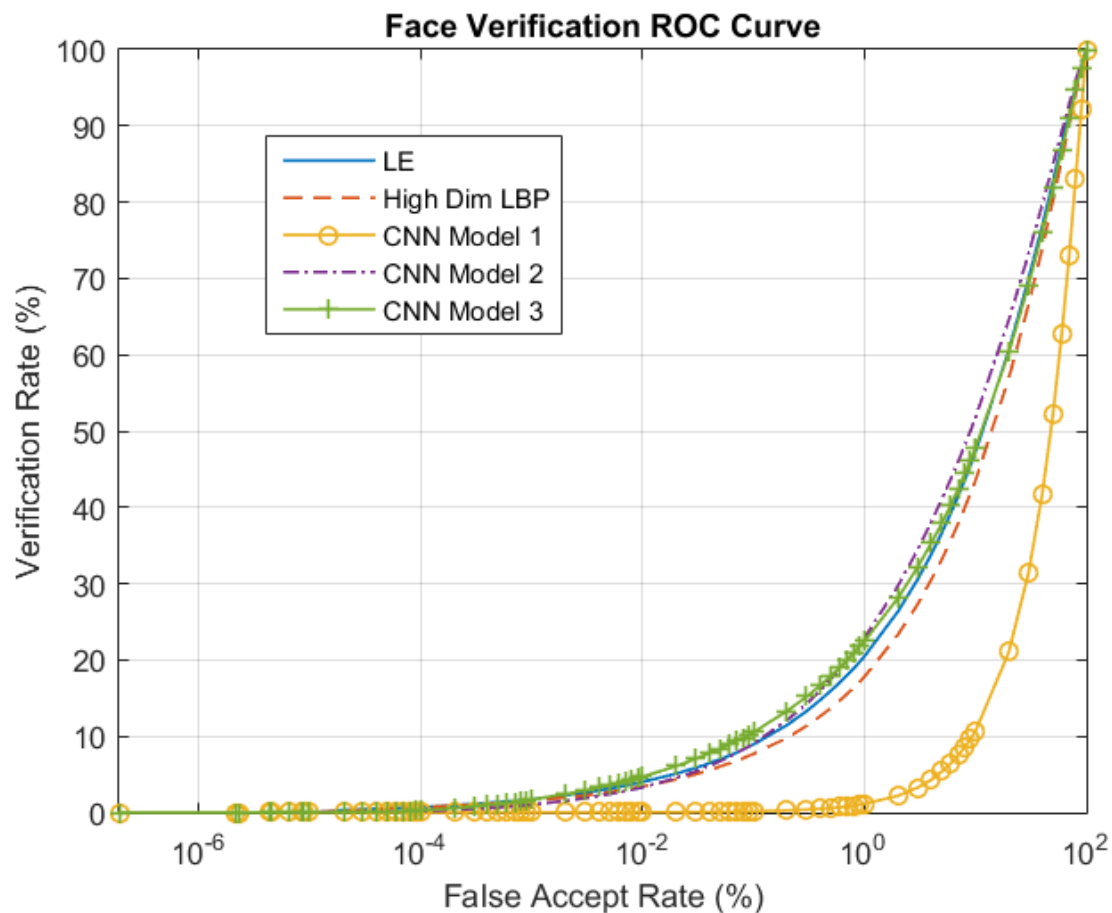
- 25 epochs trained
- Performance did not improve too much over time
- LDA > PCA
- Testing result may not improve if we train more epochs



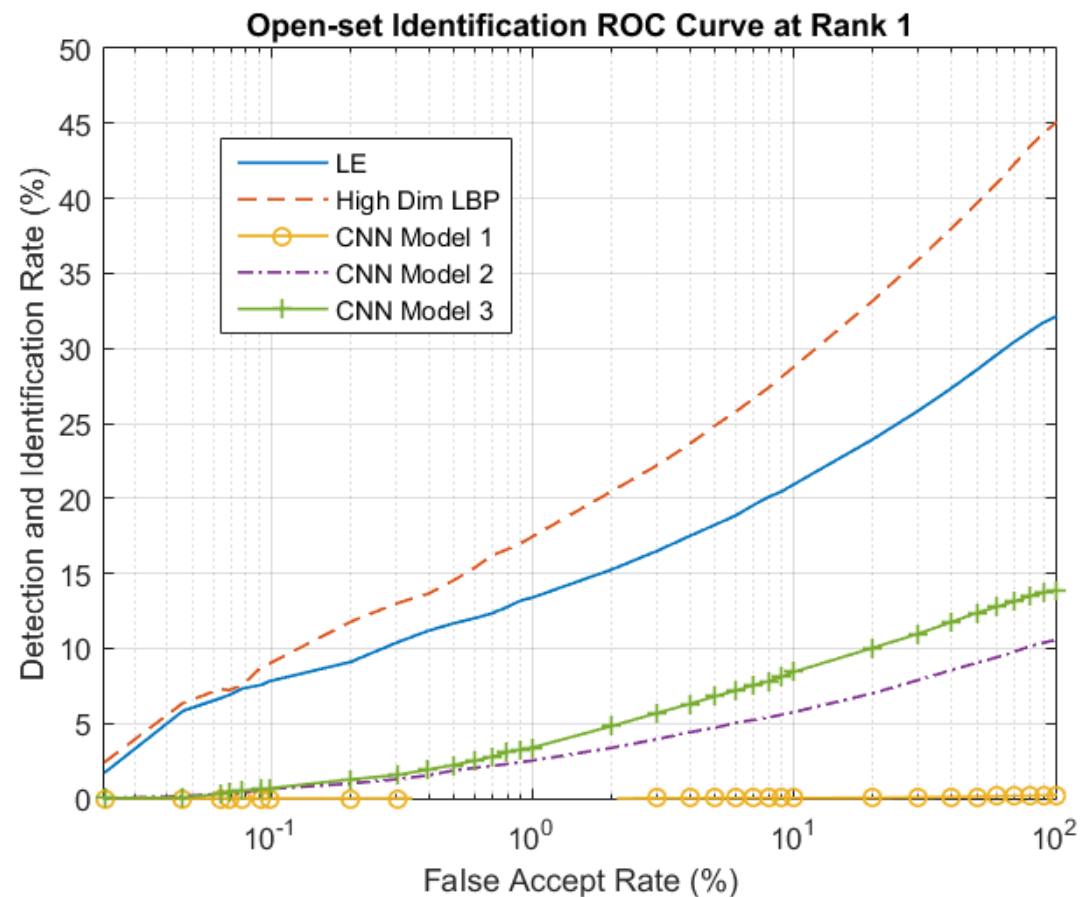
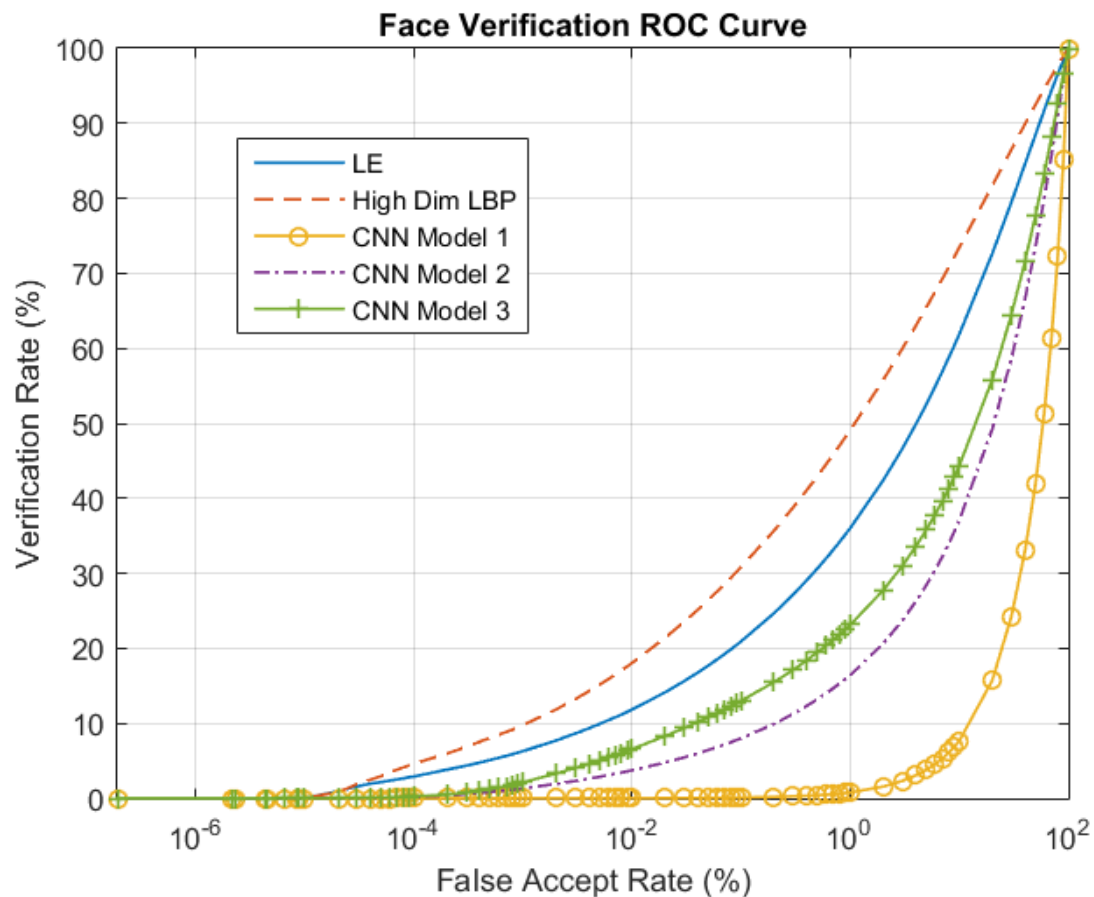
# Comparison to Other Method

	LE (Learning-based descriptors)		High Dimensional LBP		CNN Model 3	
Feature size	20736		127440		3456	
Method	PCA	LDA	PCA	LDA	PCA	LDA
Verification @ FAR = 0.1%:	8.94%	20.95%	8.20%	30.86%	10.54%	13.02%
Open-set Identification @ Rank = 1	4.93%	13.39%	5.32%	17.46%	2.66%	3.39%

# PCA Feature Extraction – VR & DIR ROC



# LDA Feature Extraction – VR & DIR ROC



# Improvement

Try different:

- Activation function
- Optimizer (learning rate, decay, momentum etc.)
- Layer
- Batch size
- Batch normalization
- Residual Network

# Conclusion

- Training method is important
- LDA outperforms PCA most of the time in BLUFR
- High dimensional LBP has the best performance at present
- Many things can be done to improve VR/DIR



Thanks!