FORENSICS @ NIST #NISTForensics

Deep Learning based Feature Extractors for Shoe Print Matching

By
Sarala Padi

National Institute of Standards and Technology
U.S. Department of Commerce

Outline

- Introduction
  - Shoe print matching
  - Deep Learning models
- Problem statement
  - Deep learning for shoe print matching?
    - Lack of data for training?
    - Transfer learning
- Approach used for shoe print matching
- Database used
  - Dataset from West Virginia University (WVU)
  - FBI Boots data
- Experimental Evaluations
- Conclusions
- References

INTRODUCTION (1):

Shoe print matching in Forensic Science crime scene analysis usually involve two categories of shoe prints

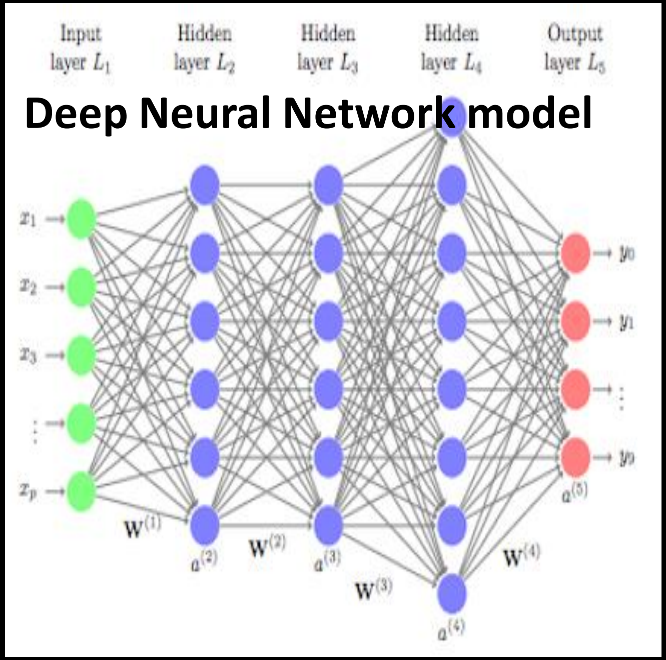
- **Crime Scene Impressions:** Footwear impression taken from a crime scene
- **Reference Impression:** Footwear impression taken from a shoe of interest
- **Current Approaches:**
  - **Investigative:** Automatically finding the make and model for the given crime scene impression from the library of impressions
  - **Evidential:** Evaluate the level of correspondence between crime scene and reference impressions by comparing size, outsole design, wear patterns, and randomly acquired characteristics (RACs)
- **Concerns:**
  - Manually done and need professional experts
  - Subjective measure and it can be easily biased

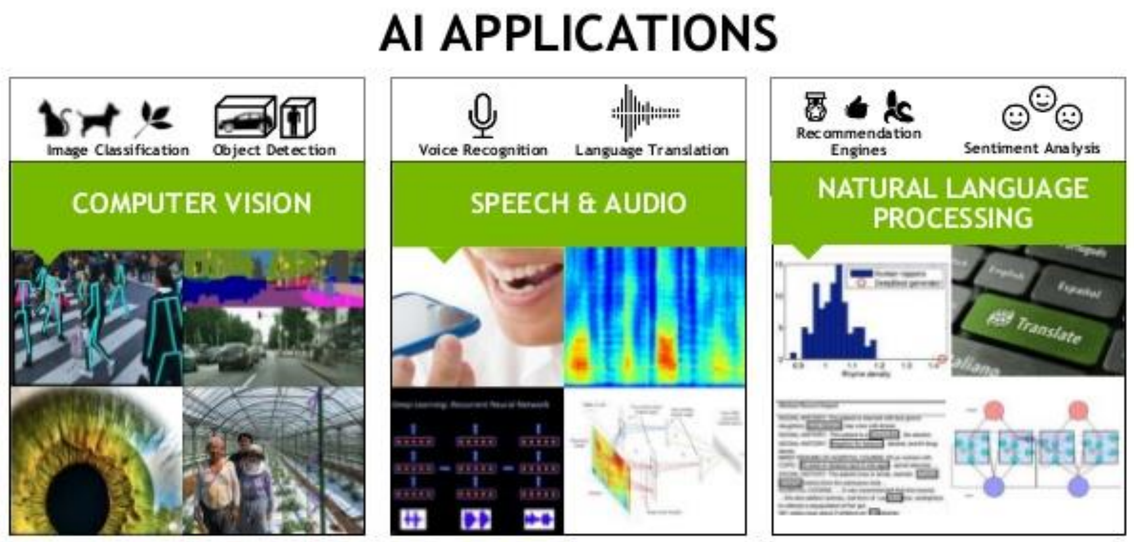
INTRODUCTION (2):

- **Objective Measures:** Automatically finding the correct match for the given crime scene impression by directly getting the features from the impressions
- **Features:** Finding and extracting the right kind of features to compute the similarity between pair of images is an important and crucial step
- **Current Approaches:** Most of the approaches quantify the degree of correspondence
  - by computing a similarity score on the original impressions
  - or suitable transforms such as Fourier, Gabor, Mellin, etc.
- **Recent Work:** Kong et.al have shown that
  - Resnet model features can lead to good performing similarity measures
  - Multi Channel Normalized Cross Correlation (MCNCC) metric is used for finding the similarity between the pair of impressions

INTRODUCTION (3):

- ## Deep Learning:
  - Deep neural network models are shown to be successful in extracting features that are more informative for comparison purposes
  - State of the art include many frameworks and pretrained models that are easily adapted to domain specific applications
- ## Requirements:
  - Building such models require large amount of data for training
  - Computationally expensive to build such models



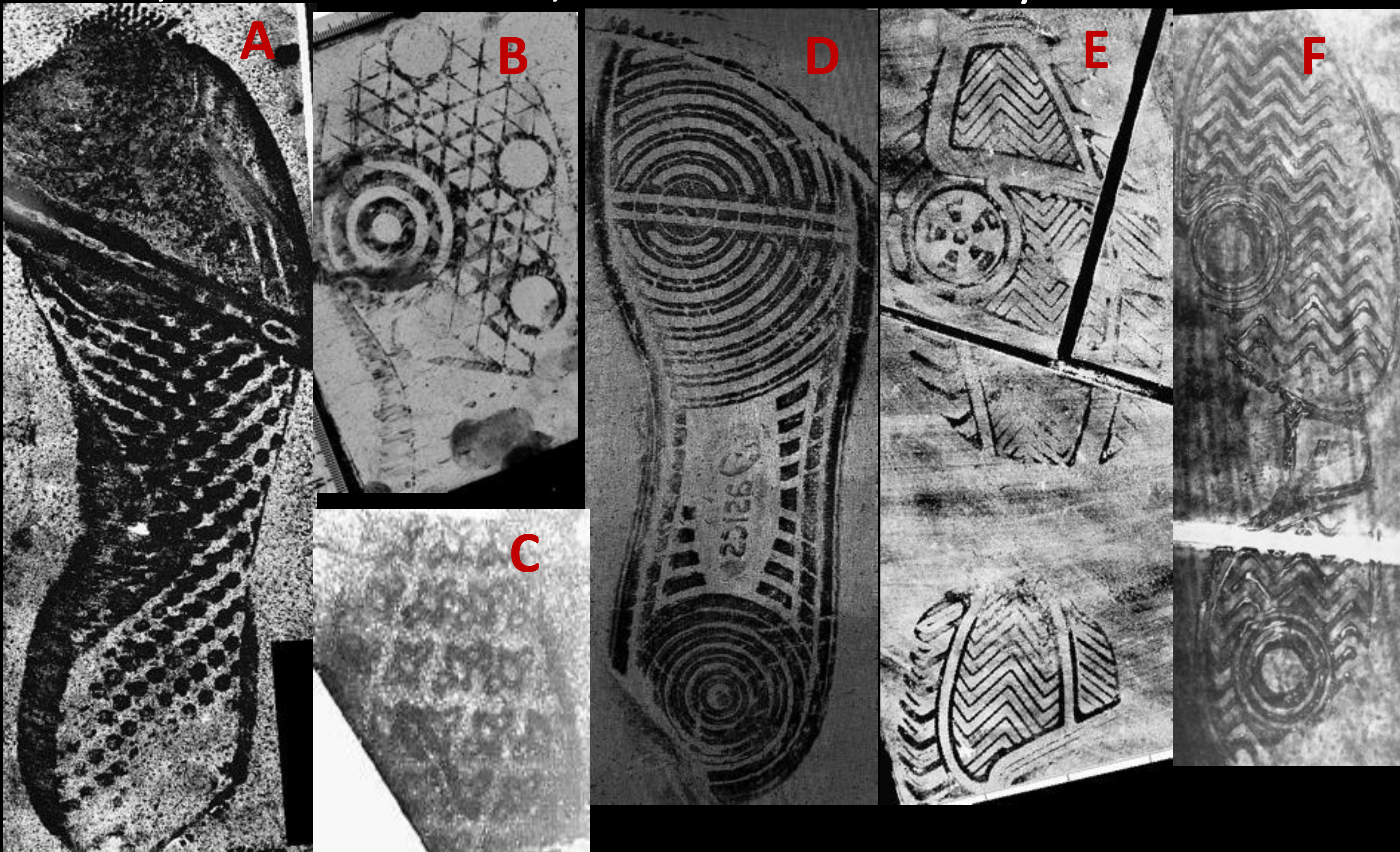


Source: AI CONNECT CONFERENCE 2017 (from NVIDIA)

PROBLEM STATEMENT:

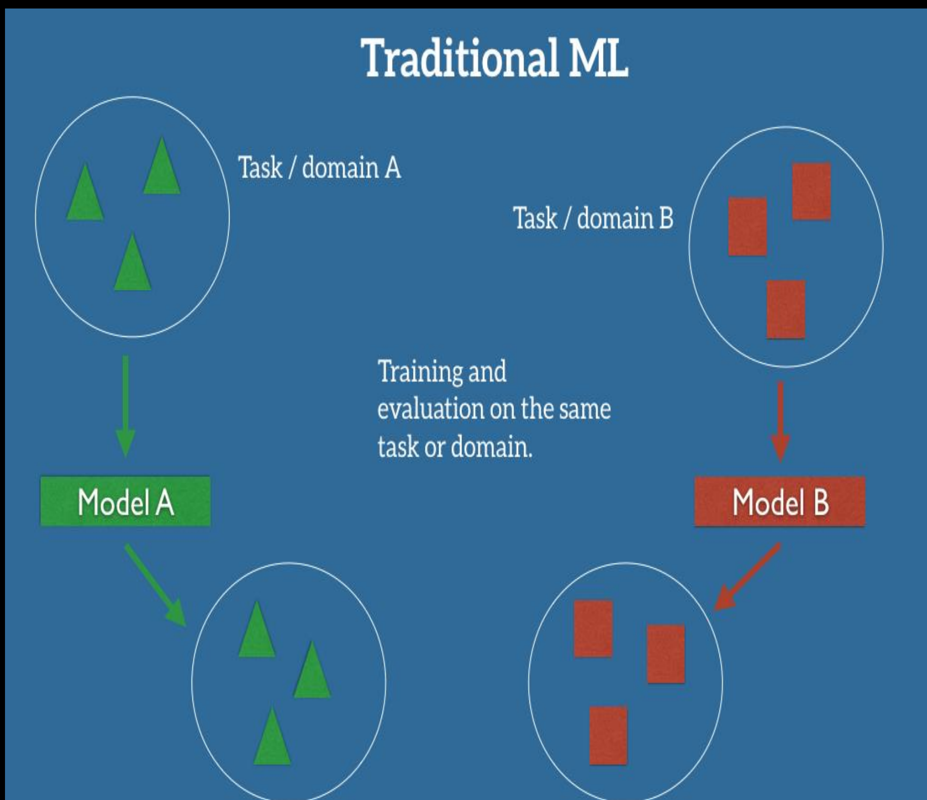
- **Challenges in applying DL models to shoe print matching:**
  - Unavailability of Datasets for modeling
  - Available datasets are small, low in quality, partial and varied in size, resolution, scale, modality, etc.
- **Proposed Method:**
  - **Pretrained models** with transfer learning is used for shoe print matching
  - **Resnet-50 model** is used to extract features with Multi Channel Phase Only Correlation (MCPOC) similarity metric to find the degree of similarity between the crime scene and reference impression
  - **Resnet model features** are trained with Least Absolute Shrinkage and Selection Operator (LASSO) regressor to get weighted average scores for finding the similarity between the pair of impressions

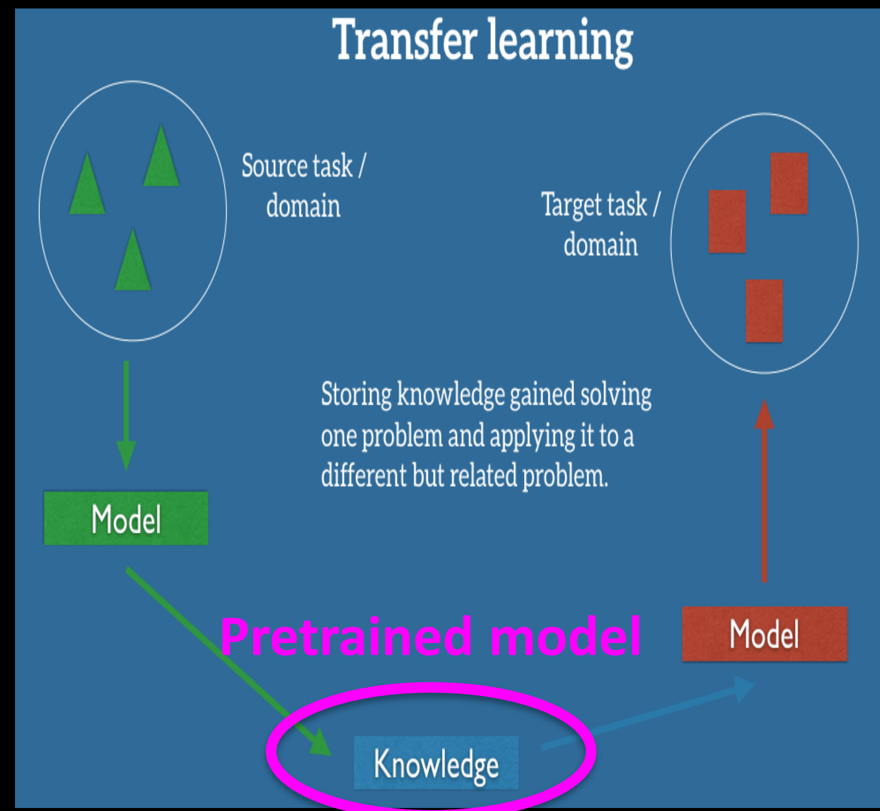
Crime scene impressions:
Partial, different in size, scale and modality



Transfer Learning (1):

Transfer learning and domain adaptation refer to the situation where what has been learned in one setting is exploited to improve generalization in another setting.



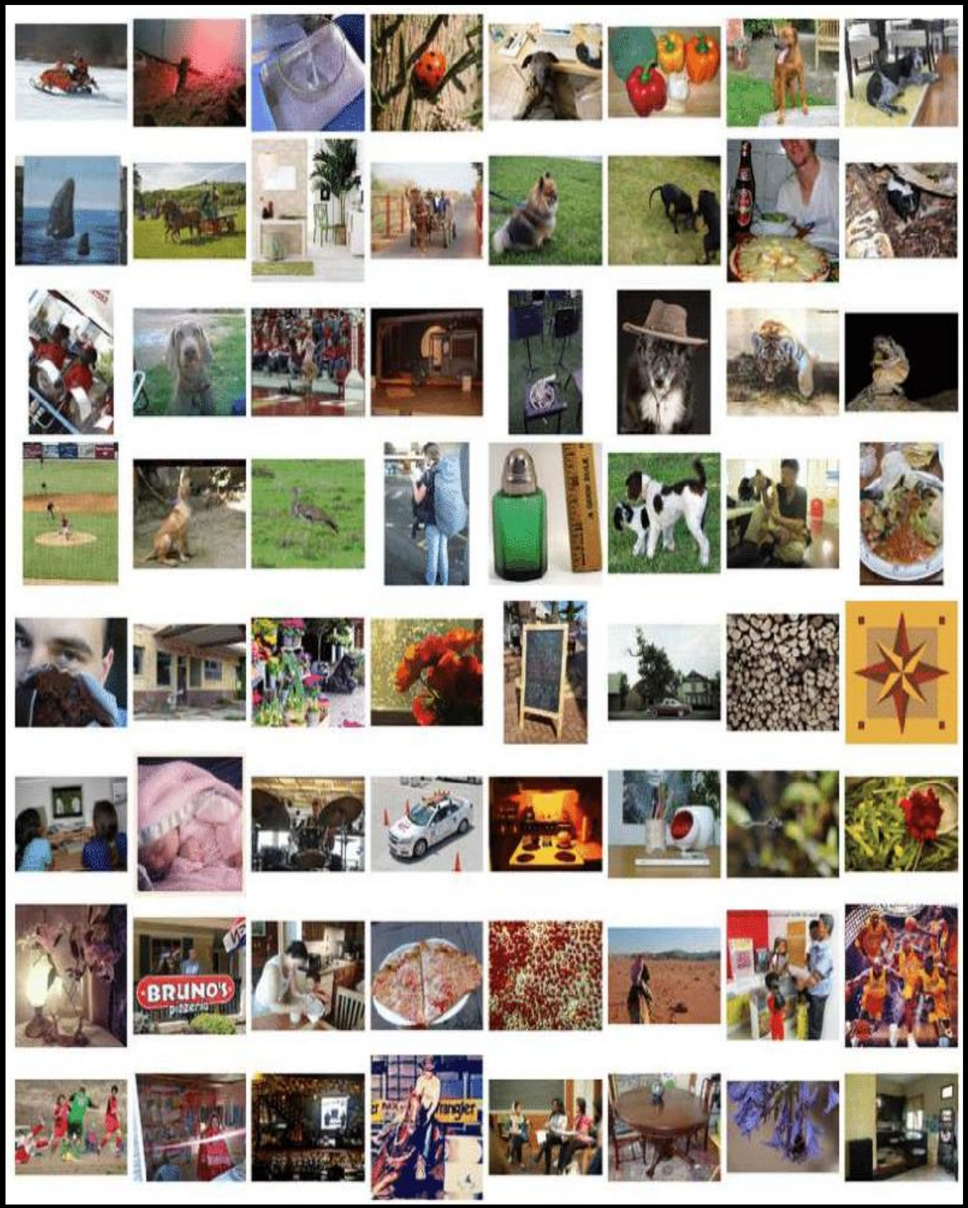


Transfer Learning (2):

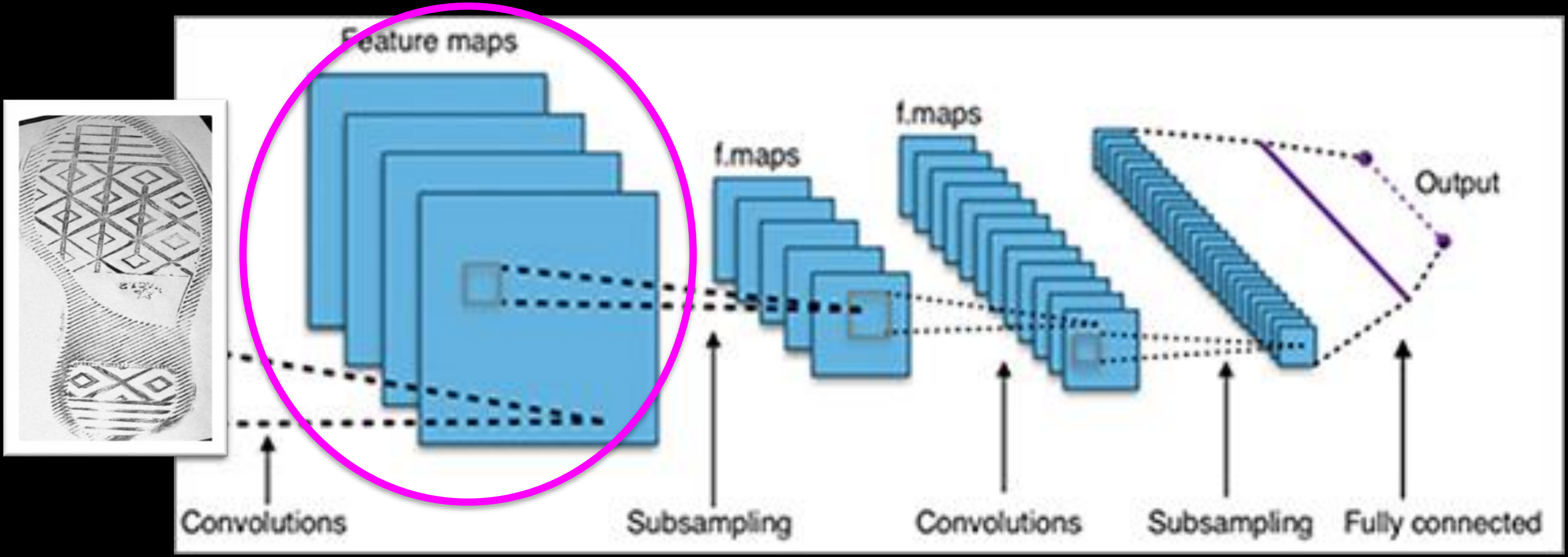
Pretrained model:

- A model is created to solve a problem
- When we try to solve a similar problem
  - Use the trained model as a starting point
- **ImageNet:**
  - Contains 1.2 million images
  - with 1000 categories
  - Animals, birds, trees, sports, vegetables, people, etc.
- Pretrained models built on ImageNet dataset that are available for use
  - **Lenet-5, VGG16**
  - **AlexNet, Resnet50**
  - **Inception, GoogleNet**

Sample Images from ImageNet Dataset

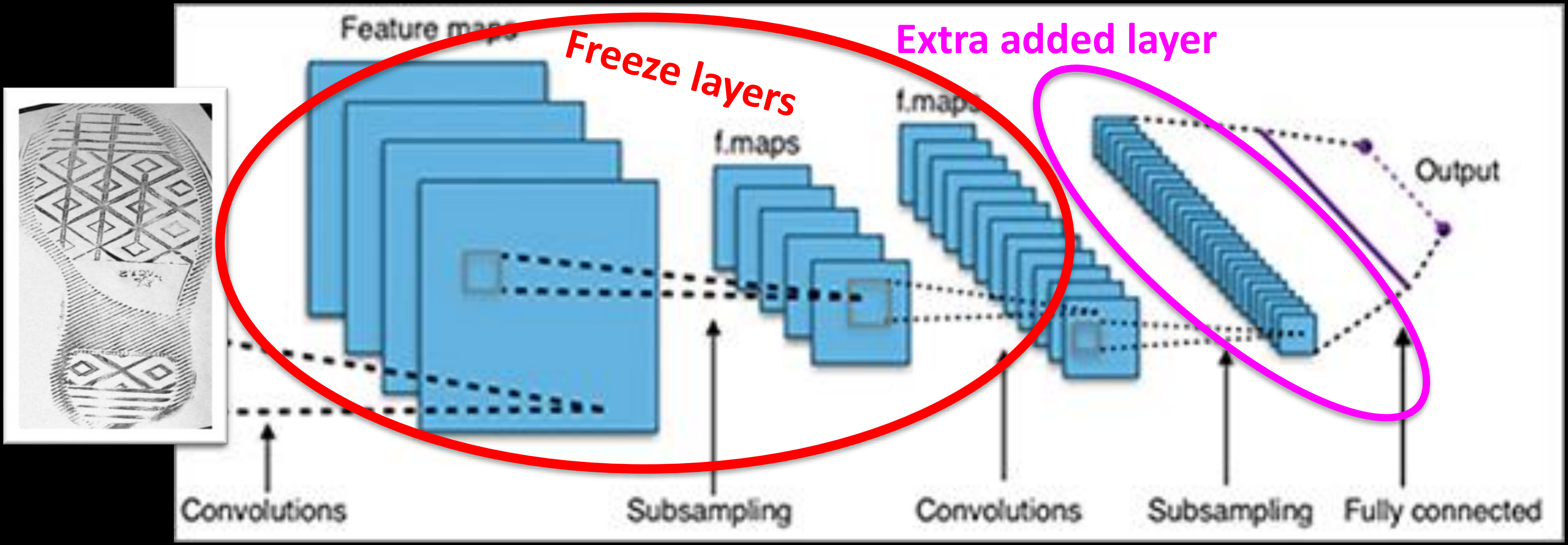


Transfer Learning: With Fixed Feature Maps



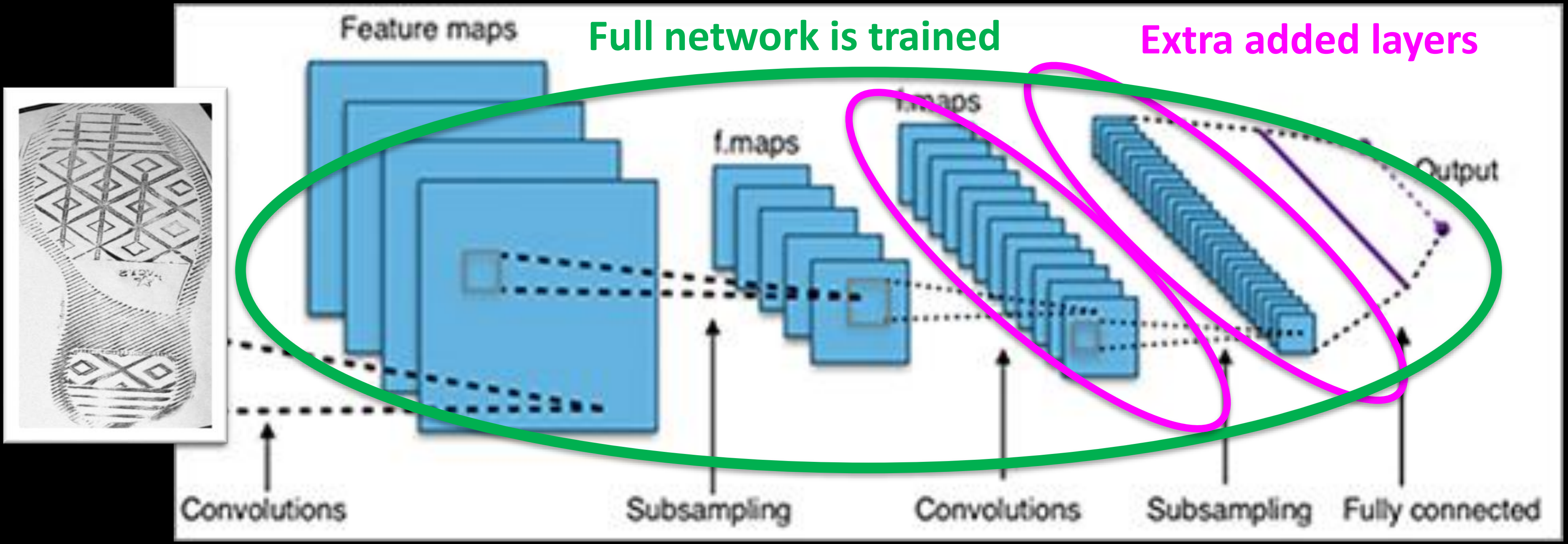
- Pretrained model with Fixed feature vectors
- Training is not required
- Initial layers can be used as feature extractors

Transfer Learning: one or two Extra added layers



- Pretrained model with one or two extra layers
- Training only the added layers and freeze the other layers
- Require small amount of data for training
- Model can be used for solving similar problem

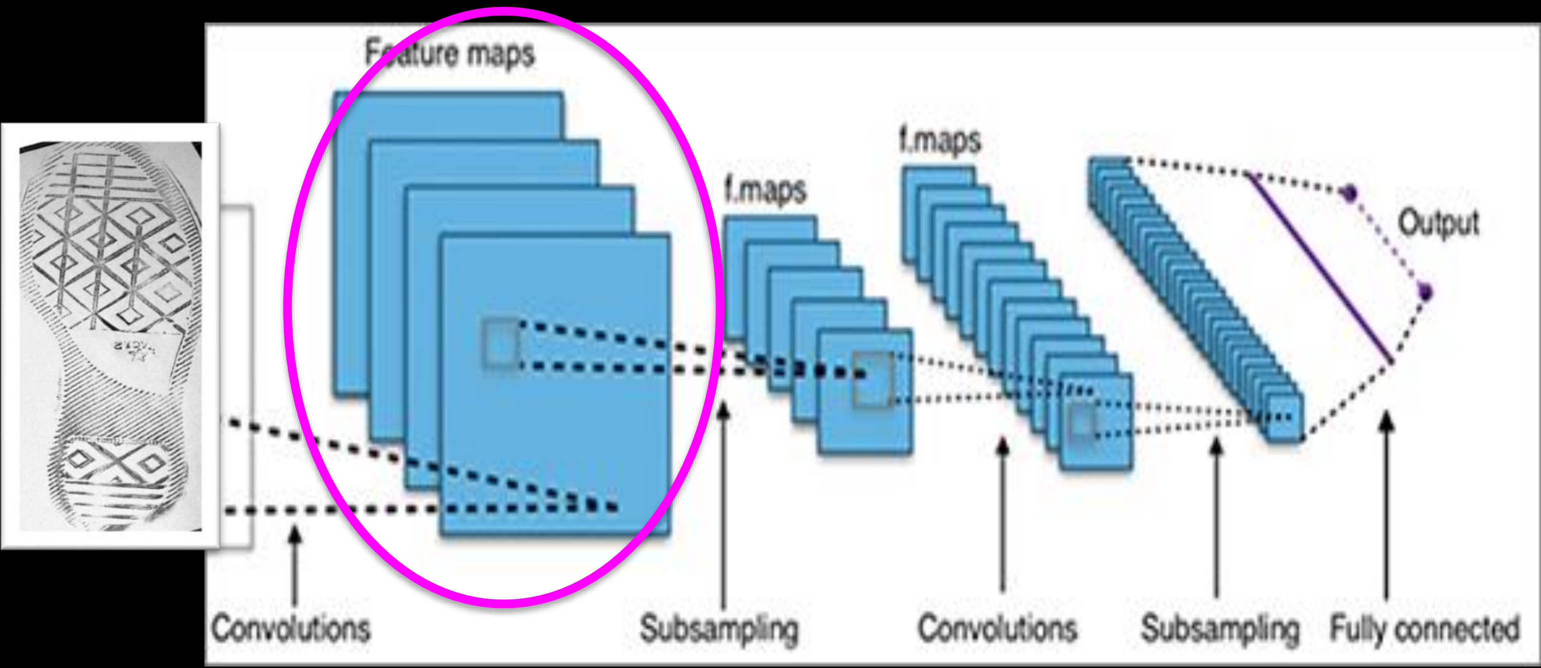
Transfer Learning: Extra added layers

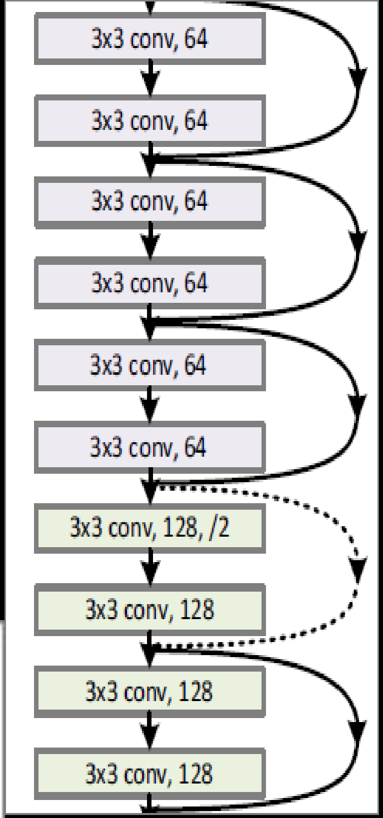


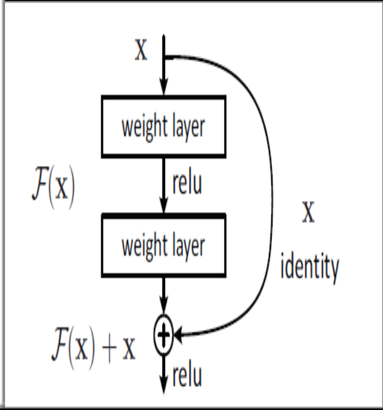
- Pretrained model with extra added layers
- Training the full network
- Require a large amount of data for training
- Model can be used for task of interest

Resnet-50 Model:

- **Model Architecture:** Convolutional neural network model
- **Number of blocks:** 24 blocks with two convolutions in each block
- **Residual :** Input is feed forwarded to each block (24 blocks)
- **Number of layers:** 50 Layers
- **Layer considered to extract features: Initial layer**
- **Initial layers:** extract edge like features and these features can be generalizable to new datasets



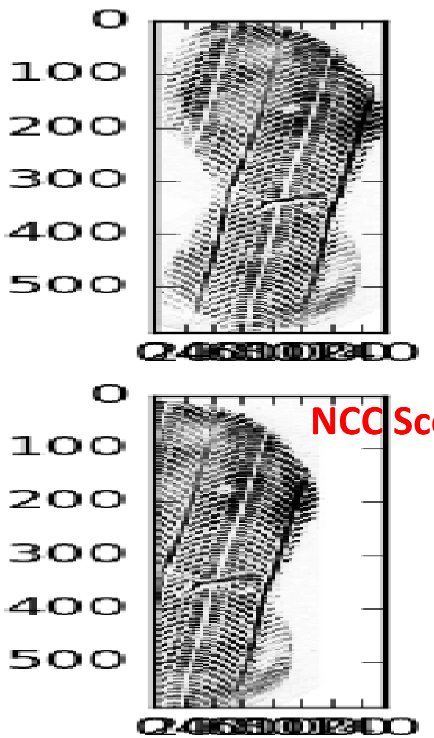




Similarity Metrics:

Normalized Cross Correlation

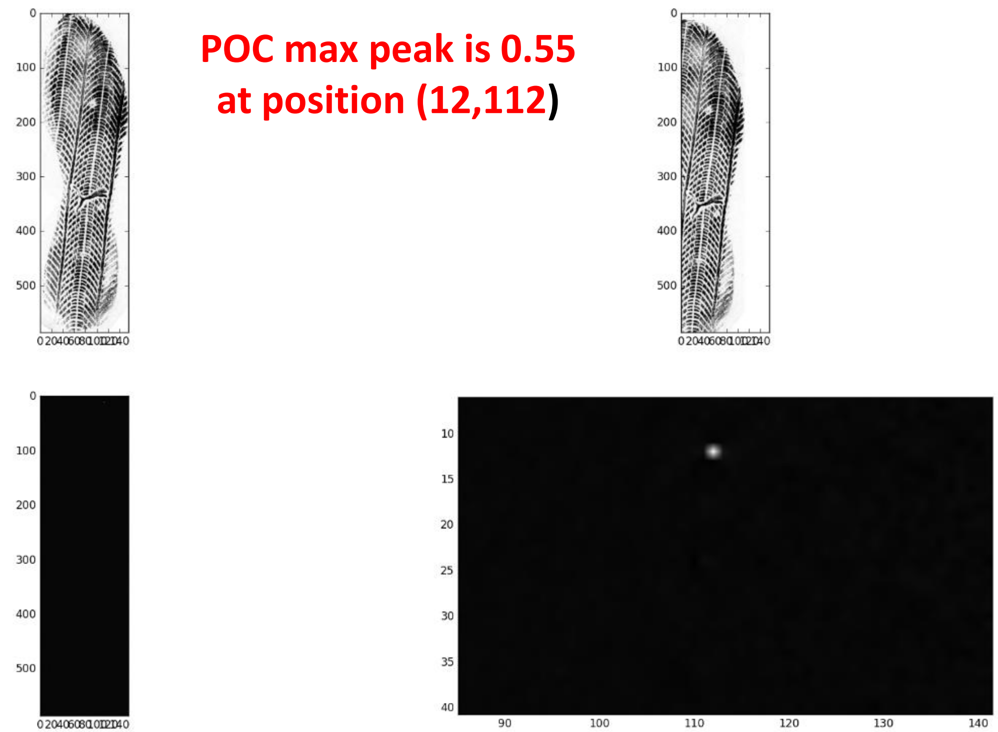
$$MCNCC=\frac{1}{N}\sum_{c=1}^{N}\frac{\sum\left(I_{1_c}-\mu_{I_{1c}}\right)\cdot\left(I_{2_c}-\mu_{I_{2c}}\right)}{\sqrt{\sum\left(I_{1_c}-\mu_{I_{1c}}\right)^2}\cdot\sqrt{\sum\left(I_{2_c}-\mu_{I_{2c}}\right)^2}}$$



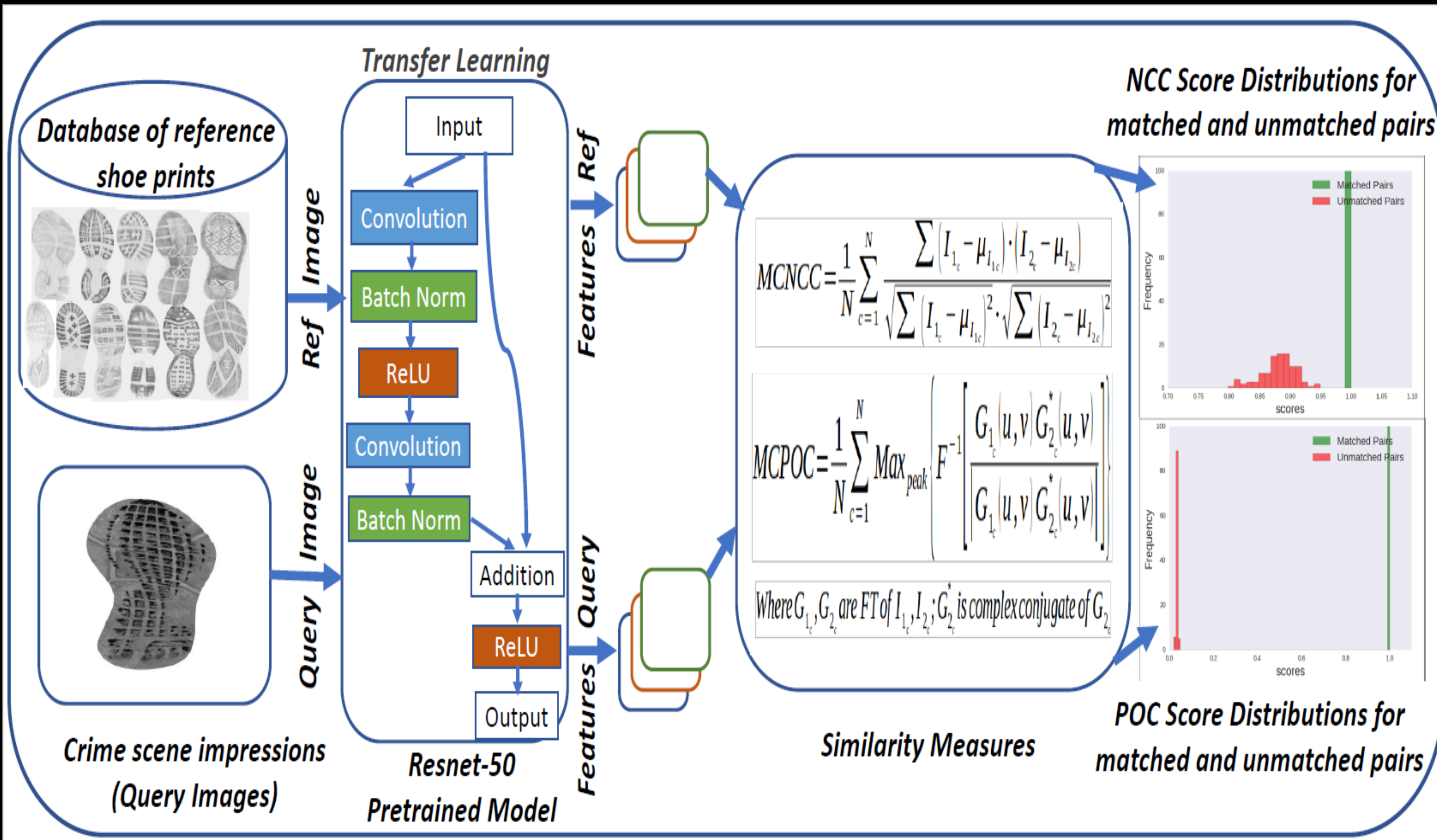
Phase Only Correlation

$$MCPOC=\frac{1}{N}\sum_{c=1}^{N} Max_{peak}\left\{F^{-1}\left[\frac{G_{1_c}(u,v)\,G^{*}_{2_c}(u,v)}{\left|G_{1_c}(u,v)\,G^{*}_{2_c}(u,v)\right|}\right]\right\}$$

$Where\ G_{1_c}, G_{2_c}\ are\ FT\ of\ I_{1_c}, I_{2_c};\ G^{*}_{2_c}\ is\ complex\ conjugate\ of\ G_{2_c}$



Approach used for shoe print matching:

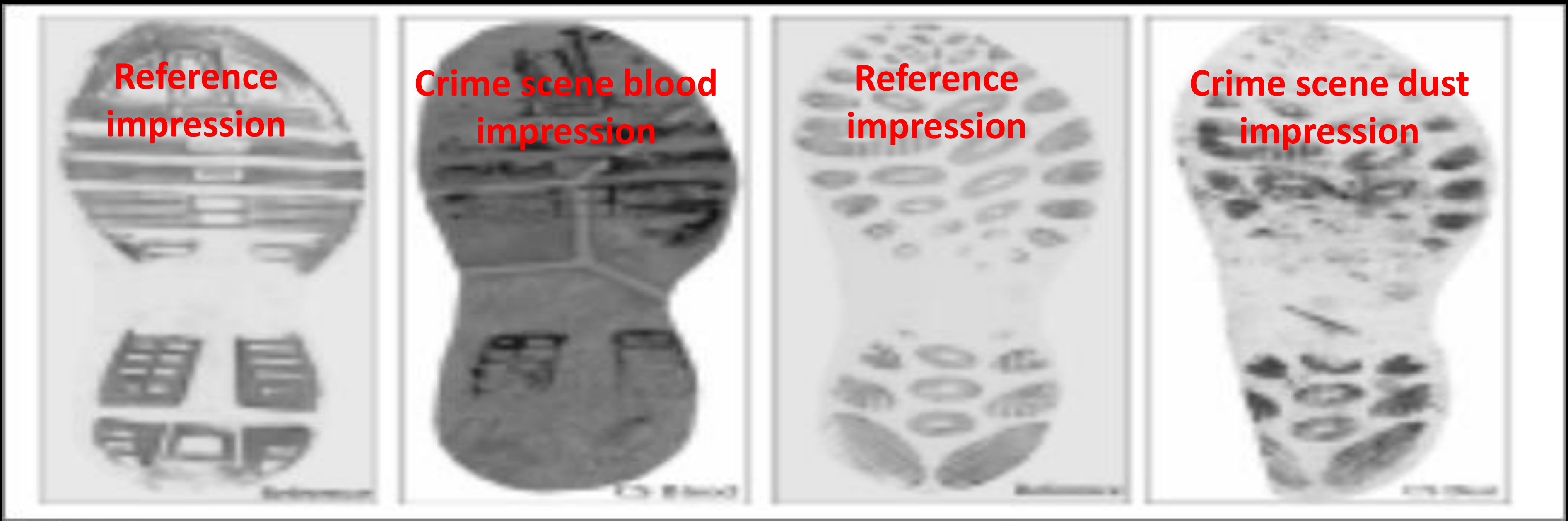


Experimental Setup:

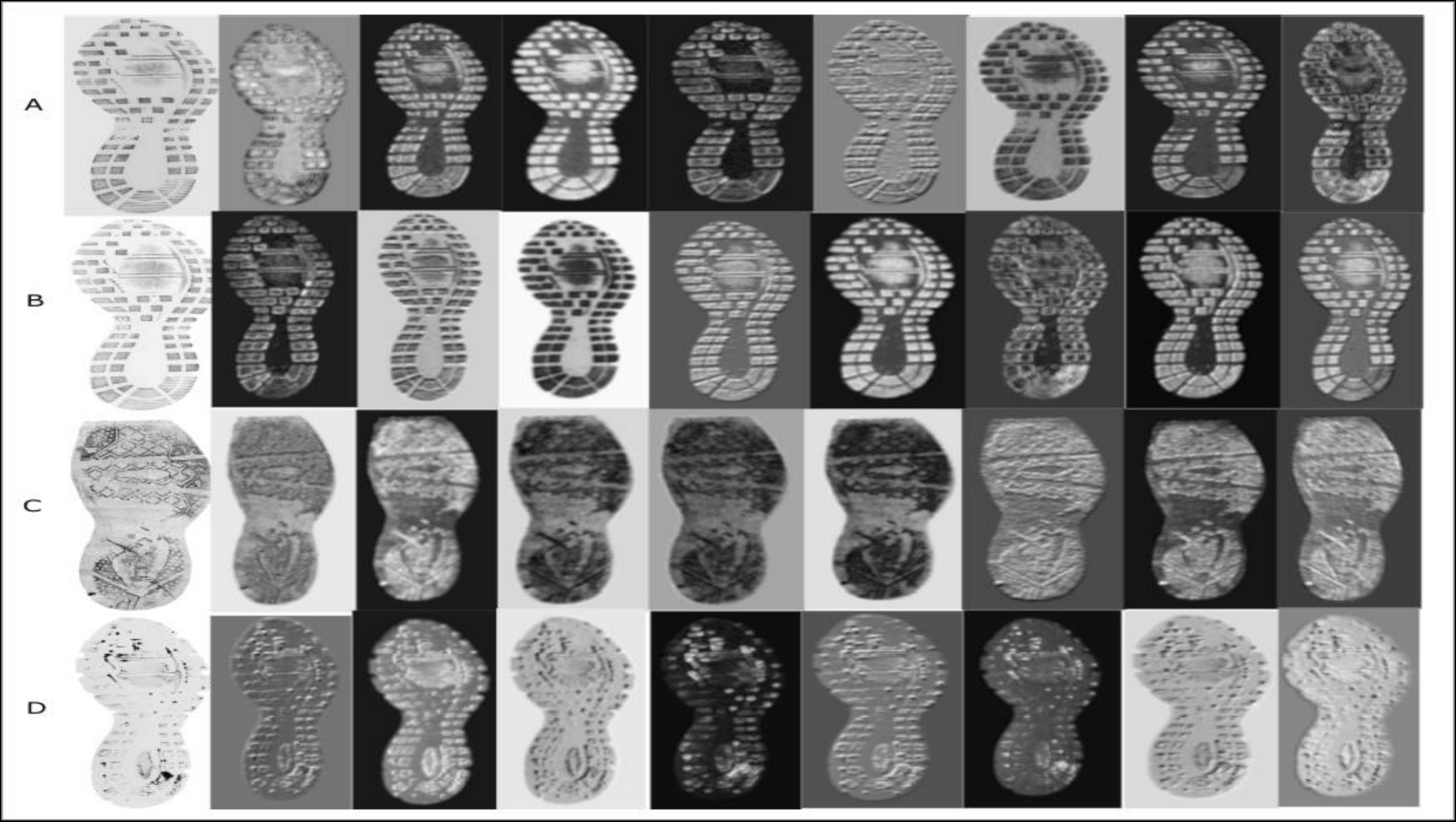
- **Datasets:** Experiments were evaluated on two sets of datasets
  - Shoe prints from WVU dataset
  - Shoe prints from FBI Boots data
- **DL framework:** **Keras DL** framework is used for experiments
- ***Model Used:*** ***Resnet – 50*** (pretrained on ImageNet data) model is used to extract features
- ***Layer:*** ***Res2a–branch–2c*** layer is considered for feature extraction (initial layer)
- **Similarity Metrics:** **MCPOC and MCNCC** scores are computed for matched, unmatched, close-nonmatched pairs
- **Feature Maps:** **256 channel features** were extracted from Resnet model
- **Scores computed:** **Average and weighted average** channel scores are used for separating the matched , unmatched and close nonmatched pairs
- **Model Evaluation:** **Receiver Operating Characteristics (ROC)** is used to evaluate the model performance

Dataset: Shoe impressions from West Virginia University (WVU)

- Nicole et.al created the Crime scene impressions using blood and dust together with three different substrates; Ceramic, Vinyl, Acetate
- This dataset is used to separate matched and unmatched pairs
- High Quality Reference impressions: **100**
- Crime Scene Dust impressions : **66**
- Crime Scene Blood : **53**
- Crime scene blood impressions were enhanced using leuco-crystal violet(LCV)
    - Crime scene Blood + LCV : **53**

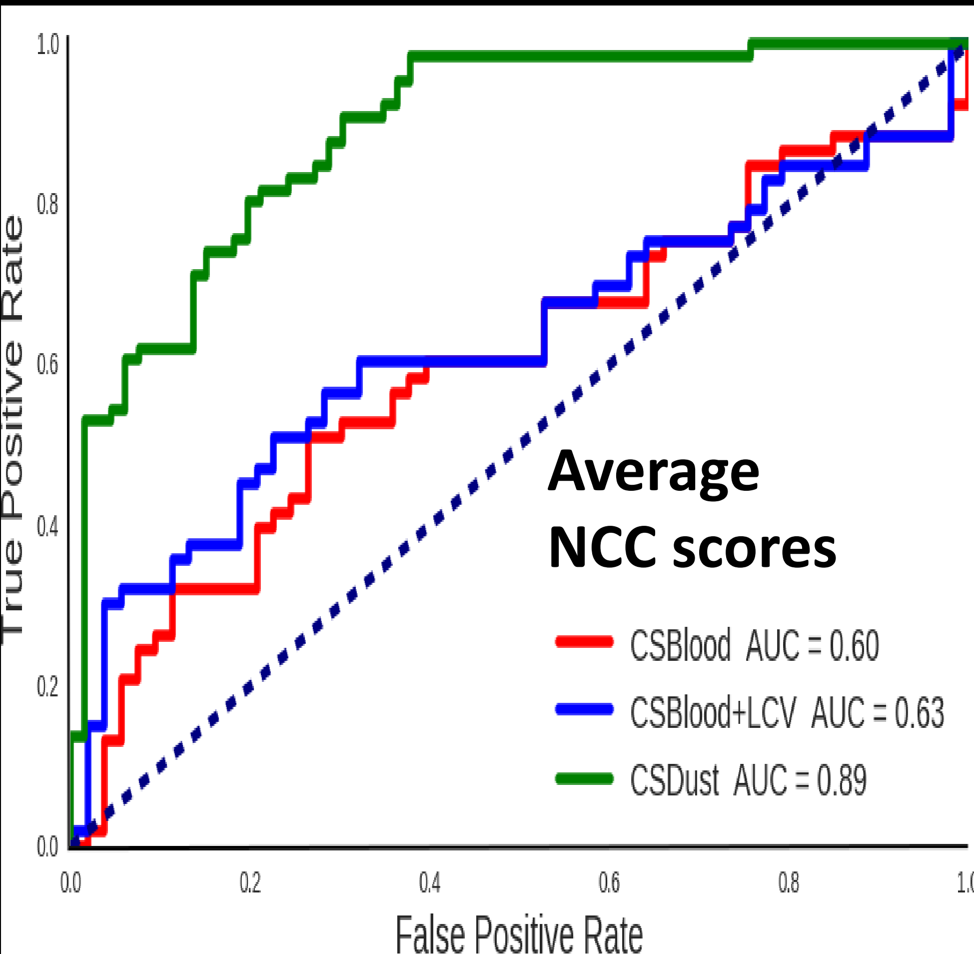


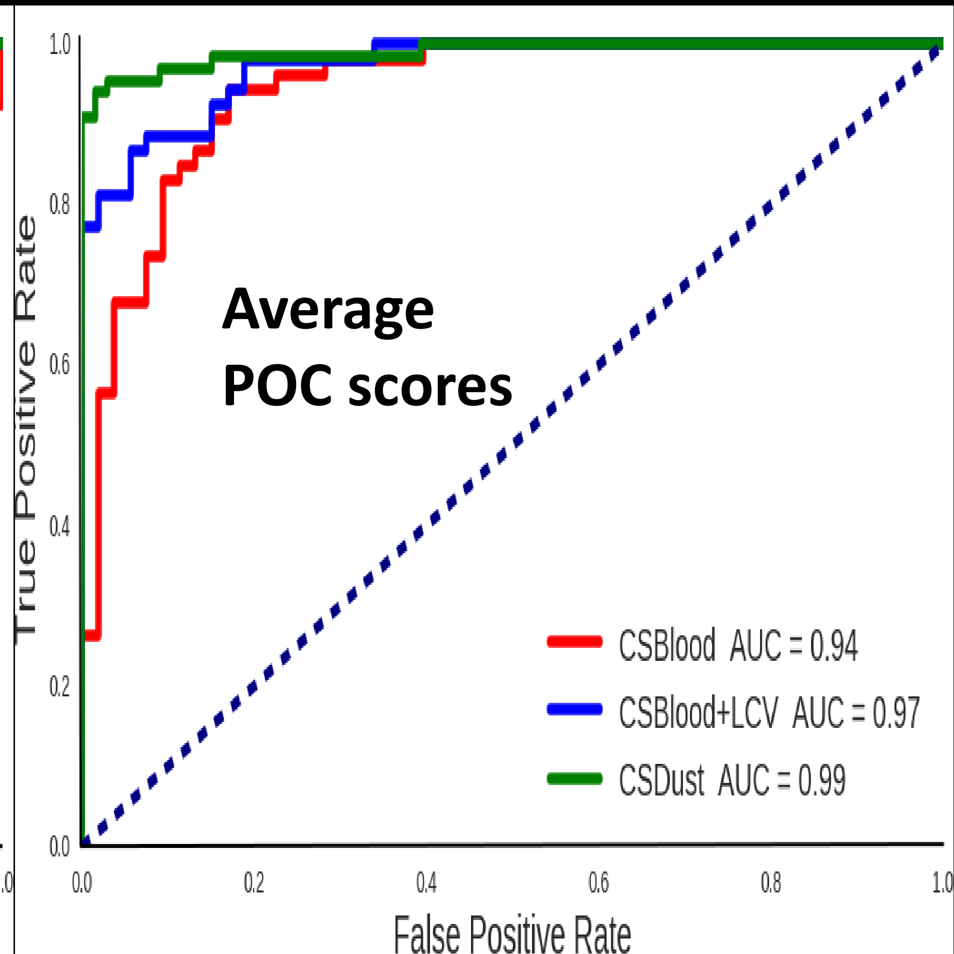
Resnet-50 model features:



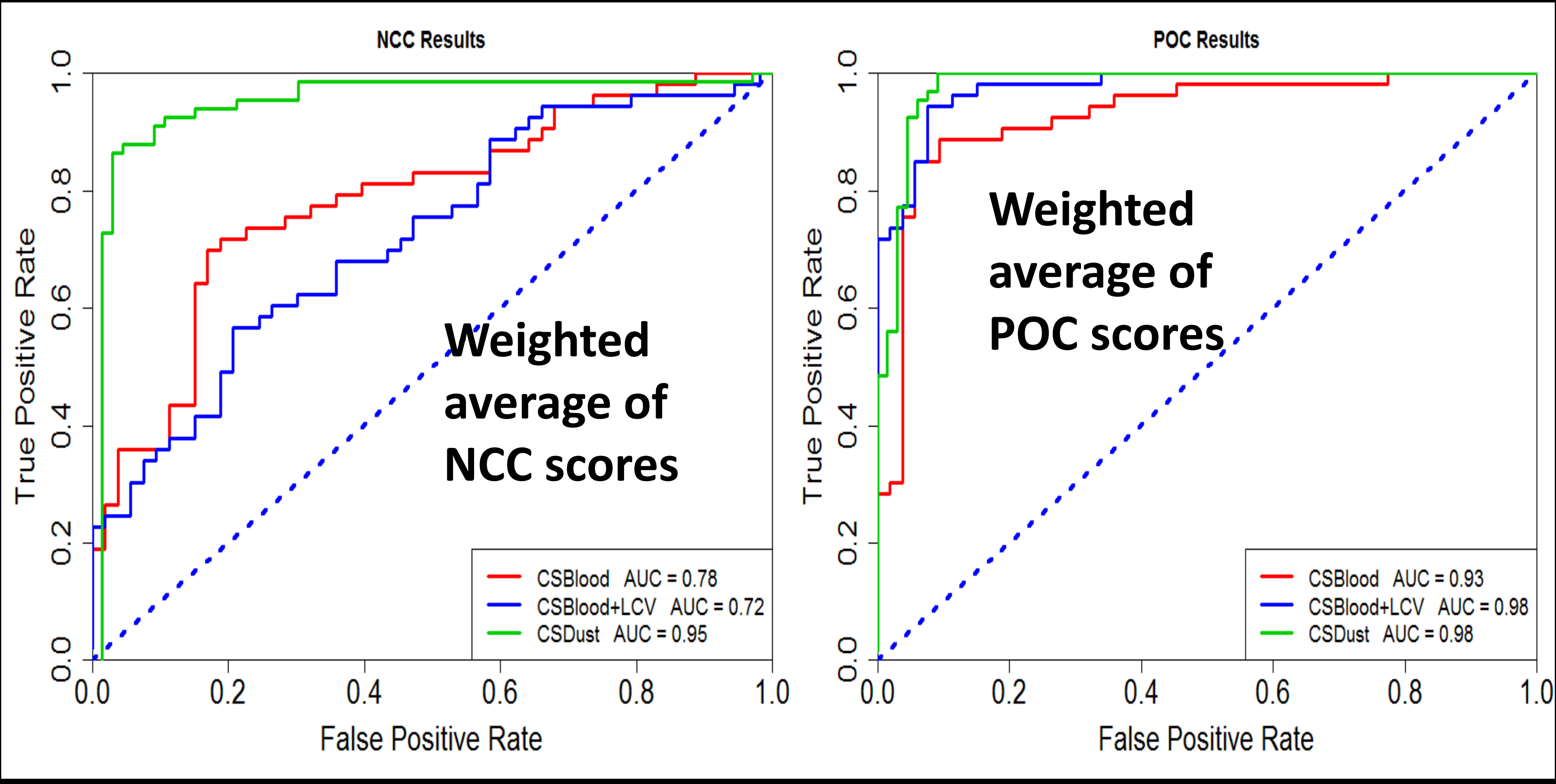
Images of WVU dataset and Resnet model features. A) High quality reference B) Query Impressions C) Crime Scene Dust D) Crime scene Blood

Experimental Results (1) : WVU dataset



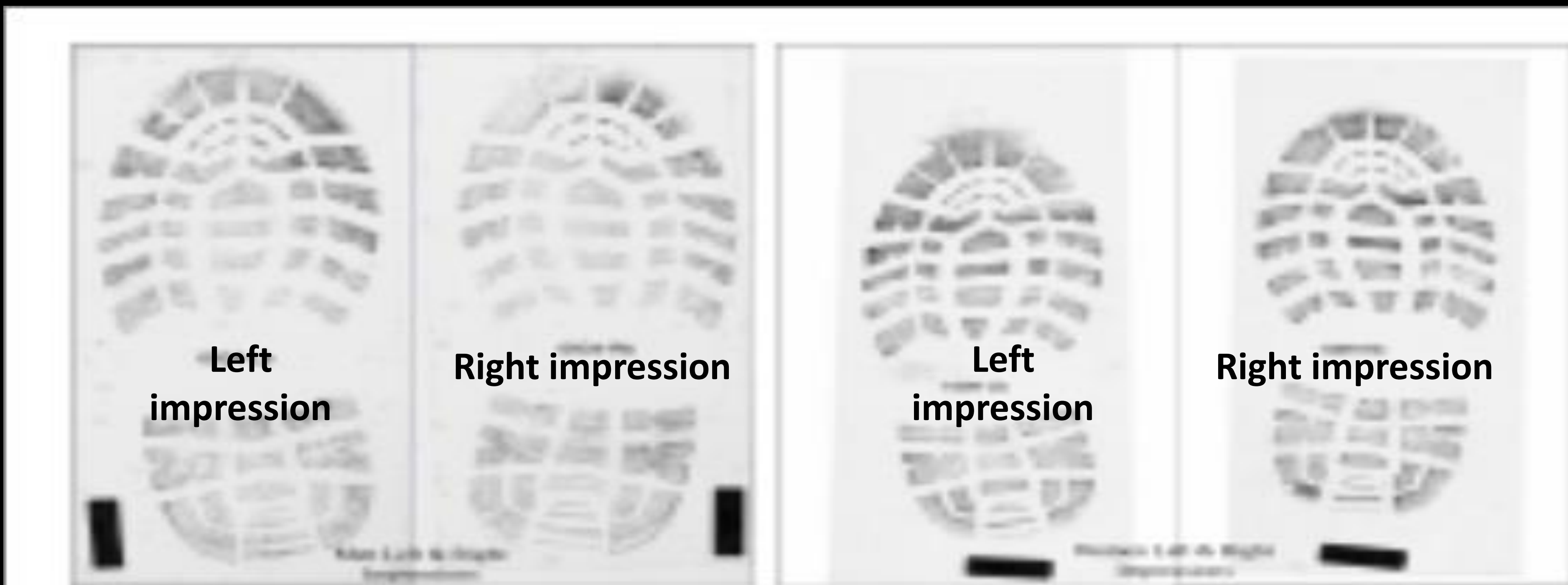


Experimental Results (2) : WVU dataset with LASSO regressor

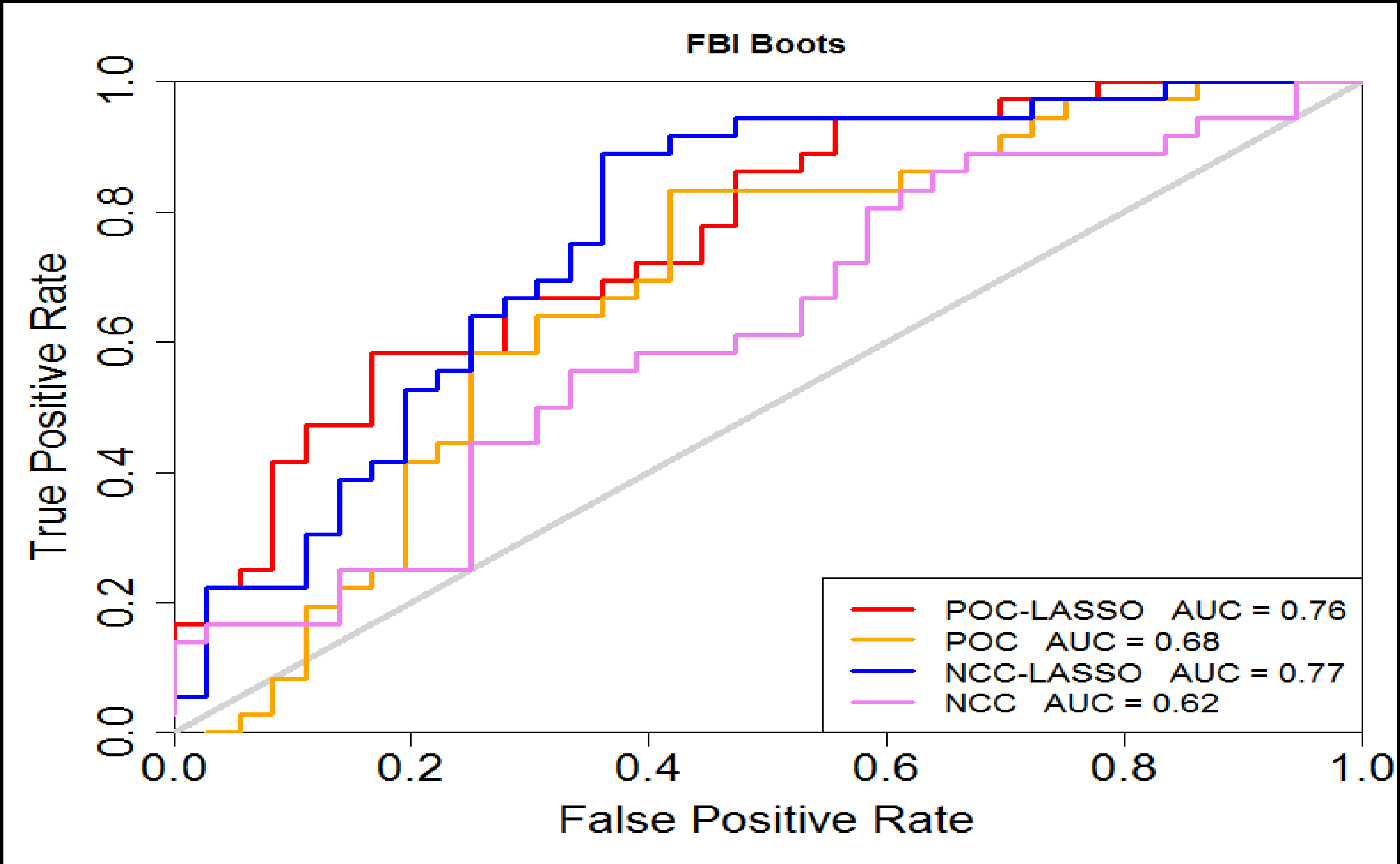


Dataset: FBI Boots dataset

- This dataset is used to separate matched vs close nonmatched pairs
- There are **72** pairs of impressions with same make and model
- Size and wear conditions vary
- There are **36** left shoe and **36** right shoe impressions
- These impressions are used to study the how well Resnet model features can discriminate between matched and close nonmatched pairs



Experimental Results (3) : FBI Boots dataset



Conclusions:

- **Matched vs Unmatched pairs:** DL based feature descriptors show good promise in separating matched and unmatched pairs
- **Matched vs Close nonmatched:** The separation of matched pairs from close-nonmatched pairs is not as good as separation of matched pairs from general non-matched pairs. This is to be expected and indicates that unique features (RACs) are important for discrimination in such cases.
- **Similarity metric:** Multi-channel phase-only correlation performs better than multi-channel normalized cross correlation
- **Future Work:**
  - As pretrained models are successful for shoe print matching, it is worth to explore DL models to address current challenges, namely, alignment, scale and modality differences
  - It is also worth exploring the additional training of these models specifically for separating matched and close nonmatched pairs

References:

Deep Learning:

- He, Kaiming, et al. ***Deep residual learning for image recognition***. *IEEE conference on computer vision and pattern recognition*. 2016.
- Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. ***How transferable are features in deep neural networks?*** In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, NIPS'14, pages 3320–3328, Cambridge,MA, USA, 2014. MIT Press.
- Goodfellow Ian, Bengio Yoshua, Courville Aaron and Bengio, Yoshua. ***Deep Learning***. MIT press Cambridge, volume 1, year 2016.

Shoe Print Matching:

- Bailey Kong, Deva Ramanan, and Charless Fowlkes. ***Cross-domain forensic shoeprint matching.*** In *British Machine Vision Conference (BMVC)*, 2017.
- Nicole Richetelli, Mackenzie C Lee, Carleen A Lasky,Madison E Gump, and Jacqueline A Speir. ***Classification of footwear outsole patterns using fourier transform and local interest points***. *Forensic science international*, 275:102–109, 2017a.

ACKNOWLEDGEMENTS:

I would like to thank Marty Herman, Hari Iyer , Steve Lund, and Clement Driard for helpful discussions and support.

Thank You