Fingerprint Image Enhancement Using Unsupervised Hierarchical Feature Learning

Presented in partial fulfilment of the requirements of MS by Research

Mihir Sahasrabudhe
200802023
Adviser: Dr. Anoop Namboodiri

International Institute of Information Technology - Hyderabad

October 17, 2014

Outline

Part 1

Background

Biometrics
Fingerprint Recognition
Scope of the Thesis

Biometrics

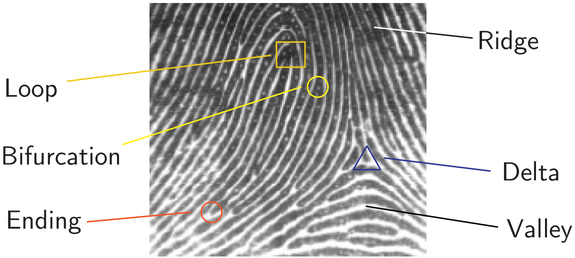
- Biometric traits: anatomical and/or behavioural
- Identify/verify individuals
- *Employing biometric traits for automatic recognition/verification of individuals*
- Face, fingerprint, iris, hand geometry, signature, voice, etc.

Fingerprints

- Impression left by a finger
- Caused by presence of skin patterns and moisture/dirt/oil
- Have been used for over a century
- Uniqueness is assumed

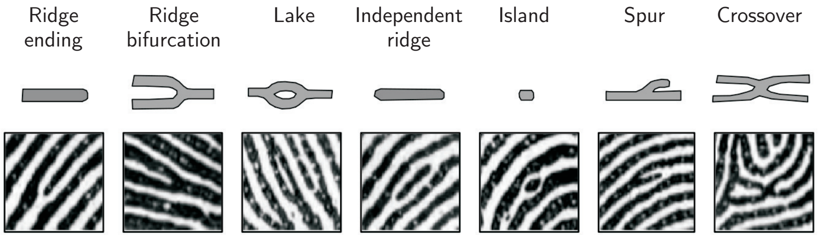
Patterns and Minutiae

- A fingerprint is a pattern of ridges and valleys
- Points of interest include ridge-endings and ridge-bifurcations
- Ridges might form loops and deltas (singularities)



Patterns and Minutiae

Most commonly found ridge patterns in fingerprints:

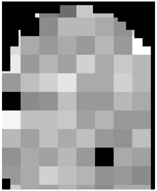


Intrinsic Images

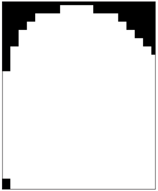
- Intrinsic properties: orientation field, frequency and region mask.
- **Orientation**: ridge direction at a point
- **Frequency**: number of ridges per pixel, $\perp$ to orientation
- **Region mask**: places where a fingerprint is present

Original

Orientation Field



Frequency Image



Region Mask

Fingerprint Enhancement

- Improve ridges and valleys
- Make them easily distinguishable
- Remove creases, cuts, and improve too wet/too dry regions.

Figure: from *left* to *right*: Fingerprints of decreasing quality

Fingerprint Enhancement

- *Low-pass filtering along the orientation* and *band-pass filtering $\perp$ to the orientation*

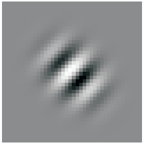
Figure: *left:* the original image; *centre:* the enhanced image; *right:* the binarised image

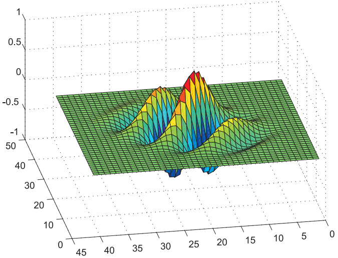
Fingerprint Enhancement Using Contextual Filters

Using Gabor filters [Hong et al. (1998)]: Two-dimensional gaussian, multiplied by a cosine along one dimension:

$$G(x, y : f, \theta, \phi, \sigma_1, \sigma_2) = \exp\left(-\frac{x'^2}{2\sigma_1^2} - \frac{y'^2}{2\sigma_2^2}\right)\cos\left(2\pi x' f + \phi\right)$$

where $x' = x\cos\theta + y\sin\theta$, and $y' = -x\sin\theta + y\cos\theta$





Fingerprint Enhancement in Frequency Domain

STFT analysis: Ridge pattern in a small region is like a 2-dimensional sine wave (Chikkerur *et al.* 2006)

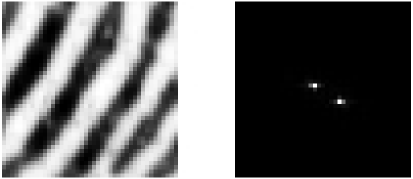


Figure: *left*: A patch from a fingerprint; and *right:* its power spectrum

- Compute and apply appropriate filters

Scope of the Thesis

- Fingerprint image enhancement
- Unsupervised feature learning
- Feature learning on fingerprint images
- Enhancement of fingerprints using learnt features
- Estimation of intrinsic images of fingerprints

Part 2

Learning Local Orientation Fields

Idea

- Consideration of neighbourhood of a point in enhancement
- A limited number of ridge patterns
- Learn patterns from prints, and use the learning to estimate noisy regions
- Using feature learning, we can “correct” damaged patterns
- Unsupervised feature learning to learn and correct patterns

Idea

- RBMs can reconstruct patterns based on learning
- Demonstrated on MNIST:

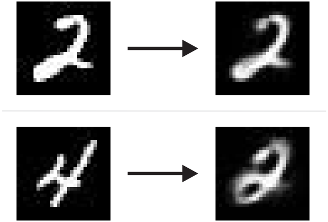
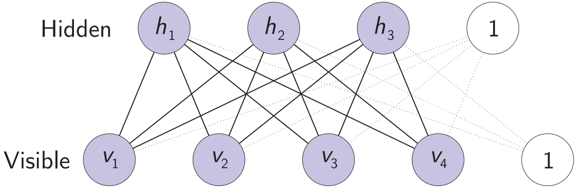


Figure: An RBM trained on images of handwritten 2s reconstructs a 4 to look like a 2.

Restricted Boltzmann Machine (RBM)

- Smolensky (1986)
- Became popular after introduction of fast learning algorithms (Hinton)
- 2-layered networks that learn a non-linear subspace



Restricted Boltzmann Machine (RBM)

- Can be thought of as dimensionality reduction models
- Each unit in the hidden layer learns a feature
- Unsupervised learning

Features learnt from a dataset of handwritten digits:

Restricted Boltzmann Machine (RBM)

- Infer hidden units from visible, and vice versa. $V \leftrightarrow H$
- $H = \sigma(W \cdot V + C)$; $V = \sigma(W' \cdot H + B)$
- Weight updates at each iteration of training
- $\Delta W_{i,j} \propto \langle v_i h_j \rangle^0 - \langle v_i h_j \rangle^\infty$

Continuous Restricted Boltzmann Machine (CRBM)

- RBMs are classically only binary-valued
- Continuous RBMs can take inputs in $[0, 1]$
- Chen and Murray (2003)

Gabor-based Enhancement

- Hong *et al.* (1998)
- Enhancement using oriented Gabor fitlers
  1. Segmentation
  2. **Compute orientation field using gradient-based method**
  3. Estimate frequency image
  4. Prepare bank of Gabor filters
  5. Apply appropriate filters from the bank

Gabor-based Enhancement

- Significant neighbourhood not considered
- Irregularities, if noise is present

Figure: *left*: original image; *centre*: bad orientation estimate; *right*: better estimate (Rama Reddy, 2011)

The Model

- Need: A framework to learn local orientation fields
- We will use continuous RBMs to learn
- Training data is clean, noise-free orientation fields

The Model

- Break down each training orientation field θ into $(s(\theta), c(\theta))$
- Orientation fields might have sharp transitions

Patch

Θ

- Sharp transitions are removed in the decomposed images

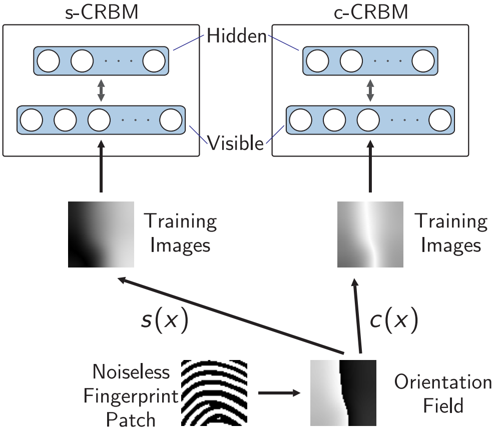
$s(\Theta)$

$c(\Theta)$

- Learning and reconstructions are made easier

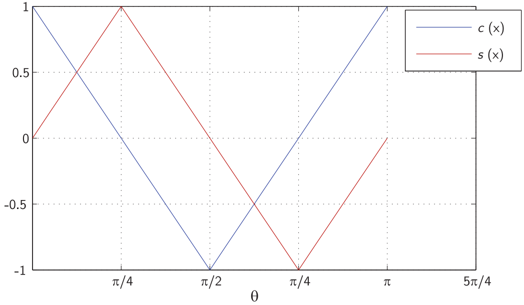
The Model

The model, illustrated with a diagram:



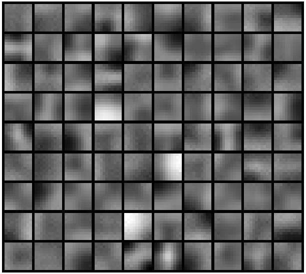
The Model

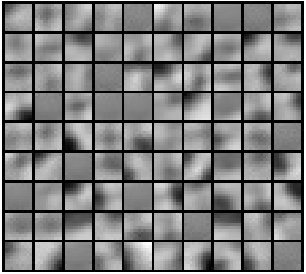
$s(x)$ and $c(x)$ are:



Training the Model

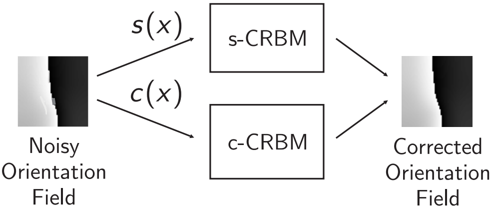
- Train both CRBMs with the respective training data
- Training patches are resized to 10×10 from 60×60
- Each resized image is flattened into a vector
- Learnt weights for c-CRBM and s-CRBM:





Using the Trained Model

- Split and feed test orientation fields extracted by gradient-based method to CRBMs
- Use the reconstructions and combine them
- Compute appropriate Gabor filters and apply them



Qualitative Results

Some examples of enhancements using Continuous RBMs:

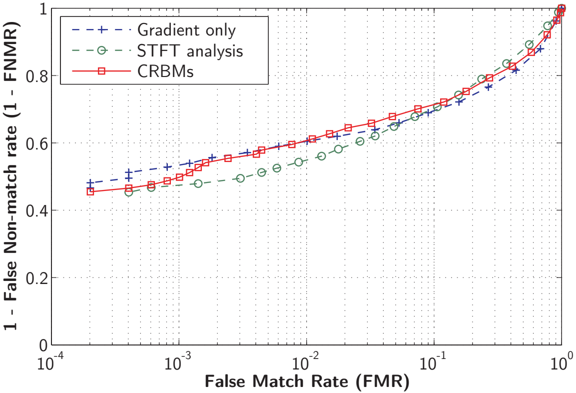
Patches from greyscale images

Gradient-only enhancement

After correction using CRBMs

Quantitative Results

Receiver Operating Characteristics on FVC 2002 Db3_a:



Quantitative Results

Spurious and missing minutiae analysis (FVC 2002 Db3_a):

Method	Gradient-only	STFT	CRBMs
Ground Truth	19032		
Detected	53389	48936	**41674**
Spurious	38415	33096	**26324**
Missing	4058	**3165**	3592
EER	24.34	21.99	22.65

Part 3

Hierarchical Learning of Fingerprint Features

Idea

- Feature extraction directly from greyscale images
- Hierarchical learning
- Reconstruction/inference using learnt features
- Estimating orientation field, frequency image and mask

Idea

Need

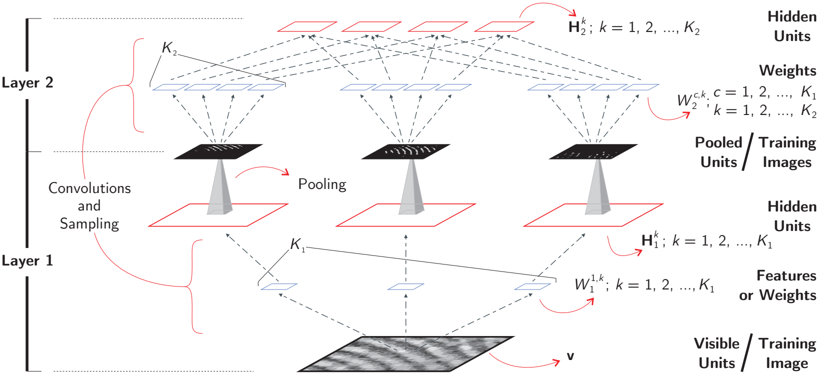
- Unsupervised feature learning from greyscale images
- A hierarchical model
- A scalable model
- A generative model
- A way to combine top-down and bottom-up signals

Convolutional Deep Belief Networks

- Conv DBNs fit the requirements
- Lee *et al.* (2009)
- Hierarchical, scalable and generative
- Hierarchical probabilistic inference

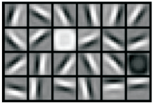
Convolutional Deep Belief Networks

A two-layered network:



Convolutional Deep Belief Networks

Layers learn edges, object parts, and objects:



Layer 1

Layer 2

Layer 3

Lee *et al.* (2009)

Convolutional Deep Belief Networks

- $H_k \leftarrow f(V * \tilde{W}_k + b_k)$; f is a non-linearity
- Probabilistic max-pooling to generate pooling units
- $V \leftarrow \sum_k H_k * W_k$
- $\Delta W_k \propto \frac{1}{N_H^2}\left(\langle H^k * V\rangle^0 - \langle H^k * V\rangle^1\right)$

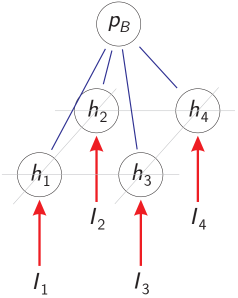
Probabilistic Max Pooling

- Downsample images while preserving features

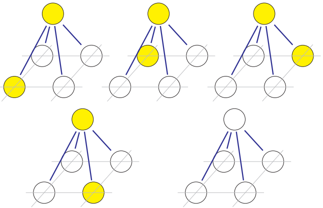
Bottom-up inference:

$$P(h_i = 1) = \frac{\exp(I_i)}{1 + \sum_j \exp(I_j)}$$

$$P(p_B = 1) = \frac{\sum_j \exp(I_j)}{1 + \sum_j \exp(I_j)}$$



Possible combinations:



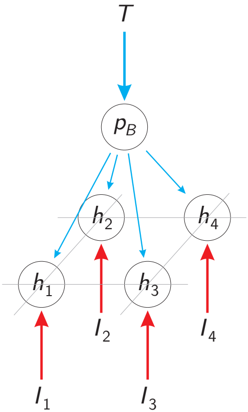
Probabilistic Max Pooling

Combines top-down and bottom-up signals (signals from two layers)!

$$P(h_i = 1) = \frac{\exp(T + I_i)}{1 + \sum_j \exp(T + I_j)}$$

$$P(p_B = 1) = \frac{\sum_j \exp(T + I_j)}{1 + \sum_j \exp(T + I_j)}$$

where T is the top-down signal.



Model Parameters

We used the following model in this work:

- 2-layered network
- 60 features in each layer
- 200 epochs of training for each layer
- L2 and sparsity regularisation
- 15 training images from different datasets

Training

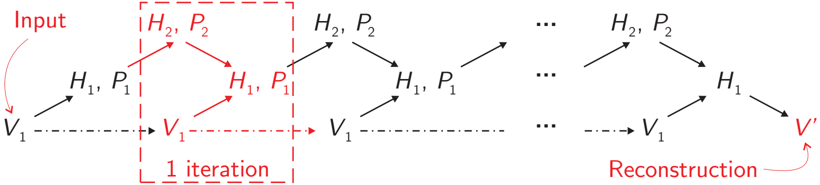
- Training is done similar to deep belief nets, Hinton *et al.* (2006)
- Greedy layer-wise
- Lower layer's weights frozen once trained
- Pooling units serve as training data for next layer

Hierarchical Probabilistic Inference

- Reconstructing an image from its representation
- Hidden(ℓ) conditioned on Visible(ℓ) + (signal from $\ell + 1$)
- 'Odd' numbered layers grouped together, and so are 'even' numbered ones

Enhancing Fingerprints

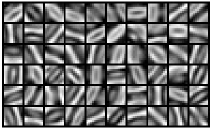
- Use HPI to reconstruct an image
- 20 iterations
- Hidden units in both layers are affected at each iteration
- Final activations of L1's hidden units used in reconstruction



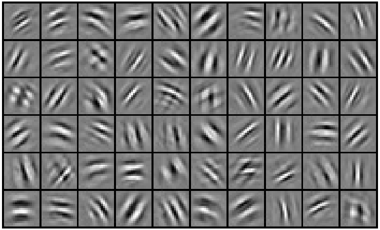
Reconstruction, $V' = \sum_k W_1^{1,k} * H_1^k$, where H_1^k denotes the k-th set of hidden units of Layer 1 after 20 iterations

Learnt Features

- Layer 1: oriented ridges
- Layer 2: Local ridge structures



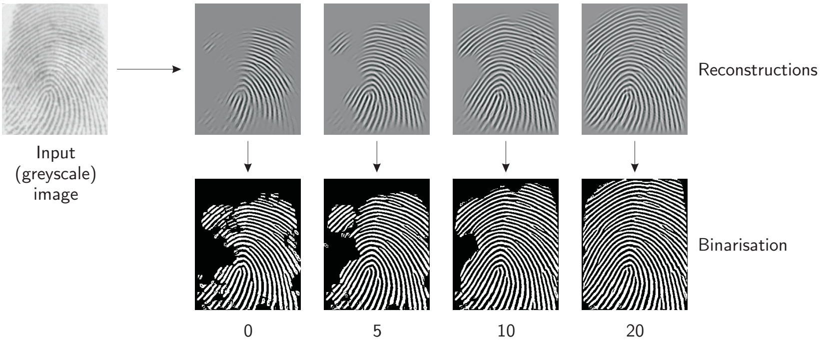
Layer 1



Layer 2

What Happens During Enhancement (Reconstruction)

- Layer 1 detects learnt features in the input image
- Layer 1 alone is unable to interpret some regions
- These are estimated by Layer 2 using surrounding, interpreted regions (combining top-down and bottom-up signals)

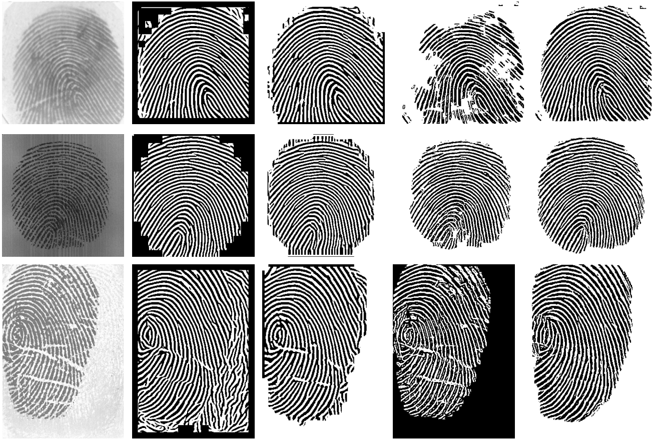


Estimating Intrinsic Images

- Orientation field, region mask, and frequency image estimated by the network
- $\mathbf{D} = \frac{1}{2} \arctan \frac{\sum_k \sin(2d_k) \left(W_1^{1,k} * H_1^k \right)}{\sum_k \cos(2d_k) \left(W_1^{1,k} * H_1^k \right)}$ followed by smoothing
- $\mathbf{F} = \left(\sum_k f_k \left(W_1^{1,k} * H_1^k \right) \right) \oslash V'$ followed by smoothing
- f_k and d_k computed for each $W_1^{1,k}$ from its Fourier transform
- $[\mathbf{M}]_{x,y} = ([V']_{x,y} \neq 0)$

Qualitative Results

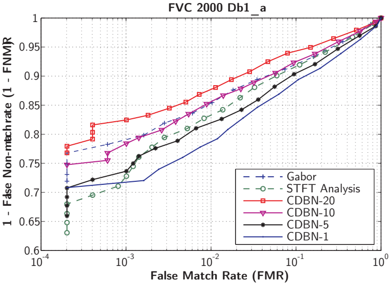
Enhancements using CDBN:

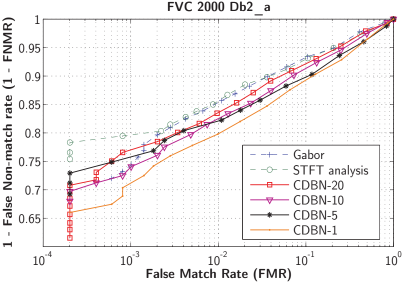


Original Gabor STFT CDBN-1 CDBN-20

Quantitative Results

Receiver Operating Characteristics:





Quantitative Results

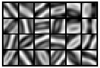
Equal-error rates on several datasets:

Dataset	Gabor	STFT	CDBN-20	CDBN-10	CDBN-5	CDBN-1
2000 Db1_a	8.47	8.79	6.62	8.19	9.59	10.57
2000 Db1_b	9.46	6.67	5.82	6.55	11.65	13.11
2000 Db2_a	7.71	7.98	8.52	9.14	10.24	10.66
2000 Db2_b	14.00	9.24	10.44	17.32	16.29	19.65
2002 Db3_a	24.34	21.99	23.95	25.00	25.45	24.48

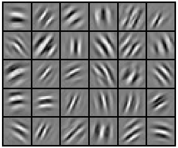
With different number of features

- Two networks: 20 features in L1, 30 in L2; 40 in L1, 60 in L2

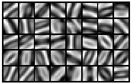
Layer 1



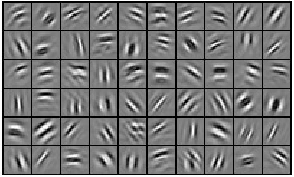
Layer 2



20 L1; 30 L2



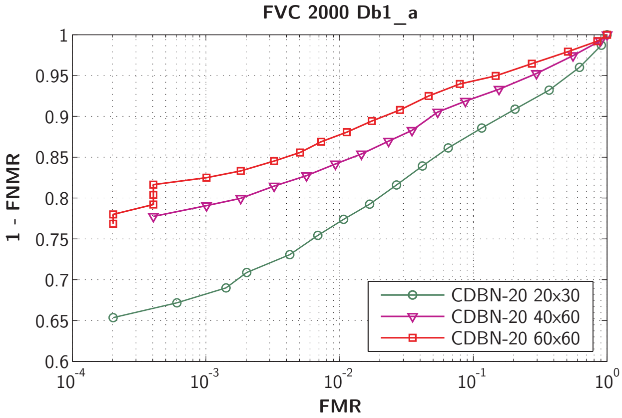
Layer 1



Layer 2

40 L1; 60 L2

With different number of features



With different number of features

- Networks with few weights aren't able to reconstruct completely
- Some features might not be learnt

Original

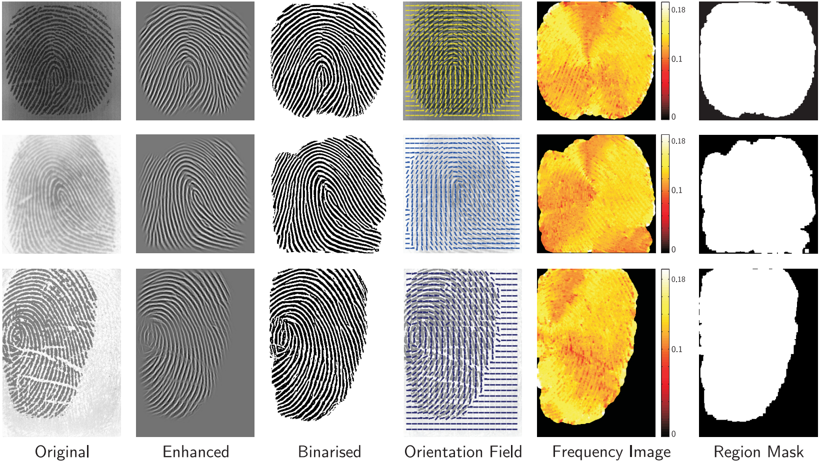
20 x 30

40 x 60

60 x 60

Intrinsic Images

Estimated intrinsic images:



Part 4

Epilogue

Conclusion

- Feature learning applied to fingerprints
- Feature extraction using neural nets
- Orientation field learning and reconstruction
- Hierarchical learning boosts enhancements
- Same model computes intrinsic images

Future Work

- Addings more layers
- Synthetic fingerprint generation
- Segmentation
- Minutiae detection
- Classification

Related Publications

1. **Learning Fingerprint Orientation Fields Using Continuous Restricted Boltzmann Machines**
   Mihir Sahasrabudhe and Anoop M. Namboodiri.
   2nd Asian Conference on Pattern Recognition, 2013.

2. **Fingerprint Enhancement Using Unsupervised Hierarchical Feature Learning** (Oral)
   Mihir Sahasrabudhe and Anoop M. Namboodiri.
   9th Indian Conference on Vision, Graphics and Image Processing, 2014. (to appear)

Thank you!