# Face Parts Labelling

CS365 – COURSE PROJECT

Under the guidance of Prof Amitabha Mukherjee
Krishna Chaitanya (11372)
Vikas jain (13788)

# Introduction

AIM - Segmenting an image into three regions and labelling them with hair, skin and background labels

For e.g.

Original Image

Segmentation


*

Labelling

MOTIVATION - Used for mid-level computer vision task: **Face recognition**

*Example from reference [1]

# Preliminaries

## Conditional Random Field (CRF)

They are graphical models well suited to modelling local interactions among adjacent regions(e.g. superpixels).
The conditional distribution and the energy function can be defined as follows:

$P_{crf}(Y|X) \propto exp(-E_{crf}(Y, X))$
$E_{crf}(Y, X) = E_{node}(Y, X_V) + E_{edge}(Y, X_E)$, *where*

*Y* refers to the region label,
$X_V$ refers to the region node feature vector,
$X_E$ refers to the edge region feature vector,
$E_{node}$ and $E_{edge}$ are the linear functions of labels, feature vectors and edge weights($\Gamma,\psi$).

## Restricted Boltzmann Machine (RBM)

It is a bipartite undirected graphical model composed of visible and hidden layers of nodes.
The joint distribution can be defined as follows:

$P_{rbm}(Y, h) \propto exp(-E_{rbm}(Y, h))$, *where*

*Y* refers to the label of the visible units,
*h* refers to the label of the hidden units.
$E_{rbm}$ is a linear function of labels, edge weights(*W*) and biases(*b,c*) of the hidden and visible nodes.

# APPROACH

We combine local consistency of CRF and global consistency (shape prior) of RBM to get the best labelling.

This is done by describing the condition likelihood of the labels *Y* given the **superpixels** features *X* as follows :

$P_{gloc}(Y|X) \propto \sum_h exp(-E_{gloc}(Y, X, h))$

$E_{gloc}(Y, X, h) = E_{crf}(Y, X) + E_{rbm}(Y, h)$, where the terms are as defined before.

The model parameters {*Γ, Ψ, W, b, c*} are trained to maximize the conditional log likelihood of the training data

$$\max_{(W,b,c,\Gamma,\psi)} \sum_{m=1}^{M} \log P_{gloc}(Y^{(m)}|X^{(m)})$$

# ALGORITHM

LEARNING

- We maximize the conditional log likelihood using contrastive divergence.
- It relies on the approximation of the gradient of the log likelihood based on a short Markov chain.

INFERENCE

- Since the joint inference of superpixel labels and hidden nodes is intractable, we use mean-field approximation
- The approximated distribution is such that it minimizes the Kullback-Leibler distance between the approximate and original distribution.

# Features of Superpixels

## NODE FEATURES

1. Color
2. Position
3. Texture

## EDGE FEATURES

1. Sum of probabilities of boundary along the border
2. Euclidean Distance between mean color histogram
3. Chi-squared distance between texture histograms

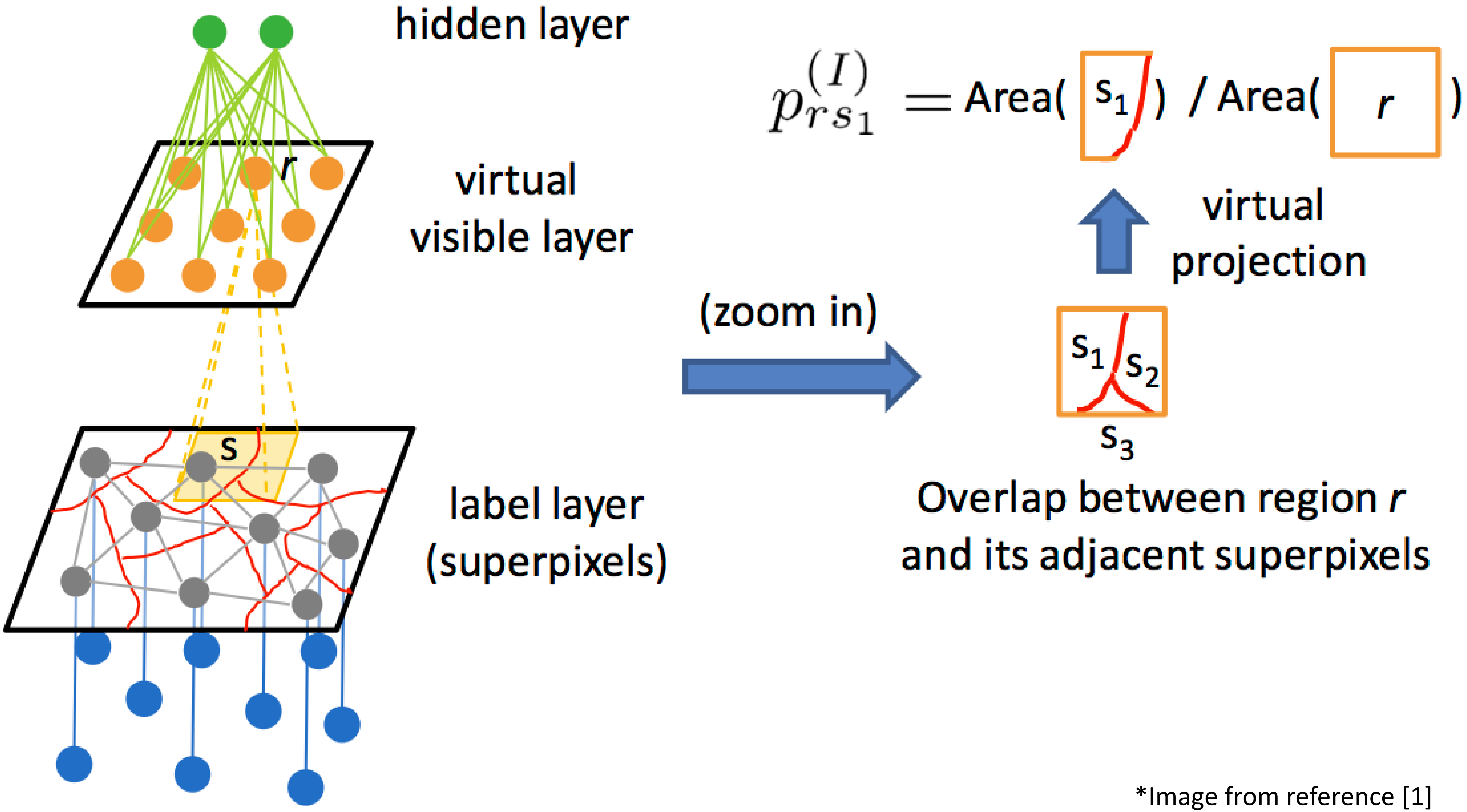


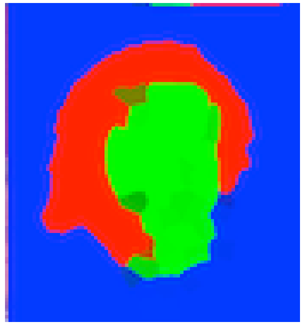

*Image from reference [1]

# EXPERIMENT AND RESULTS

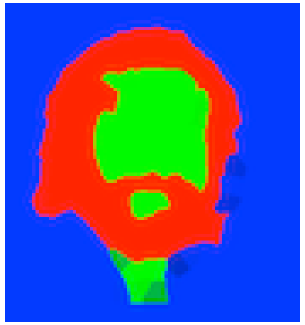
Number of Visible nodes - 576 (24×24)
Number of hidden nodes – 400
Image Size - 250×250 pixels
Number of Superpixels – 200-250 per image

| Method | Accuracy | Error Reduction |
|---|---|---|
| CRF | 93.2256% | - |
| GLOC | 94.946% | 25.39% |

Original Image



CRF



GLOC

# FURTHER WORK

- Creating a platform/Application to label a given image according to the training of the model.
- Extending the model to deeper architecture using Deep Boltzmann Machine(DBM).

# REFERENCES


- Kae, Andrew, et al. "Augmenting CRFs with Boltzmann machine shape priors for image labeling." *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*. IEEE, 2013.
- Asia Fischer and Christian Igel. An introduction to restricted boltzmann machines. In Luis Alvarez, Marta Mejail, Luis Gomez, and Julio Jacobo, editors, *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, volume 7441 of *Lecture Notes in Computer Science*, pages 14{36. Springer Berlin Heidelberg, 2012..
- Database used: Part Label Database http://vis-www.cs.umass.edu/lfw/part_labels/
- Code used: http://vis-www.cs.umass.edu/code/gloc/gloc.zip

# THANK YOU

QUESTIONS??