Face Part Labelling

Krishna Chaitanya, 11372   Vikas Jain, 13788
Department of Computer Science and Engineering, IIT Kanpur
Mentor – Prof Amitabha Mukherjee

INTRODUCTION

*Image segmentation is the process of dividing a digital image into **superpixels**. Region labelling, then assigns specific names to those segments. They are generally used for locating objects and deciding boundaries. In this work, the task of segmenting and labelling **face images** into 3 regions : **hair**, **skin** and **background** is addressed.*


Original Image


Segmentation


Labelling

[1]

MOTIVATION

- Grouping and organising image regions into logical and consistent parts are critical mid-level computer vision tasks.
- It is primarily used for
  - Face Recognition
  - High-level features extraction such as hair length, gender and pose.[2]

PREVIOUS WORK

- Conditional Random Fields (CRF) were used for segmenting and labelling images. [3]
- CRFs are useful to model the local interactions among labels for superpixels but fail in case of indistinct boundaries.

FEATURES

Node Features	Edge Features
Color	Sum of probabilities of boundary
Texture	Euclidean Distance between mean color histogram
Position	Chi-squared distance between texture histograms

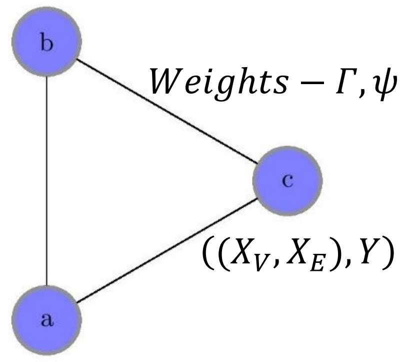
DATASET

Part Label Database
http://vis-www.cs.umass.edu/lfw/part_labels/

TECHNIQUES INVOLVED

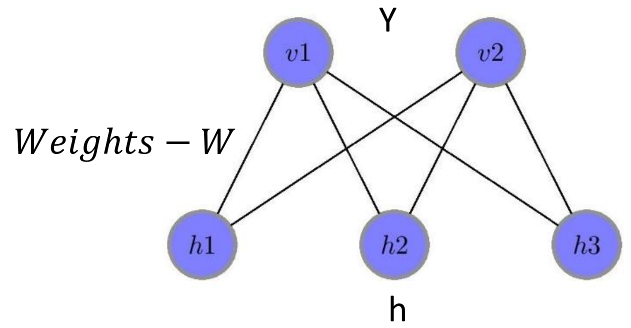
The proposed model combines two graphical models which are:
Conditional Random Fields(CRFs)
They are graphical models well suited to modelling local interactions among adjacent regions(e.g. superpixels).
The conditional distribution and the energy function can be defined as follows:

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



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:

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



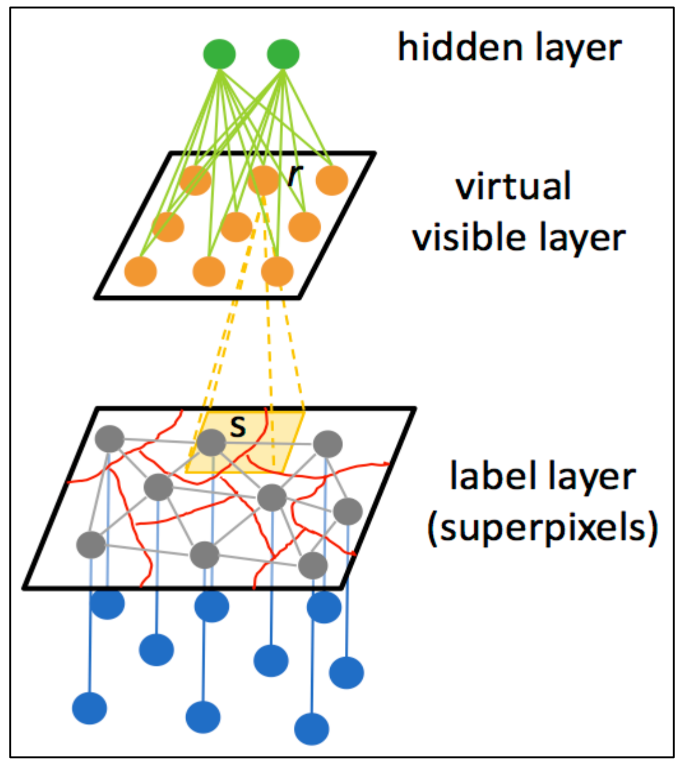
GLOC MODEL APPROACH

We combine *local consistency* of **CRF** and *global consistency* (shape prior) of **RBM** to get the best labelling (GLO + LOC) [1] .
This is done by describing the condition likelihood of the labels *Y* given the **superpixels** features *X* as follows :

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

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

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



*from [1]

ALGORITHM USED

Superpixel and Feature Generation
- Standard Algorithm were used. Code available on http://vis-www.cs.umass.edu/code/gloc/gloc_features.zip

Learning
- Maximize of 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, mean-field approximation is used.
- The approximated distribution is such that it minimizes the Kullback-Leibler distance between the approximate and original distribution.

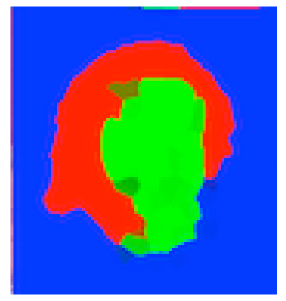
EXPERIMENTS AND RESULTS

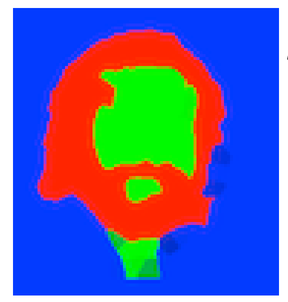
CRF and GLOC model are run on the database.
Number of Data Set Images:
- Training – 1500
- Validation – 500
- Test – 927

Parameters for GLOC model were:
- No. of visible nodes: 576(24 X 24)
- No. of hidden nodes: 400
- Image Size: 250 X 250 pixels
- No. of Superpixels: 200 – 250 per image

Results obtained on the above mentioned are:
1. CRF Model
   - Accuracy: 93.3356%
2. GLOC Model
   - Accuracy: 94.946%
   - Error Reduction: 25.39%


Original Image


CRF


GLOC

[1]

FURTHER POSSIBILITY

- Generating own dataset images(around 20)
- Generating superpixels on the images and their features
- Labelling superpixels by running GLOC Model on the images.

REFERENCES

1. 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
2. Gary B Huang, Manjunath Narayana, and Erik Learned-Miller. Towards unconstrained face recognition. In Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on, pages 1–8. IEEE, 2008.
3. He, Xuming, Richard S. Zemel, and M. A. Carreira-Perpindn. "Multiscale conditional random fields for image labeling." *Computer vision and pattern recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE computer society conference on*. Vol. 2. IEEE, 2004.
4. Code used: http://vis-www.cs.umass.edu/code/gloc/gloc.zip