Multi-class semantic segmentation of faces using CRFs

Khalil KHAN[1,*], Nasir AHMAD[2], Khalil ULLAH[3], Irfanud DIN[4]

[1]Department of Electrical Engineering, University of Poonch, Rawlakot, Pakistan

[2]Department of Computer Engineering, University of Engineering & Technology, Peshawar, Pakistan

[3]Department of Electrical Engineering, National University of Computer & Emerging Sciences, Peshawar, Pakistan

[4]Department of Information Engineering, Inha University, Tashkent, Uzbekistan

*Correspondence: e.khalilkhan@gmail.com

Abstract: Multi-class semantic image segmentation is widely used in a variety of computer vision tasks such as objects segmentation and complex scene understanding. As it decomposes an image into semantically relevant regions, it can be applied in segmentation of face images. In this paper using the idea of Conditional Random Fields (CRFs), an algorithm based on multi-class semantic segmentation of faces (MSS-CRFs) is proposed. In the proposed model each node corresponds to a super-pixel while the neighbouring super-pixels are connected to nodes through edges. Unlike previous approaches which rely on three or four classes, the label set is extended here to six classes, i.e. hair, eyes, nose, mouth, skin, and background. The proposed framework is evaluated on standard face databases FASSEG, FIGARO and LFW. Experimental results reveal that the performance of the proposed model is comparable with state of the art techniques on these standard databases.

Key words: Multi-class face segmentation, conditional random fields, feature extraction, classification.

1. Introduction

Face segmentation is useful in many facial applications of computer vision such as estimation of gender, expression, age and ethnicity. Multi-class face segmentation is used as a front-end for the estimation of all mid-level vision features for these applications. In the recent years, face segmentation techniques have attracted much attention with the development of many new algorithms [1, 2, 3]. The notable factors influencing face segmentation are variations in lighting conditions, facial expressions, face orientation, occlusion and image resolution. These and many more factors make the development of an efficient segmentation algorithm a challenging task.

Many researchers around the world have solved many complicated problems of segmentation using the idea of semantic segmentation. Extensive research work has been carried out to investigate the problem with a major contribution from PASCAL VOC challenge [4].

The work reported in Huang et al. [5] has tackled the joint study of face segmentation and pose estimation. The authors have suggested that high level features such as pose, gender, and expression can be predicted easily starting from the labelling of face image into hair, skin and background. They proved that such segmentation provided useful information for the estimation of pose. Experiments were performed on a small database of 100 images. They worked on three simple poses, i.e. left, right and frontal face.

The relationship between face parts and pose is well established from psychology literature as well [6]. Also, there are compelling evidences that facial features provide useful information for human visual system to recognize the face identity [7, 8]. Hair modelling, synthesis and animation are already active research topics in computer graphics [9, 10]. Research work on face processing applications such as virtual make-up

[11], skin color beautification [12] and skin smoothing [13] have also been reported. All these applications require the precise knowledge for each face segment at pixel level. We argue that the proposed framework is a better solution as compared to state of the art for all these applications.

In the work presented in this paper, we have developed an algorithm for face segmentation using the idea of semantic segmentation and CRFs. This work is based on our previous research where a new method of face segmentation was introduced, called the Multi-Class Face Segmentation (MFS) [14]. In our previous work the problem of face segmentation was thoroughly investigated using a small database of high resolution frontal images. A built model returns a class label and probability value for each pixel. Our present work is extension of the MFS work which tries to cover the main weaknesses of MFS. Unlike the previous work, here experiments are performed on a large database of low resolution images. Manual labelling of the face segments is performed with excellent manual labelling tool. One of the main problems of MFS is the processing time. To solve the speed problem, we integrated the pipe-line with super-pixel segmentation algorithm. Similarly a conditional hierarchy for various face segments is added to the proposed new framework.

2. Related Work

A number of models for face parts segmentation and face labelling have been proposed in literature. The work of Yacoob and Davis [15] addressed problem of hair labelling. The authors adapted region growing algorithm by building a Gaussian Mixture Model (GMM). They compared appearance of hair of different people using their model. However, performance of their proposed method was affected badly with the significant changes in hair color. GMM model was further extended by Lee et al. [11]. Their algorithm segmented face image into background and hair regions. They also contributed a database of 150 manually labelled images (hair, face and skin). A super-pixel based CRFs [15] was introduced by Huang et al. [5]. They trained standard CRFs on the images taken in the wild to provide facial image labels for hair, skin and background. Kae et al. [16] combined the strength of CRFs and Shape Boltzmann Machine [17] introducing a new model named GLOC (GLObal and LOCal). Authors of the paper claimed that this hybrid model produced results better than those of CRFs alone.

Work of Yali et al. [18] focused mainly on the hair style representation and its segmentation from facial regions. Scheffler and Jean [19] work was related to segmentation of hair, skin, background and clothing. Local label consistency was enhanced by the combination of CRFs and spatial prior of each label. Matteo et al. [1] introduced a multi-classifier approach for face segmentation. They exploited color and texture information to partition a face image into four-classes (skin, hair, clothes and background). Their study focused on adaptation of the proposed technique in electronic identity documents.

A deep learning based face labeling method was proposed by Luo et al. [2]. They combined several trained models separately in which facial parts are labeled only. The method proposed by the authors does not provide complete face labeling. Sifei et al. [3] proposed a deep convolution network which models likelihoods (pixel-wise) and label dependencies through an objective learning method denoted as Multi-Objective through GraphCut (MO-GC). The framework proposed in this method uses a single deep convolutional network. Two non-structured loss functions were used: first one encodes the label likelihoods and second one encodes label dependencies. To the best of our knowledge, this is latest proposed method providing face labeling till date.

Differently from all the mentioned approaches, our previously proposed approach MFS is a new method for face segmentation which extended the label set into six semantic-classes. A data-set of 70 manually labelled images was built and made publicly available. A new model was trained using the extracted features. The best possible configuration was investigated by changing various parameters and spatial setting in those experiments. We observed during experiments three major problems faced with MFS. Firstly, we did not include any kind of conditional hierarchy or global modelling of face regions in the framework. In the proposed MSS-CRFs model, we included a conditional hierarchy for six facial regions which boosted performance of the whole framework. Secondly, processing time of the MFS is very long due to providing labels for each pixel individually. MSS-CRFs is using the idea of super-pixels which reduces the processing time of a testing image. Lastly, the testing set of the MFS is only 70 images (high resolution frontal images), out of which 20 were used for training and 50 for testing. Along with MFS comparison, we also performed experiments on three other data-sets FASSEG V- 4, FIGARO [20] and LFW [3]. FASSEG V- 4 consists of low resolution frontal images taken from Pointing'04 [21] and SliblingDB [22] databases with image dataset of 182 images.

3. Proposed Face Segmentation Model

Our previously proposed face segmentation method (MSF) divides a given image into patches with a fixed step size. After patches creation, features are extracted from each patch. Using the extracted features a Random Decision Classifier is trained and tested. This method does not consider any conditional hierarchies such as location of various face parts and their relationship with each other. For example it is very unlikely to happen mouth region near eye region. Unlike MSF, we formulate a CRFs model which couples labels of face parts in a scale hierarchy. Another serious problem with MSF is speed, since providing a class label to each pixel within an image takes a long time. Instead of providing class labels to each pixel individually, a given image is first divided into super-pixels. All pixels within the super-pixel get the same class label and as a result reducing processing time of the framework.

The presentation of the proposed algorithm is divided into two parts: feature extraction is presented in subsection 3.1 and segmentation via CRF and energy optimization is explained in subsection 3.2.

3.1. Feature Extraction Methods

Super-pixel algorithm over-segments an image by grouping pixels into small meaningful patches that belong to the same object. Instead of using just pixels, many image processing applications benefit from working with super-pixels. The number of entities to be labeled in semantic segmentation are reduced immensely by super-pixels. Each super-pixel has multiple visual features. A single image is represented by a multiple visual feature space after segmentation. We use SEEDS [23] algorithm to over-segment an image into super-pixel. SEEDS is faster than previously proposed super-pixel segmentation methods [23]. Moreover, according to standard error metrics, quality of super-pixel segmentation for SEEDS is also higher than SLIC and other methods [23]. Main problem of our previously proposed MFS method is speed; hence SEEDS is the best choice in our experiments.

For the number of super-pixels we did a large number of experiments. During these experiments we noted better results with a super-pixels number 700. The actual number of super-pixels is of course smaller than this number due to certain restrictions. The actual number of super-pixels depends on the image size and number of block levels used in the super-pixel extraction process. The number of block levels defines the blocks which the algorithm is using in the optimization process. If the numbers of levels are increased, the super-pixel segmentation is more accurate; but this results more memory and time consumption by the CPU. The SEEDS parameters we set are; number of block levels = 3, histogram bins = 5 and each block level is iterated twice for better accuracy.

For node features we use three different features extraction methods: color, shape and spatial information. Different parameterization and settings for features are explored to find the best possible configuration. We investigated these parameters in our previous work MFS.

For spatial information, relative location of the center pixel of each patch is used as a feature. Relative location of a pixel at position (x, y) is defined as $f_{loc} = [x/W, y/H] \in R^2$; where W is the width and H is height of the image.

HSV color histogram is adapted as color features. All the three values in HSV (hue, saturation and variance) are concatenated to form a single feature vector. Patch dimension of 16×16 (D_{HSV} = 16×16) is used with 32 bins (N_{bins} = 32). Using these values each patch generated a feature vector $F_{HSV} \in R^{96}$ for color information.

To account for shape features, widely used HOG [24] is utilized. Dimension of the patch for extracting HOG is kept 64×88 (D_{HOG} = 64×88). With this dimension, a feature vector $f_{HOG} \in R^{2520}$ was produced.

Spatial, color, and shape feature vectors were concatenated to form a single feature vector $f \in R^{2618}$.

3.2. Multi-class Segmentation via CRFs and Energy Optimization

To estimate face segments we use CRFs. The proposed CRFs model encodes the probability of segmentation S with image features Z. The segmentation S is represented by S = { s_1 , ..., s_m}. Where m is the total number of super-pixels in the image. s_i can take one of the six values corresponding to 'mouth', 'eyes', 'back-ground', 'nose', 'hair' and 'skin'. Z consists of node features Z^m and edge features Z^e . We compute F_m features for i_th super-pixel, so Z_i^m is a vector having length F_m. For pair of neighbouring super-pixels i, j we compute F_e features resulting a single vector $Z_{i,j}^e$ having length F_e.

Now, the log linear CRFs model developed will have node energies $\psi(s_i,Z_i^m)$ and edge energies $\psi(s_i, s_j ,Z_{i,j}^m)$. Both of these quantities can be represented as follows,

$$\psi(s_i = l, Z_i^m) = \sum_{f=1}^{F_m} (X_l^m)_f \, (Z_i^m)_f$$

$$\psi\left(s_i = l_1, s_j = l_2, Z_{i,j}^e\right) = \sum_{f=1}^{F_e} \left(X_{l_1,l_2}^e\right)_f \left(Z_{i,j}^e\right)_f$$

Where a set of node weights is represented by X^m and edge weights X^e for each label l and pair of labels (l_1, l_2) respectively.

Now probability of the segmentation S if Z is given will be

$$p(S \setminus Z) = \frac{exp\left(-\sum_{i=1}^{m} \psi(s_i, Z_i^m) - \sum_{i,j} \psi\left(s_i, s_j, Z_{i,j}^m\right)\right)}{N(Z)}$$

The second sum in the above equation is for neighbouring super-pixels and N(Z) is the partition function which is used to normalize the distribution.

For the partition function we use the log likelihood through Bethe approximation [25]. Similarly for marginal approximation of each s_i we use loopy belief propagation. We add Gaussian prior for regularization of weights. For estimating segmentation, we utilized loopy belief propagation in order to find maximum posterior marginals. To evaluate labelling accuracy of segmentation estimates, we applied L1 error on each segmentation estimate. By this way each super-pixel is penalized according to the difference between probability of correct label and probability value 1.0. For example if the estimated super-pixel is given a probability of 0.7 being skin, which is in fact skin, a penalty of 0.3 would be incurred as a result.

4. Experimental Results and Discussion

The only dataset available for six classes is FASSEG [14]. FASSEG is available in four different versions. It can be downloaded from the website: http://khalilkhan.net/face-segmentation-dataset/. FASSEG V- 2 contains high resolution frontal images with low level of variability. FASSEG V- 4 contains images which are low resolution and also there is a variability factor such as candidates with beard, moustaches, glasses etc. We performed our experiments with FASSEG V-2 and FASSEG V- 4. The promising results show that the proposed model is capable of segmenting facial parts successfully from facial images.

Some of the images segmented with proposed MSS-CRFs model are shown in Figure 1. Images shown in Figure 1 are segmented into its corresponding face parts efficiently. However in some cases segmentation results of the proposed MSS-CRFs algorithm are comparatively poor. Figure 2 show images from the database with poor results. Testing image shown in row 1 is the case where face passed to the framework is not compatible with training data images. Facial parts nose, eyes and eyebrows are more concentrated to the upper part of the image. As a result segmentation results of the eyes and eyebrows are very poor. If a testing image has glasses, there is problem in segmentation with nose, eyes and eyebrows specifically (testing image in row 2). Similarly if a testing image has beard or moustaches, there is also segmentation problem (testing image in row 3). The proposed framework is unable to segment face parts such as moustaches and beard.

In the following paragraphs we conclude results obtained during experiments while using FASSEG V-2, FASSEG V- 4, FIGARO and LFW databases.

4.1. Face Segmentation V-2

FASSEGV-2 contains 70 images. We used this version of the database in our previous method MFS. Figure 3 shows comparison of the MFS and proposed MSS-CRFs results. From the Figure 3 it is clear that there is improvement of pixel labelling accuracy (PLA) for all classes with proposed method.

Performance of the MFS is not poor on the majority classes (hair, background and skin); however results for minority classes (eyes, nose and mouth) were not satisfactory. Our main target in the present work was improving PLA of the rare and difficult classes. The most advantageous classes in MSS-CRFs are eyes, nose and mouth regions. PLA of the minor classes increased in the present work with two valid reasons; firstly, manual labeling was not performed properly in MFS. Class labels were particularly not given properly to these rare classes due to their complex shapes. Due to their less area in the whole face image; training data for these classes was not provided properly. Here we performed manual labeling with extreme care using manual labeling software Photoshop. Particularly for nose, just tap of the nose was labeled previously in MFS. Label for nose followed different convention here i.e; extending the nose label to the mid of two eyes. As a result PLA of the nose jumped from 29.83% to 68.97% (as shown in Figure 3). Secondly, as MFS is not considering any conditional hierarchy about various face parts; previously these minor classes were mostly miss-classified with majority classes.

Moreover, processing time for single image segmentation is reduced with proposed method. A substantial increase in speed - of an order of magnitude - is obtained by using super-pixels, since the number of patches to be classified by the model is greatly reduced. In case of MFS, a class label is provided for each pixel individually while MSS-CRFs assign class label to super-pixels only. All pixels within the super-pixels then get the same class label. We used a single CPU (2.8 GHz , Core i7 and RAM 8 GB). We do not use any GPU or some dedicated hardware for our experiments. A single image having size 520×480 is divided into super-pixels in 1.51 seconds with SEEDS.

The total framework runs in 2 seconds in the proposed approach which was 49 seconds in the previously proposed MFS method.

4.2. Face Segmentation V-4

Along with FASSEG V-2 images, we added 182 more frontal images to the database. These images were taken from Pointing'04 [21] and SliblingDB [22] databases. Size of the images was kept the same as in MFS (constant height H = 512 and width W varied accordingly to keep ratio of original image). Out of the total images, 20 were taken randomly and used for training. Remaining 152 images were used for testing. Figure 4 shows confusion matrix for the results obtained for every class. From the Figure 4 it is clear that obtained PLA for all the classes except nose is really impressive.

4.3. FIGARO and LFW-PL Databases

Along with FASSEG database, we also conducted experiments with two other databases FIGARO [20] and LFW [3]. For fair comparison we kept the same setting and same set of images in testing phases as in Svanera et al. [20] and Sifei et al. [3]. However for training phase we used images as in the experiments conducted in first phase.

Authors of FIGARO [20] used only hair class during their experiments. FIGARO is a comparatively small database with 840 images in total. All these images are collected from web pages. Authors of the paper included different kind of variations in hair styles; reporting seven hair classes (straight, curly, wavy, kinky, short-men, braids, dreadlocks).

LFW is a big database with a large variety of images. All images in the LFW database are captured in the unconstrained environment where a large number of variations are present due to various environmental factors. For LFW database we conducted experiments with three classes (hair, skin and background) as in Sifei et al. [3]. Reported accuracy for this case is at pixel-level for all the three classes.

Figure 5 shows comparison of the proposed method with FIGARO and LFW database results. From the Figure it is clear that we have better results on FIGARO database. However, our reported results on LFW are lower as compare to previously reported results. All training images in the FASSEG are captured in a controlled lab environment, while testing images in LFW are in unconstrained condition. We believe that if such variations are included in the training data, we can obtain better results as compared to state of the art results on LFW database as well.

Main advantage of the proposed method is providing class labels for complete face. Unlike state of the art methods which were considering only few classes; MSS-CRFs provides segmentation of all face parts. Hair segmentation is reported as a difficult task comparatively in previous literature [26, 27]. Previously reported methods were not able to segment hair properly due to its complex geometry and larger variability varying from person to person. However, reported results for hair are encouraging and confirm effectiveness of the proposed method. Reported results also show that the proposed method is robust to lighting variations as some of the images used in testing phase are captured in uncontrolled lighting conditions. Our proposed method provided class labels for all six face parts. In some applications class label is needed for specific part only. In that case the algorithm proposed can be used according to the need and application.

The proposed method has some minor drawbacks as well. While creating the database the labeling is performed manually by a human. Providing a class label in the transition region between two classes is very uncertain in such conditions. Similarly patch sampling for training phase is based on random criterion. However the number of pixels from rare classes is insufficient for training. We believe that this results in poor performance as compare to majority classes; having sufficient training data. Also, we noted the proposed framework is un-suitable in cases having beard and moustaches in face images. Providing a separate class label for each of these parts may solve the problem.

5. Conclusion

Semantic segmentation of faces using CRFs is introduced in this paper. We combine features of position, HSV color and shape information to build a CRFs model. A great deal of information is provided about the face parts skin, hair, nose, eyes, mouth and background by a CRFs estimated model. Experimental results show that the proposed model not only outperform state of the art results on FASSEG and FIGARO databases, but also achieves improvement in the previous results by a big margin.

The future work can be extended in two directions. Firstly, improving the current model to get better pixel-labelling accuracy. A higher level of variability can be added to training and testing data to make the framework suitable for unconstrained conditions. Secondly, applying the current segmentation model to certain mid-level vision features estimation. We believe that immense sources of information are provided for many hidden variables such as pose, gender, expression, ethnicity, age, beardedness, balding in human and so on.

References

[1] Ferrara M, Franco A, Maio D. A multi-classifier approach to face image segmentation for travel documents. Exper Syst Appl 2012; 9: 8452-8466.

[2] Liu C, Yuen J, Torralba A. Nonparametric scene parsing via label transfer. IEEE T Pattern Anal 2011; 33: 2368-2382.

[3] Liu S, Yang J, Huang C, Yang MH. Multi-objective convolutional learning for face labeling. In: Proc CVPR IEEE; June 8-10, 2015; Boston, Massachusetts, USA: IEEE. pp. 3451-3459.

[4] Everingham M, Van GL, Williams CK, Winn J, Zisserman A. The Pascal visual object classes challenge. Int J Comput Vision 2010; 2: 303-338.

[5] Huang, Gary B, Manjunath N, Erik L. Towards unconstrained face recognition. In: Proc CVPR IEEE; 24-26 June 2008; Alaska, USA: IEEE. pp. 1-8.

[6] Zebrowitz, Leslie A, Joann M. Social psychological face perception: Why appearance matters. Social and Personality Psychology Compass 2008; 3: 1497-1517.

[7] Shepherd J, Ellis H, Davies G. Perceiving and remembering faces. 1st ed., Chicago USA: Academic Press, 1981.

[8] Sinha P, Balas B, Ostrovsky Y, Russell R. Face recognition by humans: nineteen results all computer vision researchers should know about. In: Proc of the IEEE 2006; 11: 1948-1962.

[9] Moon JT, Marschner SR. Simulating multiple scattering in hair using a photon mapping approach. ACM T Graphic 2006; 25: 1067-1074.

[10] Ward K, Bertails F, Kim TY, Marschner SR, Cani MP, Lin MC. A survey on hair modelling: styling, simulation, and rendering. IEEE T Vis Comput Gr 2007; 2: 213-234.

[11] Lee C, Schramm MT, Boutin M, Allebach JP. An algorithm for automatic skin smoothing in digital portraits. In: IEEE Image Proc; 07-10 September 2009; Cairo Egypt: IEEE. pp. 3113-3116.

[12] Chen CW, Huang DY, Fuh CS. Automatic skin color beautification. In: International Conference on Arts and Technology; 10-12 December 2010; Berlin Heidelberg: IEEE. pp. 157-164.

[13] Xu L, Yangzhou D, Yimin Z. An automatic framework for example-based virtual make-up. In: IEEE Image Proc; 15-18 September 2013; Melbourne, Australia: IEEE. pp. 3206-3210.

[14] Khalil K, Massimo M, Riccardo L. Multi-class semantic segmentation of faces. In: IEEE Image Proc; 27-30 September 2015; Québec, Canada: IEEE. pp. 827-831.

[15] Lafferty J, McCallum A, Pereira F. Conditional random fields: probabilistic models for segmenting and labelling sequence data. In: International Conference on Machine Learning; June 28 – July 01 2001; Williamstown, MA, USA: pp. 282-289.

[16] Kae A, Sohn K, Lee H, Learned ME. Augmenting CRFs with Boltzmann machine shape priors for image labelling. In: Proc CVPR IEEE; 25-27 June 2013; Oregon convention centre, Portland: IEEE. pp. 2019-2026.

[17] Eslami SA, Heess N, Williams CK, Winn J. The shape Boltzmann machine: a strong model of object shape. Int J Comput Vision 2014; 2: 155-176.

[18] Li Y, Wang S, Ding X. Person-independent head pose estimation based on random forest regression. In: IEEE Image Proc; 26-29 September 2010; Hong Kong, China: IEEE. pp. 1521-1524.

[19] Scheffler C, Odobez JM. Joint adaptive colour modelling and skin, hair and clothing segmentation using coherent probabilistic index maps. In: British Machine Vision Conference; 29 August – 02 September 2011; University of Dundee UK: BMVA Press.

[20] Michele S, Umar RM, Riccardo L, Sergio B. FIGARO, hair detection and segmentation in the wild. In: IEEE Image Proc; 25-28 September; Phoenix, USA: IEEE. pp. 546-550.

[21] Gourier N, Hall D, Crowley JL. Estimating face orientation from robust detection of salient facial features. In: International Workshop on Visual Observation of Deictic Gestures; 22nd August 2004; Cambridge, UK.

[22] Laurentini A, De SM, Bottino AG, Vieira TF. A new problem in face image analysis: Finding kinship clues for siblings pairs. In: International Conference on Pattern Recognition Application and Methods; 6-8 February 2012; Vilamoura, Algarve, Portugal: Springer. pp. 153-162

[23] Van DB, Michael, Xavier B, Gemma R, Benjamin DC, Luc VG. Seeds: Super-pixels extracted via energy-driven sampling. European Conference on Computer Vision; 7-13 October 2012; Firenze, Italy: Springer. pp. 13-26.

[24] Dalal N, Triggs B. Histograms of oriented gradients for human detection. In: Proc CVPR IEEE; 20-26 June 2005; San Diego, CA, USA: IEEE. pp. 886-893.

[25] Zhao W, Chellappa R, Phillips PJ, Rosenfeld A. Face recognition: A literature survey. ACM Computing Surveys 2003; 4:399-458.

[26] Wang N, Haizhou A, Shihong L. A compositional exemplar-based model for hair segmentation. In: Asian Conference on Computer Vision; 8-12 November 2010; New Zealand: Springer. pp. 171-184.

[27]Lee KC, Anguelov D, Sumengen B, Gokturk SB. Markov random field models for hair and face segmentation. In: International Conference on Automatic Face & Gesture Recognition; 17-19 Sep 2008; Rode Hoed Amsterdam, Netherlands: IEEE. pp. 1-6.

Figure 1: Images from FASSEG V-4 database. First column show original RGB images, second column show ground truth images and third column show results obtained with proposed MSS-CRFs (better segmentation results).

Figure 2: Images from FASSEG V-4 database. First column show original RGB images, second column show ground truth images and third column show results obtained with proposed MSS-CRFs (poor segmentation results).

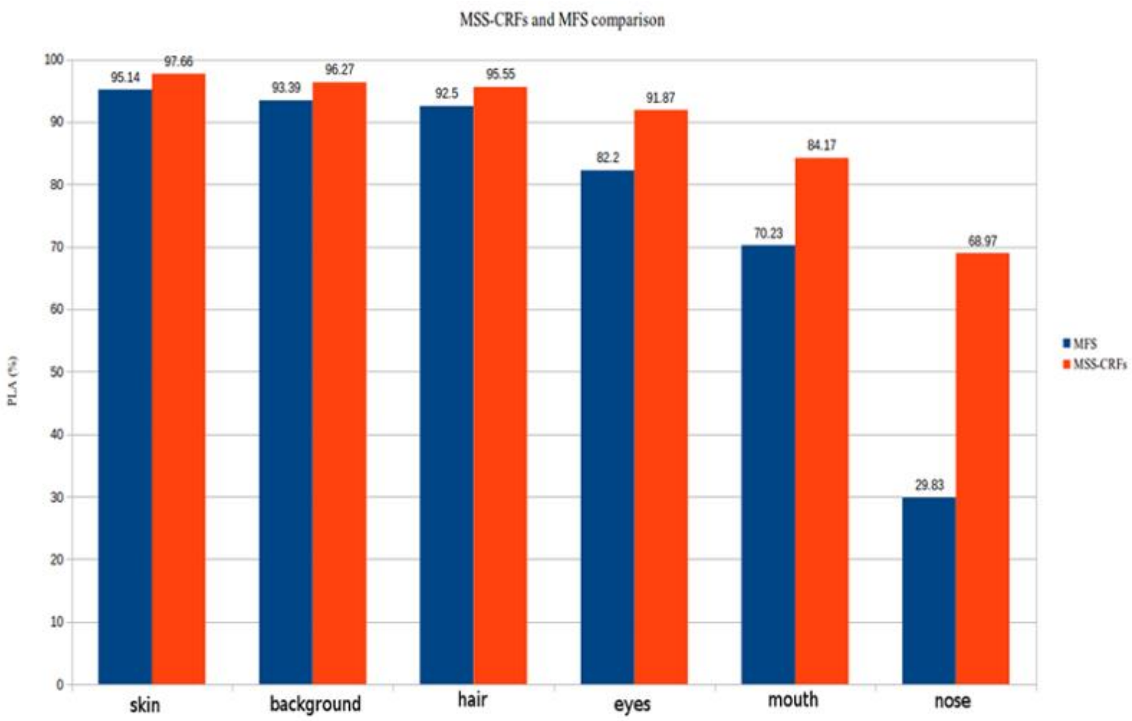

Figure 3: Proposed MSS-CRFs and MFS results comparison using FASSEG database V-2.

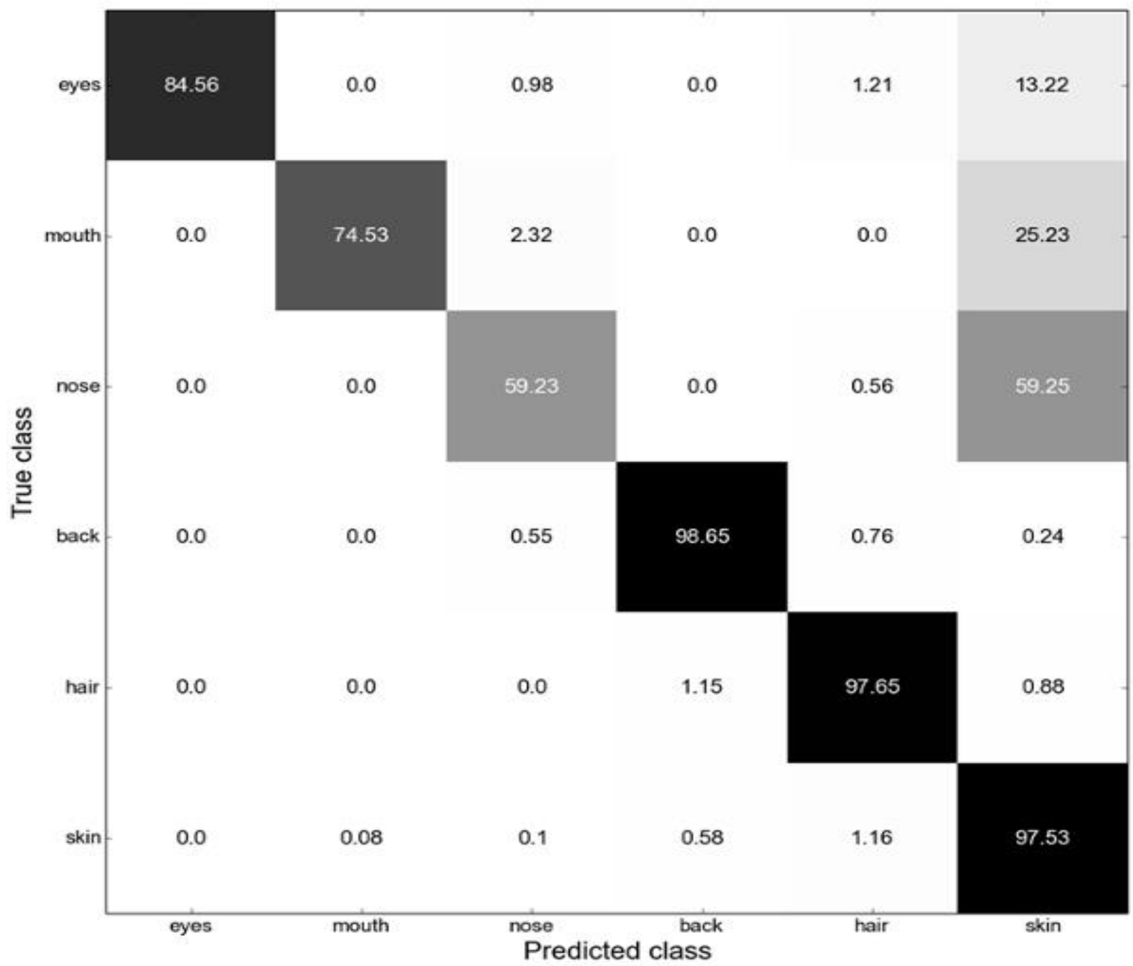

Figure 4: Confusion matrix obtained for all six classes using MSS-CRFs method and FASSEG V- 4 database.

Database	Method used	Accuracy (Percentage)
FIGARO	**Proposed approach**	**94.56**
FIGARO	FIGARO	86.20
LFW-PL	**Proposed approach**	**92.47**
LFW-PL	MO-GC with prior	95.12
LFW-PL	MO-GC	95.24

Figure 5: MSS-CRFs comparison with previously reported methods.