2020/2021 CVHCI Competition

By
Julian Hohenöcker (upqbf)
Jonas Steinhäuser (uoxoa)
Giulio Broghammer (uorvl)

Judge User: shadow

A1 Skin Detection

1st Approach

- **Bayes Classifier** with Gaussian Density Model in RGB space 1)
  - Use each skin/non skin pixel to set up mean and covariance matrix for skin/non skin model
  - Classifies pixel as skin if p(x|skin) >= ϑ * p(x|nonskin)
- Pre-processing
  - Histogram Normalization
- Post-processing
  - Perceptual Grouping
  - First Morphological Closing then Morphological Opening

1) Phung, Son Lam, Abdesselam Bouzerdoum, and Douglas Chai. "Skin segmentation using color pixel classification: analysis and comparison." IEEE transactions on pattern analysis and machine intelligence 27.1 (2005): 148-154.

2nd Approach

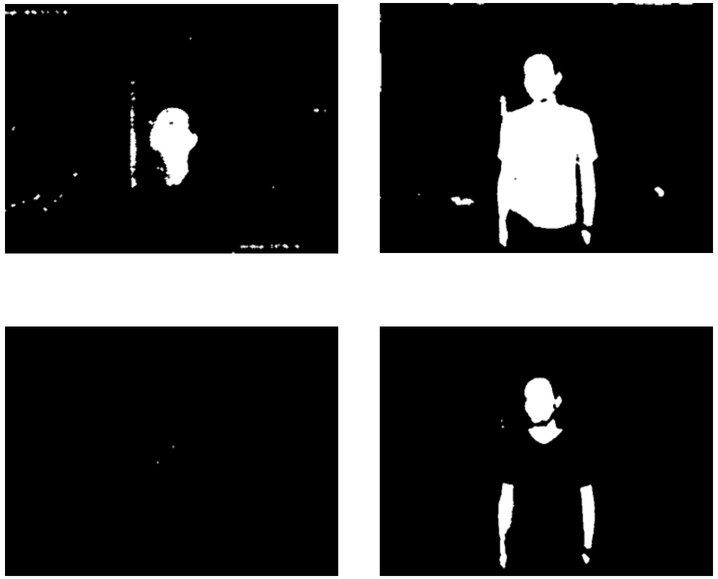
- **MLP** in RGB space [1)]
  - Input nodes for R, G and B value
  - One hidden layer since skin space not linearly separable
    - Best results with hidden layer of 25 nodes
  - Output nodes for skin and not skin rating
  - Classifies pixel as skin if R(skin) – R(non skin) >= ϑ
- Pre-processing
  - Better Results without (e.g. RGB Normalization)
- Post-processing
  - Perceptual Grouping

1) Bhoyar, K. K., and O. G. Kakde. "Skin color detection model using neural networks and its performance evaluation." Journal of computer science. 2010.

Comparison and results

- Bayes Classifier
  - Even **acceptable** results on **bright lighting** conditions
- MLP
  - **Good** results on images with **average conditions**
  - **Better** classification **for skin** like coloured objects
  - Could be improved with more training on data with different lighting conditions

Bright Lightning Normal Lightning

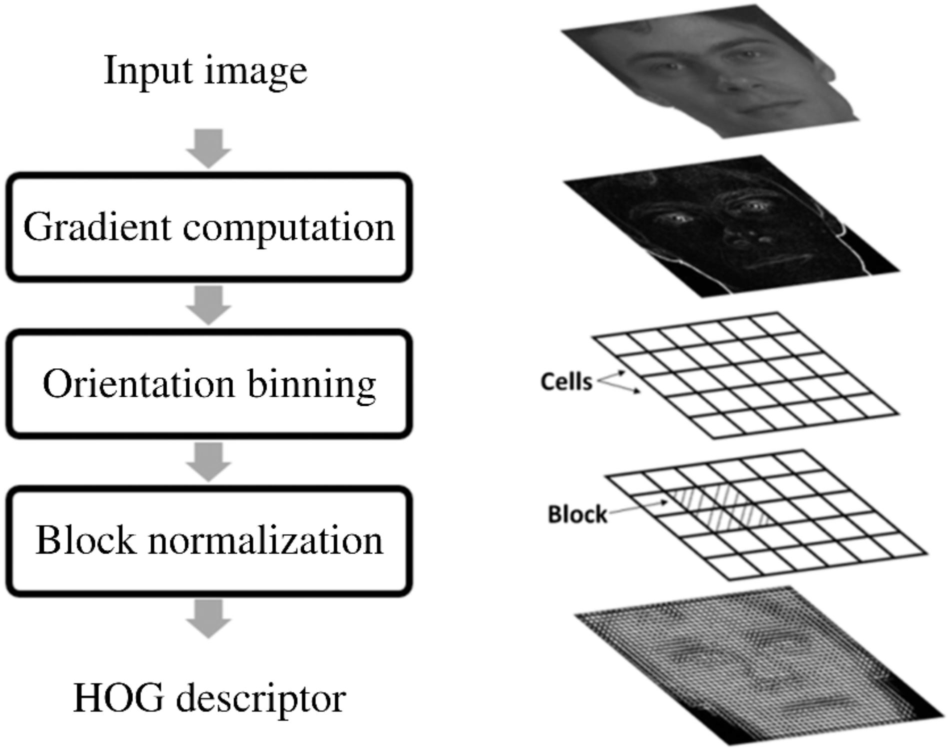


Scores	Validation Set	Test Set
Bayes Classifier	82.1635 %	64.3274 %
MLP	68.838 %	**80.753 %**

A2 Person Detector

Histogram of Oriented Gradients

- HOG Settings:
  - Image Size: 64x128 px
  - Block Size: 16x16 px
  - Cell Size: 8x8 px
  - Number of Bins: 9
  - Normalization: L2-hys
- These values worked best
- Also shown by Dalal & Triggs, 2015 1)



1) Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). Vol. 1. IEEE, 2005.

Support Vector Machine

- SVM training from scratch:
  - Bad Accuracy
- Pretrained SVMs (OpenCV)
  - Default People Detector (for 64x128px images) => our task
  - Daimler People Detector (for 48x96px images)
- Object detection > Image classification
  - OpenCV detectors create Bounding Boxes
    - We used Sliding Windows with Stride 2
  - Bounding Boxes are neglected

    => only yes or no important

HOG + SVM

- Results:
  - Accuracy of 96.97 %
- No False Positives
- Only False Negatives:

Big People
but somewhat obstructed

&

Small People
and apparently too small

A3 Face Similarity

1^{st} Approach

- Eigen Faces + LBPH
  - Pre-processing: Less Background via Cropping
  - Compute Eigen Faces of Train-Images
  - Create Local Binary Pattern Histogram
    - Add Eigen Faces pairwise
    - Distance Measure: L2 Norm
  - Project Test-Images in Eigen-Face Space
  - Use L2 Norm as Similarity Measure
- Result
  - Accuracy of 65.1899%

2nd Approach

- Dlib DNN-Face-Recognition
  - Detects multiple Faces in one image
  - Creates feature representation for each found face
- Deep Metric Learning with ResNet [1)]
  - Use modified ResNet-34 (29 conv layers)
  - Create 128-dim Feature Vector
  - Trained via Metric Loss
- Similarity via Euclidean Distance of Feature Vectors
  - Threshold of 0.6

1) He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

2^{nd} Approach

- Problem:
  - No face found => No Feature Vector created
  - (10 times in 1000+ images)
- Possible solutions:
  - Return fixed value 0 or random value (bad)
  - Use Alternative approach (e.g.: 1^{st} Approach)
- Result:
  - Accuracy of 97.3732%