APPLIED MACHINE LEARNING SYSTEM ELEC0132 19/20 REPORT

SN: 19180271

uceevai@ucl.ac.uk

ABSTRACT

This report summarises the methods, implementation and performance of four Machine Learning tasks - gender recognition, emotion recognition, face shape classification and eye colour classification. The tasks were solved separately using a combination of face detection methods, feature extraction and classification. The key was to get maximum possible accuracy on validation and test sets, while at the same time optimising performance levels. The following accuracies were obtained on each task respectively: 88%, 88.8%, 82.12% and 95%. The link to the code has been provided at footnote. [1]

1. INTRODUCTION

This section will show an overview of how the tasks were performed and the layout of the rest of the report.

Task A1 (Gender Classification) used a Viola Jones face detector to detect faces from the images belonging to CelebFaces Attributes Dataset (CelebA), which contains facial images of celebrities. Local Binary Patterns and Histogram of Oriented Gradients were used as features from the inputs. The feature space was scaled down using Principal Component Analysis and further fed to a SUpport Vector Machine (SVM) classifier using Radial Basis Function (RBF) kernel. The parameters were optimised and gender classification was performed.

Task A2 (Emotion Classification) uses the same dataset as task A1. It used a dlib based facial detector to get landmarks around the mouth region, as we are specifically interested in smile. An additional feature was created by taking the ratio of the height of the mouth against the width of the mouth called the Mouth Aspect Ratio. These features were scaled and used as input to a SVM with RBF kernel for classification.

Task B1 (Face shape classification) uses Cartoon Set as dataset, which contains cartoon avatars. Facial landmarks were extracted using dlib detector. The jaw landmarks were extracted as features. Density Based Spatial Clustering of Applications with Noise (DBSCAN) was used to detect and remove outliers from the dataset as several of the images contained beards. A SVM with RBF kernel was used to classify the images based on face shape.

Task B2 (Eye Colour Recognition) uses the same Cartoon dataset as task B1. Once again facial landmarks using dlib detector was extracted. The eye region alone was taken from the images based on the landmarks. The images were converted to Hue-Saturation-Value (HSV) space and the average hue value was taken as feature. Outilers were removed using InterQuartile Range method. The features were used to train an SVM-RBF classifier and the images were classified based on eye colour.

The report is organised as follows: Section 2 is the Literature Survey, where a study of academic papers related to the above tasks are presented briefly. The methodologies used to implement the above tasks are inspired by them. Section 3 provides a detailed description of the models used to perform each of the above tasks, along with some theoretical foundation and the rationale behind using the above models. Section 4 deals with the practical implementation of the models. It gives details regarding the functions and the libraries used, along with details about the datasets and feature space. Section 5 presents the final results and an analysis of the results. Section 6 is the Conclusion, providing a review of the results as well as room for improvement. Section 7 provides the References used for this report.

2. LITERATURE SURVEY

2.1. Gender Classification

Choosing the right features from facial images hugely determines the success of a gender detector classifier.Feature extraction methods for gender classification can be broadly categorized into geometric and appearance based approaches [1][2]. The former is based on measurements of facial landmarks. Appearance-based methods are based on some operation or transformation performed on the pixels of an image [3].Pixel intensity values can be used directly as input to train a classifier, as has been done as early as 1991 by Golomb et al.who developed a CNN called SexNet [4] with an error rate of 8.1% only. Another major stride in the field was when Moghaddam and Yang [5] used raw image pixels with nonlinear SVMs for gender classification on thumbnail faces (12x21 pixels); their experiments on the FERET database (1,755 faces) demonstrated SVMs are superior to other classifiers, achieving the accuracy of 96.6%.They were able to

[1] https://github.com/ash95sv/AMLS-1, https://drive.google.com/drive/folders/1sCeW7XhcxMgIEbvYJvFWrYOnGXydMaY3?usp=sharing

show that SVMs (with either Polynomial or Radial Basis Function (RBF) kernel) were able to out perform other more traditional classifiers, such as Linear Discriminant (LD), Quadratic Discriminant (QD), Fisher Linear Discriminant (FLD) and Nearest-Neighbours (N-N) and that the SVM system performed similarly for both high and low resolution images. This then demonstrated the robustness of system for gender classification.

Ojala et al. [6] introduced Local Binary Patterns (LBP) for greyscale and rotation invariant texture classification.Hadid et al.[7] achieved 96.3% accuracy on motion data using an appearance based approach with Local Binary Pattern (LBP) features and Support Vector Machines (SVMs).Caifeng Shan [8] adapted LBP by using Adaboost to select the best set of LBP features.Together with SVM-RBF, they were able to achieve a classification accuracy of 94.81% in uncontrolled conditions.

More recently, Annalakshmi et al.[9] used spatially enhanced local binary pattern (SLBP) and histogram of oriented gradients (HOG) with SVM classifier. This hybrid feature selection representation of texture micro-patterns and local shapes has resulted in classification accuracy of 99.1% on FERET database and 95.7% on LFW database. In [10] as well, Jiann-Der Lee et al. used LBP and HOG features on FERET database, and applied t-test to reduce feature space by 70%. SVM classifier was used to obtain an accuracy of 92.2%.

2.2. TaskA2:Emotion Classification

For automatic FER(Facial Emotion Recognition) systems, various types of conventional approaches have been studied. The commonality of these approaches is detecting the face region and extracting geometric features, appearance features, or a hybrid of geometric and appearance features on the target face [10]. For the geometric features, the relationship between facial components is used to construct a feature vector for training . Ghimire and Lee [11] used two types of geometric features based on the position and angle of 52 facial landmark points. The results on the Cohn-Kanade (CK+) facial expression database show a recognition accuracy of 95.17% and 97.35% using multi-class AdaBoost and SVM, respectively.In appearance features, Happy et al. [12] utilized a local binary pattern (LBP) histogram of different block sizes from a global face region as the feature vectors, and classified various facial expressions using a principal component analysis (PCA). The testing results indicate that by using LBP features facial expressions recognition accuracy is more than 97%.

Ghimire et al. [13] testify that SVM is suitable for recognizing facial expressions from single frame as there is no direct probability estimation in SVMs.

Binh T. Nguyen et al. [18] developed an approach to detect human emotion in real time. Their experiments resulted in a classification accuracy of 70.65% when using a range of features from the facial landmarks along with SVM-RBF. They compared the results of the SVM-RBF with a Decision Tree classifier, Random Forest classifier and a Linear-SVM. The SVM-RBF system outperformed them all.

2.3. TaskB1: Face Shape Classification

Adonis Emmanuel [14] used CNN by retraining the last layer of the Inception v3 model in classifying images of human faces into one of five basic face shapes. The accuracy of the retrained Inception v3 model was compared with that of the following classification methods that uses facial landmark distance ratios and angles as features: linear discriminant analysis (LDA), support vector machines with linear kernel (SVM-LIN), support vector machines with radial basis function kernel (SVM-RBF), artificial neural networks or multilayer perceptron (MLP), and k-nearest neighbors (KNN). 19 features were extracted per image to train the non-CNN based classifiers. The 19 features include the following: face height to width ratio, jawline width to facewidth ratio, chin-to-mouth to jaw line width ratio, and the angles that the line from the chin to each one of 16 facial points make with respect to the vertical. Whereas the CNN showed 97.8% accuracy, SVM displayed 50-55% accuracy with the above features.

Kitsuchart Pasupa et al. [15] presented a novel system for a hairstyle recommendation system that is based on a face shape classifier. They used a combination of features, geometric features obtained from facial landmarks and deep-learned features using VGG-face. Multiple Kernel Learning (MKL) was used to combine the feature sets together. They were able to achieve a classification accuracy of upto 70.33% using a SVM-RBF.

More recently, Kitsuchart Pasupa et al. [16] presented another approach in effective face shape classification for constructing a hairstyle recommendation system, based on an Active Appearance Model (AAM) and a face segmentation-technique, produced a set of features that were evaluated by Linear Discriminant Analysis (LDA), Artificial Neural Networks (ANN), and Support Vector Machine (SVM Linear and RBF). They results proved that the Support Vector Machine with Radial Basis function kernel was the best algorithm that predicted accurately up to 72%.

Luning Li et al. [17] proposed an Active Appearance Model (AAM) based method to remove face shape noise for the purpose of improving facial expression recognition performance. The feature set, which is made up of facial geometry information, is extracted using AAM. Subsequently, based on the extracted features, an SVM classifier is used to classify the face shape. They were able to achieve an average face shape classification accuracy of 89.4%.

2.4. TaskB2: Eye Colour Classification

In [18], Zezhi Chen et al. performed vehicle colour classification using an 8-bin 3D colour histogram as the vector for Polynomial-SVM classification. They were able to achieve an average classification accuracy of 97.4%.

Nakhoon Baek et al. [19] performed Vehicle Color Classification Based on the Support Vector Machine Method. They converted the vehicle image into the HSV(hue-saturation-value) space, to eliminate distortions due to the intensity changes. They constructed a histogram for the hue and saturation pairs and used the SVM(support vector machine) method to classify the feature vectors into five vehicle colors, resulting in an accuracy of 94.92%.

Cheng-Jin Du et al. [20] developed an automatic pizza sauce spread classification system using colour vision and SVMs. After segmenting the image from the background, the image was converted from the Red-Green-Blue (RGB) colour space to the Hue-Saturation-Value (HSV) colour space. Using Principle Component Analysis (PCA), they reduced the HSV feature space. They were able to achieve a classification accuracy of 96.67% using the reduced feature space and a Polynomial-SVM.

3. DESCRIPTION OF MODELS

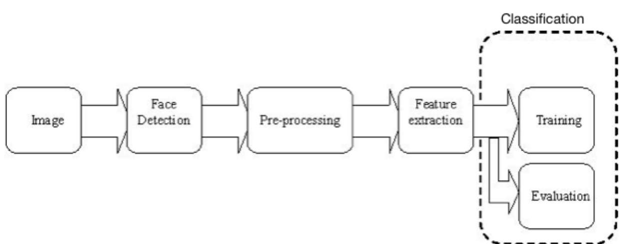

Fig. 1. General model for all classifiers

3.1. Task A1: Gender Classification

The face detector used is based on the work done by Viola and Jones in [21].The algorithm has four stages: Haar Feature Selection, Creating an Integral Image, Adaboost Training and Cascading Classifiers. Haar Like features are sub-windows of the image where one part is darker than the other. To increase speed, an Integral image is created where the sum of the pixel values is taken across the rows and columns so that pixel (i,j) is the sum of all pixels before it. Having this calculation preprocessed on the image allows the computer to calculate rectangular features with a few addition and subtraction operations instead of looping across the pixel values.AdaBoost uses all of the rectangular features and trains them against the training data set. During this process it will find and make groups of the weak features into classifiers by taking a linear combination of the weak classifiers.To classify if a sub-window of an image is actually a face, an Attentional Cascade is used. It is essentially a decision tree where each node is one of the constructed classifiers. The Viola-Jones algorithm is very robust as it offers very high true positive rates and very low false positives. It also offers high speed detection, making it suitable for real time applications. [22] After the detections were made, the region of interest (ROI) is cropped out and saved. Histogram equalization [23] was performed on the images to enhance contrast. It is a method that modifies the intensity distribution of the image histogram by distributing the frequencies of intensity evenly across pixel values.

Inspired by [9] and [10], this task uses a gender detector based on LBP and HOG features with an SVM classifier. The idea is to obtain a feature representation of texture micro-patterns (LBP) as well local gradient variations (HOG). Local Binary Pattern (LBP) is a simple yet very efficient texture operator which labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. It is robust to monotonic gray-scale changes caused, for example, by illumination variations and is computationally simple. LBP algorithm uses the following steps: Step 1: 3 by 3 pixel: For each pixel in the grayscale image, we select a neighborhood of size r, say three, surrounding the center pixel. Step 2: Binary operation: For each pixel's three by three neighbour, compare the center value and its neighbour values. If the neighbour values are greater than center, record 1 else record 0. Step 3: Decimal: Convert the binary operated values to a digit.

HOG descriptors are mainly used to describe the structural shape and appearance of an object in an image, making them excellent descriptors for object classification[24]. However, since HOG captures local intensity gradients and edge directions, it also makes for a good texture descriptor.HOG features were first introduced by Dalal and Triggs [25]. In their work, Dalal and Triggs proposed HOG and a 5-stage descriptor to classify humans in still images. The 5 stages include: Normalizing the image prior to description. Computing gradients in both the x and y directions. Obtaining weighted votes in spatial and orientation cells. Contrast normalizing overlapping spatial cells. Collecting all Histograms of Oriented gradients to form a feature vector. Since the dimensionality of the feature space is huge (more than 10000), Principal Component Analysis is used to reduce the dimensions to less than 200. Principal Component Analysis takes only those dimensions across which most of the features have maximum variance- they are the dimensions that have the highest contribution to the classifier ultimately.

The Literature Survey section testifies the robustness of using SVM as a classifier in facial recognition tasks, including gender recognition.

3.2. Task A2: Emotion Recognition

The second task deals with detecting smiles from facial images from a celebrity database. The literature survey validates the use of geometric features with an SVM classifier, and hence facial landmarks have been used as features. Facial landmarks are a set of notable points, usually located on the corners, tips or mid points of facial features [26], as seen in figure 2.

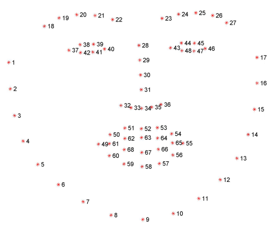

Fig. 2. Facial Landmarks

The classic dlib face detector uses Histogram of Oriented Gradients (HOG) features [25] with a linear classifier. With an image pyramid and the sliding window scheme it covers the entire image while being computationally cheap. King introduces the Max-Margin Object Detection (MMOD) as cost function [27]. Since smile is majorly a function of the mouth region, the landmarks describing the mouth are extracted using dlib face detector. Smile alters the position of these landmarks with respect to the face and hence is ideal to be used as features. Additionally, a value inspired by Eye Aspect Ratio [28] is used to define Mouth Aspect Ratio [29], which is basically a ratio of the height of the mouth to the width of the mouth.

The classifier used here is also SVM with RBF kernel, which is again validated by the robust results obtained by researchers as shown in the literature survey.

3.3. Task B1: Face Shape Detection

The images for tasks B1 and B2 require no pre-processing as they are all cartoon avatars which are all frontal, have perfect alignment and uniform intensities.

Face shape is a geometrical construct, which provides the rationale behind using facial landmarks once again as features. A close study of the cartoon dataset reveals that the difference in face shapes is primarily a function of the position of jaw landmarks. Hence the jaw landmarks were specifically chosen using the dlib face detector (explained in task A2).

The Cartoon dataset had faces with beards, which makes it nearly impossible for even humans to discern the right face shape. Including them in the training dataset could seriously hamper the performance of the classifier and hence needed to be removed. It is for this reason that Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise (DBSCAN) [30] was used to detect these outliers.

The classifier used here is also SVM with RBF kernel, which is again validated by the robust results obtained by researchers as shown in the literature survey.

3.4. Task B2: Eye Colour Recognition

The task pertains only to the eye region, and hence the eye region alone was extracted using DLIB based face detector. The cartoon dataset had 5 different eye colours, with uniform intensities, making it easier to detect than human eye colour. Observing the dataset and studying literature as mentioned in the survey gave the intuition behind using the Hue value as an estimator of eye colour [19][20]. The hue value takes a colour as it is, irrespective of it brightness or intensity. This makes it an ideal feature for colour differentiation.

In this task, the average hue value for one eye was taken for every face.The rationale behind taking average of hue values is that the dataset provides uniformity in colour unlike real human photos. While computing the average, the value corresponding to black pixels were discarded as it is forms the iris in every eye image and does not contribute to real eye colour. By taking the average, only one feature is required to describe the entire eye colour.

Once again, this dataset had images with glasses which makes it difficult to discern the real eye colour. Such images are clearly outliers which needed to be removed. It is for this reason that IQR (Inter Quartile Range) based outlier detection was used for dataset corresponding to each eye colour.

The classifier used here is also SVM with RBF kernel, due to reasons mentioned earlier.

4. IMPLEMENTATION

The following are the major libraries used to implement the models described in the previous section: 1. NumPy: is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high- level mathematical functions to operate on these arrays. 2. Pandas: is a library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series. 3. Scikit-learn: Also known as sklearn, is a machine learning library for the Python programming language. It contains modules capable of various classification, re- gression and clustering algorithms and is designed to work in conjunction with the Python numerical and scientific libraries NumPy and SciPy. 4. Scikit-image: Also known as skimage, is a collection of algorithms for image processing. 5. OpenCV: A library of programming functions mainly aimed at real-time computer vision. 6. Dlib: A general purpose cross-platform library.

4.1. TaskA1:GenderClassification

OpenCV's CascadeClassifier() module is used to detect faces from given images. This implements the Viola Jones face detector based on Haar like features. 5 types of detector objects are created- 4 for frontal face and 1 for profile face. The DetectMultiScale() object of CascadeClassifier module is used to detect face. It takes the greyscale image as the argument along with a scale factor which is the factor by which the image size is reduced at each step of the Viola Jones cascade face detector. The function returns a list of faces which have been detected in the image, along with (x,y) co-ordinates of the associated bounding box, using which the face alone is cropped out of the image. Further, the cropped image is resized to 100x100 in order to not lose resolution. At this stage, 4950 facial images are detected.

Function equalizeHist() from OpenCV was used to perform histogram equalisation on the resized faces. At this point, the list of labels are modified to have only those gender labels corresponding to detected faces.

Function local_binary_pattern() and hog() from skimage library are used to extract LBP values and HOG values respectively. local_binary_pattern() takes the processed image, radius of window and number of surrounding pixels to calculate LBP value as the arguments. In this task the radius is set as 3, which implies 24 surrounding pixels will be used to calculate LBP value of centre pixel. hog() takes the face image and size of blocks to be used for gradient computation. 8x8 sized blocks are used here. The orientation argument (9) gives the number of bins in the histogram, corresponding to angular orientation of gradient. The function returns 4356 HOG values per image. This combined with 100x100 LBP values result in 14356 features per image.

Function PCA() was used to reduce the dimensionality of feature space from 14356 to 165.

These features are split into training and validation sets using train_test_split function from sklearn in a 70-30 ratio. The features are further scaled to (0-1) range using MinMaxScaler() function from sklearn in order to optimise the convergence process.

Using Scikitlearns' SVC() function from svm module, the classification model described using a RBF kernel is implemented. Hyper parameters such as C and Gamma are required to further describe the SVM model. Therefore, to determine the best values, a grid-search of a specified parameter- space incorporating cross-validation (CV) is used. For this, SKikitlearns GridSearchCV() takes in as its parameters the classification model, the parameter-space and the number of CVs. The training data is 'fitted', and all combinations within the parameter space is tested to get the best parameter combination.

Once the best parameter combination has been obtained, a Learning Curve is plotted such that the bias and variance of the model can be determined. This is done using model selections' function learning_curve().The training curve is shown in figure. The learning curve gives insights on whether our C value makes the SVM to overfit, if the training accuracies are really high. Accordingly C value can be adjusted.

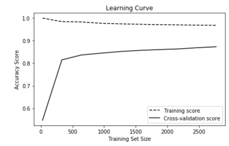

Fig. 3. Learning curve for this task

Finally, the new SVM model is trained using the training data, and tested using the validation data.

4.2. Task A2: Emotion Classification

As mention in model description section, the facial landmarks predictor is used from the Dlib library. The predictor is defined using the shape_predictor() function, with a .dat file as input. The .dat file describes the land mark structure used in the predictor. It must be noted that the shape predictor object requires both an image and an object of type dlib.rectangle. The rectangle object is the bounding box of the detected face. For simplicity, Dlib's get_frontal_face detector() is used as it returns the correct object type.The dlib detector is able to identify faces in 4822 images. Corresponding smile label values are extracted from the labels file and stored in a list.

The mouth facial landmarks are chosen in particular to determine smile. This corresponds to landmarks 48 to 67 - which amounts to a total of 20 (x,y) coordinates or 40 features per face. In addition to that, a Mouth Aspect Ratio has been used [29], which is defined as follows:

Hence each image has a total of 41 features. All the features are stored in a 2D array. The remaining steps of splitting the samples for training and validation, scaling the features and training the SVM are done as in task A1.

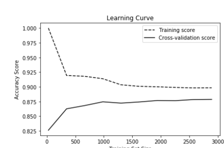

Fig. 4. Learning curve for this task

4.3. Task B1: Face Shape Detection

The Cartoon Dataset requires not much pre-processing as they are all frontal facing, uniform intensity images.

As in task A2, facial landmarks predictor from Dlib library is used to obtain all landmarks. The landmarks corresponding to the jaw are particularly extracted as they primarily determine the face shape. This corresponds to landmarks

0 to 16 - a total of 17 (x,y) coordinates or 34 feature values per face. The dlib detector is able to obtain 8090 images with all the jawline features. Corresponding labels are stored in a list.

Further, the feature set is split into 5 according to their labels in order to perform outlier detection and remove images with beards from which face shape cannot be predicted correctly. For this DBSCAN function from sklearn module is used, giving a radius (epsilon value) of 4 and minimum number of points to make a cluster as 10. DBSCAN is performed on the 5 split feature sets to detect outliers in each label and a new feature set is created without the outliers.

Further steps for training and validation are done as in the previous task.

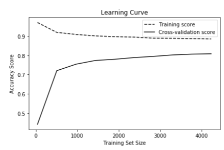

Fig. 5. Learning curve for this task

4.4. Task B2: Eye Colour Detection

As in task B1, facial landmarks predictor from Dlib library is used to obtain all landmarks. Since we are interested only in the eye colour, the landmarks corresponding to right eye alone are extracted. Using these coordinates, a rectangular bounding box is created to extract the eye. The image of the eye is saved as a jpeg file in 100x100 size.

Using skimage's color module, rgb2hsv function is used to obtain the eye image in HSV colour space. From this, the hue channel of each image alone is extracted, which corresponds to the 3rd dimension of the 3D array in which the HSV image is stored. Then the mean value of hues are calculated using numpy's mean function and stored in a list. Thus there is just one feature- the average hue value - per image. Corresponding labels are also stored in a separate list.

As mentioned in model description section, the dataset has several faces with glasses from which eye colour cannot be detected. Hence IQR based outlier detection is used to detect and remove those outliers. The feature set is split into five on the basis of their label values in order to detect the outliers corresponding to each label. For this numpy's nanquantile function is used to determine the first and 3rd quartiles. The IQR is calculated as follows:

$$IQR = Q3 - Q1$$

The outliers will be those values which lie outside the range

$$(Q1 - 1.5 * IQR, Q3 + 1.5 * IQR)$$

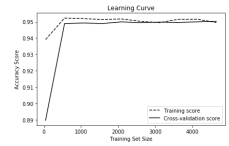

Fig. 6. Learning curve for this task

The new feature set is created after removing outliers, which consist of 8307 samples. Further steps for training and validation are done as in the previous task.

5. EXPERIMENTAL RESULTS AND ANALYSIS

The following accuracies were obtained on each task while training, validating and testing:

Task	Model	Train Acc	Val Acc	Test Acc
A1	SVM-RBF	95%	88%	88%
A2	SVM-RBF	91%	88.16%	88.8%
B1	SVM-RBF	89%	80.8%	82.12%
B2	SVM-RBF	94.5%	95%	95%

As can be seen, using SVM-RBF with chosen feature extraction methods has resulted in good accuracies on such a varied dataset. Performing DBSCAN on dataset for task B1 helped push accuracy from as low as 46% to 82% once outlier detection was performed.

6. CONCLUSION

The machine learning tasks of gender classification, emotion classification, face shape detection and eye colour recognition were performed satisfactorily using various feature extraction and classification techniques. The learning curve gave insights on how the chosen model fitted the data, and accordingly the hyperparameters were tuned.

The project also reaffirms the use of Support Vector Machine as a robust and reliable classifier for various classification tasks, with the right choice of feature vectors.

7. REFERENCES

[1] Benabdelkader C and Griffin P, "A local region-based approach to gender classification from face images.," in *Computer vision and pattern recognition-workshops*. IEEE Computer Society Conference, 2005, p. 52.

[2] Lu BL Li B, Lian XC, "Gender classification by combining clothing, hair and facial component classifiers," in *Neurocomputing*, 2012, vol. 76, pp. 18–27.

[3] Tay Y. Goi B Ng, C., "A review of facial gender recognition.," in *Pattern Anal Applic*, 2015, vol. 18, pp. 739–755.

[4] Sejnowski T Golomb B, Lawrence D, "Sexnet: a neural network identifies sex from human faces," in *Adv Neural Inf Process Syst*, 1991, vol. 3, pp. 572–577.

[5] Yang M Moghaddam B, "Learning gender with support faces.," in *Pattern Anal Mach Intell*. IEEE, 2002, vol. 24(5), pp. 707–711.

[6] Pietikainen M Ojala T, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns.," in *Pattern Anal Mach Intell*. IEEE, 2002, vol. 24(7), pp. 971–987.

[7] Pietikäinen M. Hadid, A., "Combining appearance and motion for face and gender recognition from videos," in *Pattern Recognition*, 2009, vol. 42(11), pp. 2818–2827.

[8] Shan C, "Learning local binary patterns for gender classification on real-world face images.," in *Pattern Recognition Lett*, 2012, vol. 33(4), pp. 431–437.

[9] J. Lee, C. Lin, and C. Huang, "Novel features selection for gender classification," in *2013 IEEE International Conference on Mechatronics and Automation*, 2013, pp. 785–790.

[10] M. Annalakshmi, S. Mohamed Mansoor Roomi, and A. Sheik Naveedh, "A hybrid technique for gender classification with slbp and hog features," *Cluster Computing*, vol. 22, no. 1, pp. 11–20, Jan 2019.

[11] B.C. Ko, "A brief review of facial emotion recognition based on visual information," *Sensors*, vol. 18, pp. 401, 2018.

[12] J Ghimire, D.; Lee, "Geometric feature-based facial expression recognition in image sequences using multi-class adaboost and support vector machines.," *Sensors*, vol. 13, pp. 7714–7734, 2013.

[13] A.; Routray A Happy, S.L.; George, "A real time facial expression classification system using local binary patterns..," 4th International Conference on Intelligent Human Computer Interaction, 2012, pp. 1–5.

[14] S.; Lee J.; Park S.H Ghimire, D.; Jeong, "Facial expression recognition based on local region specific features and support vector machines.," in *Multimed. Tools Appl.*, 2017, vol. 76, pp. 7803–7821.

[15] Adonis Emmanuel Tio, "Face shape classification using inception v3," 2019.

[16] Chu Kiong Loo Kitsuchart Pasupaa, Wisuwat Sunhema, "A hybrid approach to building face shape classifier for hairstyle recommender system," in *Expert Systems with Applications*, 2019, vol. 120, pp. 14–32.

[17] W. Sunhem and K. Pasupa, "An approach to face shape classification for hairstyle recommendation," in *2016 Eighth International Conference on Advanced Computational Intelligence (ICACI)*, Feb 2016, pp. 390–394.

[18] Michael Freeman Zezhi Chen, Nick Pears and Jim Austin, "A gaussian mixture model and support vector machine approach to vehicle type and colour classification," in *IET Intelligent Transport Systems*, 2013, vol. 8(2), pp. 135–144.

[19] Cheng-Jin Du and Da-Wen Sun, "Pizza sauce spread classification using colour vision and support vector machines," in *Journal of Food Engineering*, 2005, vol. 66(2), pp. 137–145.

[20] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, Dec 2001, vol. 1, pp. I–I.

[21] Matthew Wasserman, "Face/off: High speed facial tracking using the viola jones method," *https://towardsdatascience.com/face-off-high-speed-facial-tracking-using-the-viola-jones-method-f9e0ba491c9f*, Jul 2019.

[22] "Histogram equalization," *https://www.math.uci.edu/icamp/courses/math77c/demos/hist*$_e$*q.pdf*.

[23] "Histogram of oriented gradients (and car logo recognition)," *https://gurus.pyimagesearch.com/lesson-sample-histogram-of-oriented-gradients-and-car-logo-recognition/*.

[24] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, June 2005, vol. 1, pp. 886–893 vol. 1.

[25] Yun Tie and Ling Guan, "Automatic landmark point detection and tracking for human facial expressions," in *EURASIP Journal on Image and Video Processing*, 2013, vol. 1, p. 8.

[26] Z. Huang, G. Song, Y. Zhao, J. Han, and X. Zhao, "Smile recognition based on support vector machine and local binary pattern," in *2018 IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, July 2018, pp. 938–942.

[27] Davis E. King, "Max-margin object detection," 2015.

[28] Tereza Soukupova and Jan Cech , "Real-time eye blink detection using facial landmarks," in *21st Computer Vision Winter Workshop*, 2016.

[29] "Smilefie: how you can auto-capture selfies by detecting a smile," *https://www.freecodecamp.org/news/smilfie-auto-capture-selfies-by-detecting-a-smile-using-opencv-and-python-8c5cfb6ec197/*, 2017.

[30] Daniel Chepenko, "A density-based algorithm for outlier detection," 2018.