Multiple Classifiers for Age Prediction Against AAM and ASM

Musab Iqtait, Fatma Susilawati Mohamad

ABSTRACT--- In recent years, the researchers on age prediction relied on face pictures to get more attention, due to their important applications in security control and human computer interaction. Age prediction incorporates two processes: traits elicitation and prediction of machine learning. In the aspect of face traits elicitation, accurate and robust location for the trait point is convoluted and becoming a challenging issue in age prediction. Active Shape Model (ASM) can elicit the facial shape effectively and correctly. Furthermore, as the improvement of ASM, Active Appearance Models (AAM) is proposed to elicit both shape and texture traits from facial images simultaneously. In this paper, the two models are tested and compared for their performance against 6 algorithms which are Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Support Vector Regression (SVR), Canonical Correlation Analysis (CCA), Linear Discriminant Analysis (LDA), and Projection Twin Support Vector Machine (PTSVM). The experiments show that ASM is faster and gains more precise result than the AAM.

Index Terms— Active Appearance Model (AAM), Active Shape Model (ASM), age prediction, machine learning.

I. INTRODUCTION

The human face reflects many of properties for example, identity, age, gender and expression. With the time passing, the facial presence progressions as individual aging, this demonstrates individual attitude and preference. In spite of distinctive individuals are aging differently and age indicates Different structures in different ages [1].

Individual age can be specifically inferred by recognized patterns from the facial manifestation. For the same person, the photographs taken during distinctive years uncover the aging procedure on their faces. When the time interval is longer then there will be more evident changes, similarly as demonstrated in Fig. 1. Age data plays a paramount part for human computer interaction and artificial intelligence systems, also offers a significant number in other face-related errands for example, face recognition.

For late years, the attention toward human facial age prediction has fundamentally expanded [2]. An ordinary process of the existing strategies to age prediction as a rule comprises for two steps [3]: (1) collecting picture traits for age, and (2) learning the age predictor for these picture traits. Many facial traits extractor have been developed for facial age prediction. Some of them, biologically inspired features (BIF) proposed by [4] indicates the best execution for age prediction and it was wildly utilized. For the got picture traits for age, different techniques were suggested to learn an age predictor. The greater part for these methods,

age prediction is regarded as either a classification problem [5] or a regression problem [6]. Recently, machine learning schemes, particularly Convolutional neural Networks (CNNs), have been effectively utilized for a number assignments identified with facial analysis, including face detection, face arrangement [7], face confirmation [8], and demographic estimation [9]. In [10] concentrated characteristic maps got in distinctive layers Concerning illustration age offers dependent upon the deep learning model. In [11] gave a careful assessment on deep learning for age prediction and compared it with the hand-crafted combination traits.

In [23] suggested extra points viewing AAM and ASM models to facial traits extraction. They applied their experiment on the one dataset of the faces. This experiment show that ASM is speedier Also have more exact traits points than AAM, yet AAM gets a superior match of the texture. In [23] suggested an enhanced age recognition framework utilizing active appearance mode (AAM) model also three classifiers, support vector machine (SVM), K-nearest Neighbor and Support Vector Regression (SVR) to enhance the accuracy of age recognition.

In plurality pattern recognition systems, it is extremely critical to elicit the goal from the pictures. The restriction and arrangement from target item of the picture is an undertaking for extraordinary importance. In recent years, a large number of techniques proposed to manage those issue including ,Active Shape Models (ASM) [12], Active Appearance Models (AAM) [13], deformable template [14] and Gabor wavelet networks [15].

Through all these algorithms, ASM and AAM are both dependent upon statistical models which are proved that it is effective Furthermore, it powerful to picture interpretation. Over ASM, local texture on the trend perpendicular of the contour, so known as profile, it is utilized to shape that local texture from claiming each point of interest Also hunt to those points locally. Those comprehensive state models are after that connected with "correct" those nearby quest outcomes as stated by the statistical state model. Obviously, in the ASM, only local texture is utilized at the same time the global texture is not used to oblige the face state elicitation. Therefore, ASM are inclined will trap under local negligible due to vague local texture or poor beginning position.

In AAM model, global statistical shape and texture constraints would join for fabricate appearance model. The thing that will be more, a straight prediction model will be

assembled on foresee those appearance parameters for enhancement. Along these lines not just those shape as well as the texture of the aimed item a chance to be investigated by those analysis-by-synthesis standards. However, since the linear relations existed just in limited range. Therefore, AAM model might not have the capacity with localize person landmarks carefully because of complex background or picture variations.

Depend on the previous observations, this paper introduce a comparison between the ASM model and AAM model in age prediction depending on SVM, SVR, KNN, CCA, LDA and PTSVM machine learning algorithms with MORPH database.

II. METHODLOGY

This section first present overview of the proposed approach, then discuss a database used. A short introduction of facial traits elicitation is given, primarily focusing for the reason the figured out how "deep" traits would powerful clinched alongside characterizing the age labels and entryway to settle on utilization of the learned trait in age prediction issue. After that, ASM and AAM models are examined in traits elicitation to its useful execution with catch those underlying aging structure and explain the age data in a low dimensionality. Likewise, this approach explored a few prominent machine learning algorithms including Support vector machines (SVM), K-nearest Neighbor (KNN), Support Vector Regression (SVR), Linear Discriminant Analysis (LDA), Canonical Correlation Analysis (CCA) and Projection Twin Support Vector Machine (PT-SVM) for age prediction.

A. Overview of the Proposed Approach

As shown in Fig. 2, this section describes the proposed approach which consist of 3 phases:

First phase includes traits extraction from facial images which depends on ASM and AAM models. These models have three steps to extract traits which are face detection, image normalization and traits elicitation in vectors.

Second phase is classification: The elicited traits are trained and tested under different ages based on six machine learning algorithms, the predicted ages are the output of this phase.

Third phase is results comparison: This phase compared the results based on ASM model with AAM models using machine learning algorithms.

B. Face Database

In this experiment, MORPH database used [16], it contains 55,132 face pictures related to 13,000 subjects. The average value for pictures for every Subject is 4. Those ages of the face pictures range starting with 16 to 77 with an average age of 33. The facial images are taken from diverse races, around which the African facial images represent around 77%, the European facial images represent over 19%, and the remaining 4% are Asian, Indian, and other races.

C. Facial Traits Elicitation

For discover aging patterns from Different facial pictures, it require a model that have a capable learning ability.

Recently, ASM and AAM models have been utilized to elicit various traits alongside a hierarchical structure. The constructed structure might catch hierarchical Furthermore unique traits. ASM and AAM models showed very guaranteeing and possible results about various applications. This approach figure that ASM and AAM models are great toward characterizing the high level facial traits. This section describes ASM and AAM models which used for facial traits elicitation and gives details on these models.

1) Active Shape Model (ASM)

Shape models are based on a group of annotated landmarks. Those annotated landmarks to every picture would be structured as a vector. Following adjusting these vectors under a popular coordinate, principal component analysis is used to obtain a group about orthogonal foundation p. Each adjusted shape could nearly appeared like x ≈ x + pb, b is form variable. Furthermore shape model, neighborhood texture model to every point are assembled likewise. In [12], [13] utilize normalized primary derivate of profile should Fabricate the neighborhood texture model. The interval amidst a new point (its first derivate profile is g_{new}) and the model is $f(g_{new}) = (g_{new} - \bar{g})^T S_g^{-1}(g_{new} - \bar{g})$ where $\bar{g}$ is the mean of the first derivate of the profile, S_g is their covariance.

In light of those two models, search advancement could be completed. After initialization, every point in the model may be optimized by selecting the perspective with base separation specified previously in the trend vertical blonde of the shape for a specific range. Since those modern shape is potentially implausible, thus there is requirement on alter those state parameters to and in addition relative parameters in 2-D. Such steps iterated until no significant progress will be watched.

2) Active Appearance Model (AAM)

In this model, the shape model and additionally texture model is developed. By warping those picture encased by those shape of the mean shape, shape release texture is gotten. Then stratifying PCA of the shape spare texture, Also get those texture model $g = \bar{g} + P_g b_g$, where $\bar{g}$ related to the mean worth of the trained shape free texture $\{g_1 \ldots g_n\}$, P_g is the grid comprising k central orthogonal modes of the covariance in $\{g_1 \ldots g_n\}$, and b_g is texture variables regulating those difference of texture. Presence matrix is manufactured via concatenating the shape variables and texture variables $b = \begin{pmatrix} w_s b_s \\ b_g \end{pmatrix}$ where w_s in a diametrical grid permitting to those distinction between the state and gray models. Through stratifying PCA for the presence matrix {b}, the linkage amidst shape and texture may be removed, prompting for a further model $b = Qc$ so Q is the eigenvector Also, c is a vector from claiming presence variables.

AAM search is relied of the straight supposition between the texture contrast δg and the error in the model variables. The straight relations could be indicated as $\delta c = R_c \delta g$ and

$\delta t = R_t \delta g$. The variable δc will be the presence displacement. Furthermore, δt will be the position displacement comprehensive interpretation, turning, also scale displacement. Those straight grid R_c. Furthermore, R_t can be acquired eventually through offline learning. AAM optimizations might at that point be done with these straight relations.

D. Machine Learning Algorithms

Age prediction can be classified as a regression issue, since every age might be acknowledged as a regression value. In the age prediction issue could illustrate as classification issue, where it think as of every age as a class label. This approach used six machine learning methods for age prediction: KNN, SVM, SVR, CCA, LDA, and PTSVM.

1) Support vector machines (SVM)

A standout amongst the greater part imperative late researches over classifier outline may be those introduction for support vector machines classifier [17]. Those thought consists from claiming mapping the space S={x} of the information illustrations under a high dimensional trait space. SVM is essential a two-class classifier for which the streamlining paradigm is those width of the edge the middle of the classes. The streamlining about SVM comprises with minimize those amount about support vectors by expanding the edge between those two classes. The choice work inferred Toward the SVM classifier to a two class issue, this would be formulated, utilizing a Kernel function K(x,xi) of a new example x and a training example x_i , like following equation:

$$f(x) = \sum_{i \in SV} \alpha_i y_i K(x_i, x) + \alpha_0 \tag{1}$$

The SV will be the support vector set (a subset for training set). Also $y_i = \pm 1$, the mark from claiming illustration xi. Those parameters $\alpha_i \geq 0$ would optimized throughout those training procedure.

2) K-nearest Neighbor (KNN)

KNN is a standout amongst the vast majority basic and direct data mining systems [18]. It is known as memory based classification as those preparing cases require to be in the memory when the process running.

When managing constant qualities the contrast between those qualities is computed utilizing those Euclidean separation. If the to start with occurrence is (a1) and the second occurrence will be (b1), the distance between them will be computed utilizing the following equation

$$d(x, y) = (\sum_{i=1}^m ((x_i - y_i)^2))^{1/2} \tag{2}$$

The attributes x, y represents the histogram which related for X=Rm. the most training features with the age labels are inserted to this matrices.

3) Support Vector Regression (SVR)

The objective of the SVR issue is to make the hyper plane "close" to large portions of the preparing points as potential [19]. Provided N training points $(x_1, y_1), (x_2, y_2), ..., (x_N, y_N)$ with $x_i \in \mathbb{R}^n$ and $y_i \in \mathbb{R}$, i=1,...,N those hyper plane and qualities should build from claiming w and b. The hyper

plane w will be choose for a little norm, same time minimizing the aggregate of the distances from these training points of the hyper plane, measured eventually utilizing Vapnik's ε-insensitive loss function:

$$|y_i - (wx_i + b)|_\varepsilon =$$
$$\begin{cases} 0 & if \ |y_i - (wx_i + b)| \leq \varepsilon \\ |y_i - (wx_i + b)| - \varepsilon & otherwise \end{cases} \tag{3}$$

The worth of ε chosed via user, and the tradeoff between discovering the hyper-plane with a great regression execution is controlled through the provided regularization c. The QP issue connected with SVR will be portrayed concerning illustration.

$$\min_{w,b,\xi,\xi^*} \frac{1}{2} w^t w + c(\sum_{i=1}^L \xi_i + \sum_{i=1}^L \xi_i^*) \tag{4}$$
$$s.t. \ w^T \emptyset(x_i) + b - y_i \leq \epsilon + \xi_i$$
$$y_i - w^T \emptyset(x_i) - b \leq \epsilon + \xi_i^*$$

4) Linear Discriminant Analysis (LDA)

With age enrollment functions defined, they might make utilized within LDA by method for utilizing age participation functions. Concerning illustration class weighting qualities [20] as the following formula:

$$S_B = \sum_{i=1}^c N_i'(m_i - m)(m_i - m)^T, S_w$$
$$= \sum_{j=1}^c \sum_{i=1}^N \mu_j(x_i)(x_i - m_j)(x_i - m_j)^T$$
$$m = \frac{1}{N} \sum_{i=1}^N x_i, m_i =$$
$$\sum_{i=1}^N \mu_j(x_i) x_i / \sum_{i=1}^N \mu_j(x_i), N_j' = \sum_{i=1}^N \mu_j(x_i) \tag{5}$$

where N is the total value about training example. Similarly as training LDA, those ideal projections specified as below equation:

$$W_{opt} = \text{argmax}_w \frac{|W^T S_B W|}{|W^T S_W W|} = [w_1, w_2, ..., w_m] \tag{6}$$

where $[w_1, w_2, ..., w_m]$ is the group related to popularizing eigenvectors of $S_W^{-1} S_B$.

5) Canonical Correlation Analysis (CCA)

This approach Right away briefly presents the principle ideas about CCA. CCA is an approach for measuring the linear correlation the middle of two groups for attributes [21]. The SR context, the attributes are in vectors, and these vectors act a separated LR alternately HR picture, like its PCA coefficients, alternately remaining picture patches. CCA figures two bases, onto to any attributes, alongside which the corresponding correlation coefficients of the two groups of attributes are expanded.

The LR and HR attributes are changed under associated spaces, termed coherent subspaces, when anticipated under these new bases. Concerning illustration indicated after the pronouncement based SR algorithm may be more active to

these coherent subspaces, due to those improved comparability between LR also HR neighborhoods. Assume two groups are provided about m vectors P. Furthermore, Q for zero means.

$$P = \{P_j\}_{j=1}^m = [P_1, \dots, P_m], \quad Q = \{Q_j\}_{j=1}^m = [Q_1, \dots, Q_m] \tag{7}$$

In the SR issue, Pj Furthermore Qj, respectively, perform to the PCA coefficients of LR and hr forms of a face picture throughout global face reconstruction, alternately LR and hr lingering patches when performing remaining recompense. The transformed vectors U Also v are characterized as.

$$u = (W_P)^T P, \; V = (W_q)^T Q \tag{8}$$

where W_p also W_q mean the comparing basis vectors in the coherent subspaces. The goal of CCA will be to acquire basis vectors W_p and W_q which expand that correlation coefficient of u and v:

$$\rho(u, v) = \frac{W_P^T C_{PQ} W_q}{\sqrt{W_P^T C_{PP} W_P W_q^T C_{QQ} W_q}} \tag{9}$$

where $C_{PP}=E[PP^T]$ also $C_{QQ}=E[QQ^T]$ exemplify the within-set covariance matrices about p and Q, respectively, same time $C_{PQ}=E[PQ^T]$ and $C_{QP}=E[QP^T]$ mean their covariance matrices. E[.] is mathematical expectation. To discover W_p and W_q, first calculate:

$$R_1 = C_{PP}^{-1} C_{PQ} C_{QQ}^{-1} C_{QP}, \; R_2 = C_{QQ}^{-1} C_{QP} C_{PP}^{-1} C_{PQ} \tag{10}$$

6) Projection Twin Support Vector Machine (PT-SVM)

The major goal for linear projection twin support vector machine is to scan for a projection axis to all class [22], as well within-class disparity of the predicted points is decreased next the expected points of the characteristic classes are fixed. So, the fundamental problems for PTSVM are from QPPs.

$$min_{w1} \frac{1}{2} w_1^T S_1 W_1 + c_1 e_2^T \xi_2 \tag{11}$$

$$s.t. \; Bw_1 - \frac{1}{m_1} e_2 e_1^T A W_1 + \xi_2 \geq e_2, \; \xi_2 \geq 0$$

$$min_{w2} \frac{1}{2} w_2^T S_2 W_2 + c_2 e_1^T \xi_1 \tag{12}$$

$$s.t. - \left(Aw_2 - \frac{1}{m_2} e_1 e_2^T BW_2\right) + \xi_1 \geq e_1, \; \xi_1 \geq 0$$

where $c_1>0$ and $c_2>0$ are trade-off constants, $e_1 \in R^{m1}$ and $e_2 \in R^{m2}$ are both matrices of ones, and ξ_1 and ξ_2 are both non negative slack attributes vectors. S_1 and S_2 are within-class variance matrix which are expressed as

$$S_1 = \sum_{i=1}^{m_1} \left(x_i^{(1)} - \frac{1}{m_1} \sum_{j=1}^{m1} x_j^{(1)}\right) \left(x_i^{(1)} - \frac{1}{m_1} \sum_{j=1}^{m1} x_j^{(1)}\right)^T$$

$$S_2 = \sum_{i=1}^{m_2} \left(x_i^{(2)} - \frac{1}{m_2} \sum_{j=1}^{m2} x_j^{(2)}\right) \left(x_i^{(2)} - \frac{1}{m_2} \sum_{j=1}^{m2} x_j^{(2)}\right)^T \tag{13}$$

$$\tag{14}$$

III. RESULTS AND DISCUSSION

This section presents the experimental algorithm, and displays some experimental results. Also showed accuracy and performance comparison between ASM and AAM using machine learning algorithms: KNN, SVM, PT-SVM, SVR, LDA and CCA.

A. Experimental Setup

This part compare the usage of ASM and AAM Models as facial traits extractors utilizing accuracy and performance of machine learning algorithms in age prediction, the comparison depend on three distinctive measurements, which are point to point error, Mean Absolute Error (MAE) and Cumulative score (CS). Mean Absolute Error is computed utilizing the average of the absolute errors between the predicted results and the ground truth. MAE equation is follow:

$$MAE = \frac{1}{K} \sum_{k=1}^K |g_k - g_k'|, \tag{15}$$

where gk and gk0 perform the ground truth age and the predicted age separately. K will be the aggregate number from claiming test pictures. The cumulative score (CS) equation is as follows:

$$CS_a = \frac{K_a'}{K} \times 100 \tag{16}$$

where K_a' is the amount of facial pictures which have the absolute error between the predicted age and the real age is not more than α year.

Fig. 3 shows resulting point to point errors using 100 images from MORPH database. The horizontal axis is the point to point error between ASM and AAM models and the vertical axis is the percentage of image, the lines is the point to point error of all traits points. From the figure, we can consider that the point to point error of ASM model is less than AAM model except some cases.

Table 1 shows that MAE for all machine learning algorithms which used ASM model is less than machine learning algorithms which used AAM model. The best MAE value obtained on ASM model is 3.50 achieved in CCA compared to the AAM model result 4.17 achieved in CCA. The worst MAE value obtained on ASM model is 8.43 achieved in KNN, compared to the AAM model result 10.2 achieved in KNN.

Table 1: MAE for age prediction techniques using ASM and AAM models

Method / MAE	SVM	SVR	KNN	LDA	CCA	PT-SVM
AAM	5.71	4.68	10.2	5.67	4.17	4.67
ASM	4.85	4.61	8.43	5.13	3.50	4.64

With examine the execution from various machine learning algorithms in age prediction, experiments for testing execution about separate regression and

classification methods are directed over MORPH utilize AAM and ASM models as traits elicitors. The machine learning algorithms incorporate CCA, LDA, PTSVM, SVM, SVR, and KNN. The results are illustrated in Fig. 4 and 5. From the comparison of cumulative score, CCA performs best alongside predicting age based the elicited traits in both models (ASM and AAM). Experimental results exhibit that CCA performs better contrasted with other machine learning algorithms. The cumulative score results obtained by ASM traits is better than AAM traits in all machine learning algorithms.

B. Efficiency

The efficiency results conducted based on execution time calculation for all machine learning algorithms consecutively. Table 2 shows the time results for all machine learning algorithms using ASM model and AAM model separately.

Table 2: Efficiency for the age prediction algorithms using ASM and AAM models

Method Time	SVR	SVM	KNN	CCA	LDA	PT-SVM
AAM	0.059475	0.055351	0.070886	0.050661	0.245744	0.070886
ASM	0.049475	0.047014	0.019070	0.166884	0.086501	0.070886

Table 2 shows the running time outcomes for the age prediction methods using ASM and AAM separately. The test have been directed on MATLAB R2016a, computer with processor Intel(R) Core(TM) i5, 4.00 GB RAM memory and Operating System Windows 7 64-bit. In ASM model, the consequence indicates that KNN techniques get the most reduced time and the CCA method get the worst result. In AAM model, CCA get the best results and the worst result is for LDA method.

IV. CONCLUSION

In this paper, another structure to facial age prediction issue. A deep learning scheme for facial traits extraction rely on ASM and AAM models is firstly presented to this issue. We bring exhibited the viability of the recommended model over predicting age in view of those facial pictures. Our framework is assessed on one MORPH database. Diverse machine learning algorithms is additionally consolidated in the framework, indicating its effectiveness over discriminative subspace learning of the profound age label. CCA performs best around the used machine learning algorithms with both ASM and AAM models. The accuracy results obtained by ASM is better than AAM in all machine learning algorithms. The hopeful execution through the state-of-the-art researches expectations shows the possibility of the proposed approach towards real world applications.

REFERENCES

1. Z. Hu, Y. Wen, J. Wang, M. Wang, R. Hong, and S. Yan, "Facial age estimation with age difference," IEEE Transactions on Image Processing, 26(7), 2017, pp. 3087-3097.
2. H. Dibeklioğlu, F. Alnajar, A. A. Salah, and T. Gevers, "Combining facial dynamics with appearance for age estimation," IEEE Transactions on Image Processing, 24(6), 2015, pp. 1928-1943.
3. Y. Fu, G. Guo, and T. S. Huang, "Age synthesis and estimation via faces: A survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(11), 2010, pp. 1955-1976.
4. G. Guo, G. Mu, Y. Fu, and T. S. Huang, "Human age estimation using bio-inspired features," IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 112-119.
5. G. Guo, Y. Fu, T. S. Huang, and C. R. Dyer, "Locally adjusted robust regression for human age estimation," IEEE Workshop on Applications of Computer Vision, 2008. WACV 2008, pp. 1-6.
6. S. Yan, H. Wang, X. Tang, and T. S. Huang, "Learning auto-structured regressor from uncertain nonnegative labels," IEEE 11th International Conference on Computer Vision, 2007, pp. 1-8.
7. Y. Sun, X. Wang, and X. Tang, "Deep convolutional network cascade for facial point detection," IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 3476-3483.
8. Y. Taigman, M. Yang, M. A. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1701-1708.
9. M. Yang, S. Zhu, F. Lv, and K. Yu, "Correspondence driven adaptation for human profile recognition," IEEE Conference on Computer Vision and Pattern Recognition, 2011, pp. 505-512.
10. X. Wang, R. Guo, and C. Kambhamettu, "Deeply-learned feature for age estimation," IEEE Winter Conference on Applications of Computer Vision, 2015, pp. 505-512.
11. I. Huerta, C. Fernández, C. Segura, J. Hernando, and A. Prati, "A deep analysis on age estimation," Pattern Recognition Letters, 68, 2015, pp. 239-249.
12. T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," Computer Vision and Image Understanding, 61(1), 1995, pp. 38-59.
13. T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(6), 2001, pp. 681-685.
14. A. L. Yuille, "Deformable templates for face recognition," Journal of Cognitive Neuroscience, 3(1), 1991, pp. 59-70.
15. V. Krüger, and G. Sommer, "Gabor wavelet networks for object representation," in Multi-Image Analysis, R. Klette, G. Gimel'farb and T. Huang, Eds. Berlin: Springer, 2001, pp. 115-128.

Retrieval Number: B10800782S319/19©BEIESP
DOI : 10.35940/ijrte.B1080.0782S319

460

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

16. K. Ricanek, and T. Tesafaye, "Morph: A longitudinal image database of normal adult age-progression," IEEE 7th International Conference on Automatic Face and Gesture Recognition, 2006, pp. 341-345.
17. C. J. Burges, "A tutorial on support vector machines for pattern recognition," Data Mining and Knowledge Discovery, 2(2), 1998, pp. 121-167.
18. T. Cover, and P. Hart, "Nearest neighbor pattern classification," IEEE Transactions on Information Theory, 13(1), 1967, pp. 21-27.
19. A. J. Smola, and B. Schölkopf, "A tutorial on support vector regression," Statistics and Computing, 14(3), 2004, pp. 199-222.
20. Y. Yan, E. Ricci, R. Subramanian, G. Liu, and N. Sebe, "Multitask linear discriminant analysis for view invariant action recognition," IEEE Transactions on Image Processing, 23(12), 2014, pp. 5599-5611.
21. H. Hotelling, "Relations between two sets of variates," Biometrika, 28(3/4), 1936, pp. 321-377.
22. X. Chen, J. Yang, Q. Ye, and J. Liang, "Recursive projection twin support vector machine via within-class variance minimization," Pattern Recognition, 44(10-11), 2011, pp. 2643-2655.
23. M. Iqtait, F. S. Mohamad, and M. Mamat, "Feature extraction for face recognition via Active Shape Model (ASM) and Active Appearance Model (AAM)," IOP Conference Series: Materials Science and Engineering, 332(1), 2018, pp. 1-8.