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.

1. INTRODUCTION

The human face reflects many of properties for example, identity, age, gender and expression. With the time passing, the facial presence progresses 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 aging procedure on their faces. When the time interval is longer then there will be more evident changes, similarly as aging procedure on their faces. When the time interval is longer then there will be more evident changes, similarly as aging procedure on their faces.

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 multitude 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 obligle 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. METHODOLOGY

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 $x \approx x + pb$, 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 $g_{\text{new}}$) and the model is $f(g_{\text{new}}) = (g_{\text{new}} - \bar{g})^T S_g^{-1} (g_{\text{new}} - \bar{g})$ 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 $g = \bar{g} + Pg_{\text{bg}}$, where $\bar{g}$ related to the mean worth of the trained shape free texture $\{g_1, \ldots, g_n\}$. $P$ 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 $b = (w_g b_g)$ where $w_g$ is a diamentrical 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 $dg$ and the error in the model variables. The straight relations could be indicated as $\delta c = R_d \delta g$ and
\( \delta t = R \delta g \). The variable \( \delta c \) will be the presence displacement. Furthermore, \( \delta t \) will be the position displacement comprehensive interpretation, turning, also scale displacement. Those straight grid \( R \). Furthermore, \( R \) 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 that 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, the process running. The worth of \( \varepsilon \) chose 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.

\[
\begin{align*}
\min_{w,b,\xi_1^+,\xi_1^-} & \frac{1}{2} w^T w + c \sum_{i=1}^N (\xi_i^+ + \xi_i^-) \\
\text{s.t.} & w^T \phi(x_i) + b - y_i \leq \varepsilon + \xi_i^+ \\
& y_i - w^T \phi(x_i) + b \leq \varepsilon + \xi_i^- 
\end{align*}
\]

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:

\[
S_B = \sum_{i=1}^c \sum_{j=1}^N (m_i - m)(m_i - m)^T, S_w = \sum_{j=1}^c \sum_{i=1}^N \mu_j(x_i)(x_i - m_j)(x_i - m_j)^T
\]

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

\[
W_{opt} = \arg \max_w \frac{w^T S_B w}{w^T S_w w} = [w_1, w_2, ..., w_m]
\]

where \( [w_1, w_2, ..., w_m] \) is the group related to popularizing eigenvectors of \( S_B^{-1} S_w \).

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...
MULTIPLE CLASSIFIERS FOR AGE PREDICTION AGAINST AAM AND ASM

In the SR issue, \( P_j \) Furthermore \( Q_j \), 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.

\[
u = (W_p)^T P, \quad v = (W_q)^T Q
\]

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 \( v \):

\[
\rho(u, v) = \frac{w_p^T C_{pq} w_q}{\sqrt{w_p^T C_{pp} w_p} \sqrt{w_q^T C_{qq} w_q}}
\]

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:

\[
R_1 = C_{pp}^{-1} C_{pq}, \quad R_2 = C_{qq}^{-1} C_{qp}
\]

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 PT-SVM are from QPPs.

\[
\text{min}_{w_1} \frac{1}{2} w_1^T S_1 W_1 + c_1 e_1^T \xi_2
\]

s. t. \( B W_1 - \frac{1}{m_1} e_2 e_1^T A W_2 + \xi_2 \geq 0, \quad \xi_2 \geq 0
\]

\[
\text{min}_{w_2} \frac{1}{2} w_2^T S_2 W_2 + c_2 e_1^T \xi_1
\]

s. t. \( (A W_2 - \frac{1}{m_2} e_1 e_2^T B W_2) + \xi_1 \geq 0, \quad \xi_1 \geq 0
\]

where \( c_1 > 0 \) and \( c_2 > 0 \) are trade-off constants, \( e_1 \in \mathbb{R}^{m_1} \) and \( e_2 \in \mathbb{R}^{m_2} \) are both matrices of ones, and \( \xi_1 \) and \( \xi_2 \) are both non negative slack attributes vectors. \( S_1 \) and \( S_2 \) are within-class variance matrix which are expressed as

\[
S_i = \sum_{j=1}^{m_i} \left( x_j^{(1)} - \frac{1}{m_i} \sum_{j=1}^{m_i} x_j^{(1)} \right) \left( x_j^{(1)} - \frac{1}{m_i} \sum_{j=1}^{m_i} x_j^{(1)} \right)^T
\]

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

<table>
<thead>
<tr>
<th>Method</th>
<th>SVM</th>
<th>SVR</th>
<th>KNN</th>
<th>LDA</th>
<th>CCA</th>
<th>PT-SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASM</td>
<td>4.85</td>
<td>4.61</td>
<td>8.43</td>
<td>5.13</td>
<td>3.50</td>
<td>4.64</td>
</tr>
<tr>
<td>AAM</td>
<td>5.71</td>
<td>4.68</td>
<td>10.2</td>
<td>5.67</td>
<td>4.17</td>
<td>4.67</td>
</tr>
</tbody>
</table>

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, PT-SVM, 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

<table>
<thead>
<tr>
<th>Method</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SVR</td>
</tr>
<tr>
<td>AAM</td>
<td>0.059475</td>
</tr>
<tr>
<td>ASM</td>
<td>0.049475</td>
</tr>
</tbody>
</table>

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.

ACKNOWLEDGMENT

This work is financially supported by the Fundamental Research Grant Scheme (FRGS) RR227 Ministry of Education Malaysia and academically supported by Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin (UniSZA), Terengganu, Malaysia.

REFERENCES

MULTIPLE CLASSIFIERS FOR AGE PREDICTION AGAINST AAM AND ASM


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.