iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 Issue: VI Month of publication: June 2021

DOI: https://doi.org/10.22214/ijraset.2021.34869

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

Effective Refinement of Distinctive Analysis of the Facial Matrices for Automatic Face Annotation

Kavitha G L
Department of Information Science and Engineering, Atria Institute of Technology
Bangalore - 560024 , Karnataka, India.

Abstract: We deal with real world images which contains numerous faces captioned with equivalent names, it may be wrongly annotated. The face naming technique that we propose, exploits the weakly labeled image dataset, and aims at labeling a face in the image accurately. We propose this efficient face naming technique which is self regulated and aims at correctly labeling a face in an image. This is a challenging task because of the very large appearance variation in the images, as well as the potential mismatch between images and their captions.

This paper introduces a method called Refined Low-Rank Regularization (RLRR) which productively employs the weakly named image information to determine a low-rank matrix which is obtained by examining many subspace structures of the recreated data. From the recreation method used a discriminatory matrix is deduced. Also, Large Margin Nearest Neighbor (LMNN) method is used to label an image, which further leads to another kernel matrix, based on the Mahalanobis distances of the data and the two consistent facial matrices can be fused to enhance the quality of each other and it is used as a new reiterative method to infer the names of each facial image. Experimental results on synthetic and real world data sets validate the effectiveness of the proposed method.

***Index Terms: Refined* Low Rank Regularization (RLRR), Large Margin Nearest Neighbor (LMNN).**

I. INTRODUCTION

Internet based photo sharing has become advantageous to many current real world applications.

Most of the facial images shared over social media are wrongly annotated. A few approaches were projected in the literature for this image annotation problem. This paper aims at automatic image naming on the indeterminate affiliated captions.

Preliminary steps include using automated face detectors [1] and label entity detectors. The series of labels are expressed as the candidate label set. Notwithstanding these initial steps, self-regulated face labeling is very challenging because of the large appearance disparity in the images, and discrepancy between images and their captions. Moreover , the candidate label set may sometimes be disturbed and partial and thus a labeled image may not have the right labeled caption.

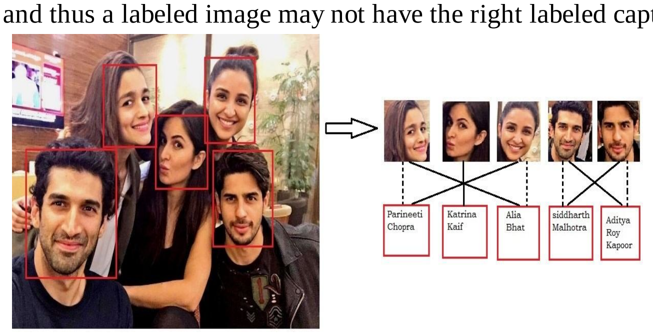


Fig. 1. An enclosure to the facial image annotation problem. The solid lines and the dotted lines represent the correctly named faces and the weakly annotated faces respectively.

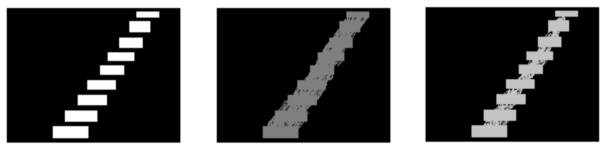
This paper introduces a new system of self-regulated face naming with label-based control aiming at automatic image naming. Two corresponding facial matrices are obtained by shaping the wrongly named images. These two matrices which are discriminated and merged into a single merged matrix based on which a reiterative plan is advanced for the self-regulated face naming.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

This paper introduces a new method called Refined Low Rank Regularization to obtain the first facial matrix by consolidating wrongly labeled image information from the Unsupervised Label Refinement (ULR) method, so that the recreated matrix can be ultimately obtained. To efficiently deduce the likeliness between the faces based on the visual appearance of the faces and the labels in the candidate label set, this paper accomplishes the subspace structures [2] among faces based on the following inference, that the faces of the same subject are present in the same subspace, and the subspaces are linearly absolute.

Universal Label Refinement (ULR) [3] is formulated to amplify the naming quality by using graph based and low-rank learning scheme. It is a structure to refine the labels of the facial images by discovering machine learning techniques.

Introducing the proposed method, the RLRR is a new refined regularized approach which combines with the caption based weak supervision into the unbiased ULR in which we reprimand the recreation of the faces using different subjects; and based on the interpreted recreated matrix we can cipher the similarity between each pair of faces.



(a) (b) (c)

Fig. 2.
(a) Original image W* according to the ground truth
(b) W* from ULR Algorithm
(c) W* from RLRR Algorithm (proposed system).

Additionally, the kernel matrix is based on the Mahalanobis distances among the faces as another equivalent facial matrix.

Large Margin Nearest Neighbor (LMNN) [4] scheme uses the Mahalanobis distances consistently improving the kNN (k nearest neighbor) classification using Euclidean distances. LMNN classification works better with PCA Principle Component Analysis than Linear Discriminant Analysis when some form of dimensionality reduction is required for preprocessing.

In consideration of RLRR and LMNN we analyze the weak supervision in discrete and creative way. The two corresponding facial matrices are combined to obtain a merged facial matrix that is employed for face labeling.

II. RELATED WORK

Utmost of the research in Automatic face labeling are focused on developing procedures for automatic image naming. Berg et al. [5] presented face clumping method to annotate the faces in news pictures. M Guillaumin [6] introduced the multiple-instance metric learning from automatically labeled bags of faces (MildML). Ozkan and Duygulu [7] developed a graph-based method by constructing the similarity graph of face. Zeng et al. [8] developed the low-rank SVM (LR-SVM) method which makes use of an assumption that the feature matrix of faces from the same subject is low rank. Luo and Orabona [9] developed learning from candidate labeling sets method for face naming.

Following is the comparison between our proposed method and existing systems:

Our proposed method RLRD is recounted to LR-SVM [8] and ULR [3]. In case of LR-SVM approach, LR-SVM considers distant supervision data in the permutation matrices, whereas RLRD utilizes regularizer that we have proposed, to deal with the recreation coefficients. In LR-SVM, data is not recreated by using itself as the base. In case of RLRD, it is related to the recreation-based approach of ULR. ULR is an unsupervised method that evaluates multiple subspace structures of data. Whereas, RLRD considers the image-level constraints to solve the face labeling problem in images.

Large-margin nearest neighbors (LMNN), is a traditional metric learning system. LMNN is constructed on appropriate supervision without any uncertainty. LMNN utilizes the hinge loss function. LMNN was proposed to learn distance metric M that supports the squared Mahalanobis distance between each training sample and its target neighbors to be smaller than those between this training sample and samples from other classes. The LMNN algorithm is built on the remark that the kNN will correctly classify an example if its k-nearest neighbors share the same label. The algorithm attempts to increase the number of training examples with this property by learning a linear transformation of the input space that precedes kNN classification using Euclidean distances

LMNN learns a distance metric that can be used to produce a facial matrix and can be fused with the facial matrix obtained from RLRR approach for the betterment of image labeling performance.

In the existing systems, such as MIL and MIML, data objects are represented as bags of instances. The distance between the data objects (bags) is a set-to-set distance. MIL makes use of class-to-bag distance, which assesses the relationships between the classes and the bags. The face labeling problem is solved by applying MIL and MIML method, in which each image is treated as a bag, faces in the image as the instances and names in the candidate name set as bag labels.

In some cases, the bag labels may be incorrect due to absence of names in the caption to which a face corresponds.

III. DIFFERENTIATION OF FACIAL MATRICES FOR AUTOMATIC FACE ANNOTATION

This paper introduces a new system of self-regulated face naming with label-based control. This is perplexing because of the inherent incongruity between various facial images and their captions. Here two facial matrices are worked on by making use of the equivocal labels, to perform image annotation based on the facial matrix obtained by fusing the two facial matrices. Further in the paper, the approach called Refinedd Low-Rank Regularization (RLRR) is briefed. The facial matrix obtained from this method is fused with the facial matrix obtained from the LMNN [4] method.

I_n is defined as the n×n similarity matrix, and 0_n , $1_n \in R_n$ as the n×1 column vectors of all zeros and ones, in the corresponding order. Also, we use I, 0 and 1 instead of I_n, 0_n, and 1_n in the case where the magnitudes are evident. tr(A) represents the trace of A and <A, B> means the dot product of two matrices. A∘B represents the element-wise multiplication of two matrices A and B (a∘b in case of vectors a and b). $\|A\|_\infty$ denotes the greatest absolute value of all the elements contained in matrix A. $\|A\|_F = (\sum_{i,j} A^2_{\ i,j})^{1/2}$ represents the Frobenious norm of the matrix A. $a \leq b$ implies that $a_i \leq b_i$ $\forall$ i = 1,...,n . $A \geq 0$ denotes that A is a positive semidefinite matrix (PSD matrix).

A. Problem Statement

Dealing with facial images which is captioned with analogous names, it may so happen that it may be wrongly annotated. The face naming technique that is proposed, is self regulated and aims at appropriately labeling a face in an image.

This wrong annotation may happen due to the variation in the images and mismatch between the images and their captions. This paper presents methods for face naming using collection of images with captions. This is carried out in the following steps:

1) Repossess all faces of a particular person from the data set.

2) Form the correct association between the names in the captions and faces in the image.

Let us assume that we have m images, each of which consists of r_i names and n_i faces, $\forall$ i = 1…..m. Let q ∈ {1,...,p} denote a name and $x \in R^d$ denote a face, where p is the total number of names in all the captions and *d* is the feature dimension. Thereafter, each image can be represented as (X^i, N^i), where $X^i = [x^i_{\ 1},...,x^i_{\ ni}] \in R^{d\times ni}$ is the data matrix for faces, that are in the *i*th image with each X^i_f being the *f*th face in the image (f = 1,...,N_i), and $N^i = \{q^i_1,..., q^i_{r_i}\}$ is the corresponding set of candidate names with each $q^i_j \in \{1,...,p\}$ being the *j*th name (j = 1,...,r_i). Further, let $X = [X^1.....X^m] \in \mathbb{R}^{d\times n}$ represent the data matrix of the faces from all m images, where $n = \sum_{i=0}^{m} n_i$.

After defining a binary label matrix $Y = [Y^1,...,Y^m] \in \{0,1\}^{(p+1)}$ with each $Y^i \in \{0,1\}^{(p+1)\times n_i}$ being the label matrix for each image X^i, the next step is to infer the facial label matrix Y based on the candidate name sets $\{N^i|_{i=1}^m\}$. When the ground-truth name of a face does not appear in the associated candidate name set N^i, we make use of the (p+1)th name to denote null class, so that the face can be assigned to the (p+1)th name. The label matrix Y^i for each image should satisfy the following image-level constraints [8].

1) Distinctiveness: In the same image, two faces cannot be annotated with the same name except the (p+1)th name, i.e., $\sum_{f=1}^{n_i} y^i_{j,f} \leq 1, \forall j = 1, ..., p$.

2) Expediency: the faces in the *i*th image should be tagged using the names from the set: $\tilde{N}^i = N^i \cup \{(p+1)\}$, i.e., $Y^i_{j,f} = 0, \forall f = 1, ..., and$ $j \notin \tilde{N}^i$.

3) Non-Pleonastic: In the ith image, each face should be tagged exactly one name from the set $\tilde{N}^i$, i.e., $\sum_j Y^i_{j,f} = 1, \forall f = 1, ..., n_i$.

B. Face Naming Using Facial Matrix

The feasible set of $\mathbf{Y}^i$ for the *i*th image, based on image-level constraints can be defined as follows:

$$y^i = \left\{ Y^i \in \{0,1\}^{(p+1)\times n_i} \;\middle|\; \begin{array}{c} 1'_{(p+1)}(Y^i \circ T^i)1_{n_i} = 0, \\ 1'_{(p+1)}Y^i = 1'_{n_i} \\ Y^i 1_{n_i} \le [1'_p, n_i]' \end{array} \right\} \quad (1)$$

The matrix $T^i \in \{0,1\}^{(p+1)\times n_i}$ has rows related to the indices of the names in N^i are all zeros and rest of rows are all ones.

The feasible set for the label matrix can be represented as $y = \{Y = [Y^1, \ldots, Y^m] \mid Y^i \in y^i \; \forall_i = 1, \ldots, m\}$.

Let A $\in \mathbb{R}^{n\times n}$ be a facial matrix, which meets the condition A=A' and $A_{i,j} \ge 0, \forall i,j$. Each $A_{i,j}$ expresses the pair-wise similarity between the *i*th face and the *j*th face. Our goal is to learn a proper A such that $A_{i,j}$ is large if and only if the *i*th face and the *j*th face share the same ground-truth name. Then, the face naming problem can be solved based on the facial matrix A obtained. We solve the following, to annotate the faces in an image:

$$\max_{Y \in y} \sum_{c=1}^{p} \frac{y_c' A y_c}{1' y_c} \quad s.t \;\; Y = [y_1, y_2, \ldots, y_{(p+1)}]' \quad (2)$$

$y_c \in \{0,1\}^n$ correlates to the *c*th row in Y. The faces with the same label are clustered as one group, and the sum of the average similarities for each group is maximized.

The RLRR method is proposed to learn the Unsupervised refined regularized Low-Rank recreation matrix. The first facial matrix is obtained from the RLRR method. Also, LMNN method is used to obtain another facial matrix. Finally, these two facial matrices are fused into one single facial matrix in order to perform image tagging.

C. Learning Discrimination of Facial Matrix with Refined Regularized-Low-Rank Dipiction (RLRR)

We will first analyse ULR (unsupervised label refinement) and then present our proposed method which is RLRR.ULR was proposed to enhance the face labelling quality via a graph-based and low-rank learning (LRR) approach. ULR makes use of content-based image search face annotation, face annotation performance on database. LRR is designed to solve the subspace clustering problem. The goal of LRR is to evaluate the structure of subspace in the given data X =[$x_1,\ldots,x_n$]$\in R^{d\times n}$. LRR attempts to obtain a recreation matrix W, which is based on an assumption that the subspaces have linearly independent vectors . This recreation matrix W is given by W = [$w_1,\ldots,w_n$]$\in R^{n\times n}$, where each w_i denotes the representation of x_i using X as the base. Because X is used as the base to recreate itself, the ideal solution W∗ of LRR encodes the pair-wise resemblance between the data matrices. The efficiency problem of LRR is given as:

$$\min_{W,E} \|W\|_* + \lambda \|E\|_{2,1} \quad \text{s.t} \;\; X = XW+E \quad (3)$$

where E $\in R^{d\times n}$ is the recreation error, $\lambda > 0$ is a tradeoff parameter, $\|W\|_*$ which is the nuclear form, is used to replace rank(W) as commonly used in the rank minimi- zation problems, and $\|E\|_{2,1} = \sum_{j=1}^{n} (\sum_{i=1}^{d} (E_{i,j})_2)^{1/2}$ is a regularizer that supports the recreation error E to be column-wise sparse. LRR performs better than sparse subspace clustering method, and hence produces better results in most of the real world applications that includes Faceprints.

Graph based method is proposed to determine the most relevant subset among the set of possible faces related to the query name, where the most relevant subset is likely to match with the faces of the queried person. Graph based method is implemented to rectify the correct faces of a queried person using both text and visual appearances. This approach eliminates the wrong tags, by applying geometrical constraint. The geometrical distance corresponding to the *i*th assignment refers to

$\sqrt{X^2 + Y^2}$ where,

$$X = \frac{\text{locX(i)}}{\text{sizeY(image1)}} - \frac{\text{locX(match(i))}}{\text{sizeX(image2)}}$$

$$Y = \frac{\text{locY (i)}}{\text{sizeY(image1)}} - \frac{\text{locY(match(i))}}{\text{sizeY(image2)}} \quad (4)$$

And locX is the X coordinate and locY is Y coordinate of the feature points in the images, sizeX and sizeY hold X and Y sizes of the images and match(i) corresponds to the matched keypoint in the second image of the ith feature point in the first image.
Unsupervised Label Refinement (ULR) task is to learn a refined label matrix $F^* \in R^{n\times m}$ to improve the initial raw label matrix Y . ULR makes use of an assumption called "label smoothness". i.e., the more similar the visual contents of two facial images, the more likely they share the same labels. The label smoothness principle is formulated as an idealization problem of reducing the following loss function $E_s(F,W)$:

$$E_s(F,W) = \frac{1}{2}\sum_{i,j=1}^{n} W_{i,j} \| F_{i*} - F_{j*}\|^2_F = tr(F^T LF) \qquad (5)$$

Where W is a weight matrix of a sparse graph, $\| \cdot \|_F$ denotes the Frobenius norm, L = D−W denotes the Laplacian matrix where D is a diagonal matrix with diagonal elements as $D_{ii} = \sum_{j=1}^{n} W_{i,j}$ and tr denotes a trace function.
We implement a new term $\|W \circ H\|^2_F$, which is called regularization term that includes the weak supervised information. Definition of $H \in \{0,1\}^{n\times n}$ depends on the candidate name sets $\{N^i|^m_{i=1}\}$. $H_{i,j} = 0$ if the following two conditions satisfy:
1) the i^{th} face and the j^{th} face has at least one name in common, in the corelated candidate name sets and
2) i = j. If not, $H_{i,j} = 1$.
And so forth, non-zero entries in W, where the corelated pair of faces have no names in common in their candidate name sets, and the entries that corelate to the situations where a face is recreated by itself, are penalized. Therefore, the resultant facial matrix W is expected to be more distinguishable, with information related to weak supervision encoded in H.
By implementing the new regularizer $\|W\circ H\|^2_F$ (5) can be reformulated into ULR, and the new optimization problem is achieved as
follows:

$$\min_{W,E} \|W\|_* + \lambda \|E\|_{2,1} + \frac{\gamma}{2}\|W\circ H\|^2_F \quad \text{s.t. } X = XW+E \qquad (6)$$

where $\gamma \geq 0$ is a used to balance the new refined regularizer with the other term. This problem is referred to as RLRR. By setting the parameter γ to zero, the RLRR problem in Eq(5) can be reduced to the ULR problem .
Once we obtain the ideal solution W^* after solving Eq(6), the facial matrix A_W can be computed as $A_W = \frac{1}{2}(W^* + W^{*'})$.
To obtain equivalent optimization problem , an intermediate variable J is introduced in Eq(6):

$$\min_{W,E,J} \|J\|_* + \lambda \|E\|_{2,1} + \frac{\gamma}{2}\|W\circ H\|^2_F \quad \text{s.t. } X = XW+E,\ W = J. \qquad (7)$$

Considering the following augmented Lagrangian function from Augmented Lagrangian Method (AML):

$$L = \|J\|_* + \lambda \|E\|_{2,1} + \frac{\gamma}{2}\|W\circ H\|^2_F + \langle U, X-XW-E\rangle + \langle V, W-J\rangle + \frac{\rho}{2}\left(\|X - XW - E\|^2_F + \|W-J\|^2_F\right) \qquad (8)$$

where ρ is a positive penalty parameter and $U \in R^{d\times n}$ and $V \in R^{n\times n}$ are the Lagrange multipliers. Notably, lets set the following parameters as follows:
$E_0 = X - XW_0$, $W_0 = (1/n)(1_n 1'_n - H)$, $J_0 = W_0$ and U_0, V_0 as zero matrices. The following steps are performed recursively at the *t*th iteration, until convergence is achieved.

1) Fix the others and update J_{t+1} by

$$\min_{J_{t+1}} \|J_{t+1}\|_* + \frac{\rho_t}{2}\left\|J_{t+1} - \left(W_t + \frac{V_t}{\rho_t}\right)\right\|^2_F$$

which can be solved in closed form using the singular value thresholding method.

2) Fix the others and update W_{t+1} by

$$\min_{W_{t+1}} \frac{\gamma}{2} ||W_{t+1} \circ H||_F^2 + \langle U_t, X - XW_{t+1} - E_t \rangle + \langle V_t, W_{t+1} - J_{t+1} \rangle + \frac{\rho_t}{2} ||X - XW_{t+1} - E_t||_F^2 + \frac{\rho_t}{2} ||W_{t+1} - J_{t+1}||_F^2 \quad (9)$$

Due to the new regularizer $||W \circ H||_F^2$
this problem cannot be solved as in [2] by using pre-computed SVD.
The gradient descent method is used to efficiently solve (7), where the gradient with respect to W_{t+1} is

$$(H \circ H) \circ W_{t+1} + \rho_t(X'X + I)W_{t+1} + V_t - \rho_t J_{t+1} - X'(\rho_t(X - E_t) + U_t)$$

3) Fix the others and update E_{t+1} by

$$\min_{E_{t+1}} \frac{\lambda}{\rho_t} ||E_{t+1}||_{2,1} + \frac{1}{2} \left\| E_{t+1} - \left(X - XW_{t+1} + \frac{U_t}{\rho_t} \right) \right\|_F^2$$

4) Update U_{t+1} and V_{t+1} by respectively using

$$U_{T+1} = U_t + \rho_t(X - XW_{t+1} - E_{t+1})$$
$$V_{t+1} = V_t + \rho_t(W_{t+1} - J_{t+1}).$$

5) Update ρ_{t+1}using
$\rho_{t+1} = min(\rho_t(1 + \Delta\rho), \rho_{max})$ where Δ_ρ and ρ_{max} are the constant parameters.

6) The iterative algorithm stops if the two convergence conditions are both satisfied

$$\|X - XW_{t+1} - E_{t+1}\|_\infty \leq \epsilon$$
$$\|W_{t+1} - J_{t+1}\|_\infty \leq \epsilon$$

where ϵ is a constant parameter.

D. *Large Margin Nearest Neighbor Classification (LMNN)*

Weinberger and Saul [4] proposed the LMNN method to learn a distance metric M that promotes the squared Mahalonobis distances between each training sample and its target neighbours to be smallers than the distance between this training sample and samples from other classes. In LMNN, the metric is trained with the goal that the k-nearest neighbors always belong to the same class and the examples from various classes are separated by a large margin. The algorithm is based on an observation that an example will be classified correctly by KNN decision rule, if its K-nearest neighbors share the same label. Large Margin Nearest Neighbor (LMNN) metric learning algorithm has been used widely in many applications and has produced promising results.

LMNN optimizes matrix M with the help of semidefinite programming. The objective is twofold: For every data point $\vec{x}_i$, the target neighbours should be close and imposters (differntly labelled) should be far away. The learned metric causes the input vector $\vec{x}_i$ to be surrounded by training instances of the same class. This optimization is illustrated in figure 3.

Let $\{(x_i, y_i)|^n_{i=1}\}$ be the n labeled samples: $x_i \in R^d$ denotes the i^{th} sample, with d being the feature dimension, and $y_i \in \{1,...,z\}$ denotes the label of this sample, with z being the total number of classes. $\eta_{i,j} \in \{0,1\}$ indicates whether x_j is a target neighbor of x_i.

i.e, $\eta_{i,j} = 1$ if x_j is a target neighbour of x_i, and $\eta_{i,j} = 0$ if x_j is a target neighbor of x_i ,$\forall$i,j $\in\{1..n\}$. $\nu_{i,l} \in\{0,1\}$ indicates whether x_l and x_i are from different classes. i.e, $\nu_{i,l} = 1$ if $y_l \neq \mathbf{y_i}$, and $\nu_{i,l} = 0$ if $y_l = \mathbf{y_i}$, $\forall$ i,l $\in\{1,...,n\}$. The squared Mahalanobis distance between t-wo samples x_i and x_j can be defined as:

$$d^2_M(x_i,x_j)=(x_i-x_j)'M(x_i-x_j).$$

LMNN minimizes the following idealization problem:

$$\min_{M\geq 0} \sum_{(i,j):\eta_{ij}=1} d^2_M(x_i,x_j) + \mu \sum_{(i,j,l)\in S} \xi_{i,j,l}$$

$$\text{s.t } d^2_M(x_i,x_l) - d^2_M(x_i,x_j) \geq 1-\xi_{i,j,l}\ ,\ \forall(i,j,l)\in S,\ \xi_{i,j,l} \geq 0\ ,\ \forall(i,j,l)\in S \qquad (10)$$

where $\xi_{i,j,l}$ is a slack variable, μ is a tradeoff parameter and $S = \{(i,j,l)|\eta_{i,j} = 1, \nu_{i,l} = 1, \forall i,j,l \in\{1,...,n\}\}$. Therefore, $d^2_M(x_i,x_j)$ is the squared Mahalanobis distance between x_i and its target neighbor x_j, and $d^2_M(x_i,x_l)$ is the squared Mahalanobis distance between x_i and x_j that belong to different classes. The slack variable can condone the cases when $d^2_M(x_i,x_l) - d^2_M(x_i,x_j)$ is smaller than one. The LMNN problem in Eq. (10) can be equivalently reformulated as the idealization problem as follows:

$$\min_{M\geq 0} \sum_{(i,j):\eta_{i,j}=1} d^2_M(x_i,x_j) + \mu \sum_{(i,j,l)\in S} |1 - d^2_M(x_i,x_l) + d^2_M(x_i,x_j)|_+$$

Where $|\cdot|_+$ is the truncation function.

Algorithm 1 summarizes the entire learning process.

Algorithm 1: LMNN

Input: Data samples $\{\mathbf{x_i,y_i}\}^N_{i=1}$,
number of target neighbors K, output
dimension m, maximum number of
optimization iterations T.
Result: matrix $L \in R^{d\times m}$
Initialize L with the first m leading eigen vectors of the covariance
matrix of the data samples $\{\mathbf{x_i}\}^N_{i=1}$;
For t=1 to T ***do***
Randomly generate subsamples S; Calculate
the descending direction d; Use line
search algorithm to find the step length λ; Update $L \leftarrow L + \lambda d$;
if the termination condition satisfies ***then*** ***break;***

IV. ANNOTATION OF FACIAL IMAGES

The first facial matrix A_w can be calculated as, $A_W = \frac{1}{2}(W^* + W^{*'})$, using coefficient matrix W^* learned from RLRR, and regularize A_W to the range [0,1]. The second facial matrix can be calculated from learnt distance metric M of LMNN as $A_K = K$, where K is a kernel matrix depending upon the Mahalanobis distance. These two facial matrices use weak supervision information in different ways. Therefore, the two facial matrices contain interdependent information which is beneficial for face annotation. Two facial matrices obtained from our RLRR and LMNN are combined to attain better accuracy, and this fused facial matrix is called as RLRR, which is our proposed method. This fused facial matrix A is the linear combination of the two facial matrices derived from RLRR and LMNN, where A is given by, $A=(1-\alpha)A_W + \alpha A_K$, where α is a parameter in the range [0, 1]. Lastly, the image face naming tagging is carried out based on A. Working on image face annotation is done by solving the following idealization problem:

$\max_{Y \in \mathcal{Y}} \sum_{c=1}^{p} \frac{y'_c A y_c}{1' y_c}$ S.T, Y = [$y_1 \ldots\ldots\ldots y_{(p+1)}$]'. (11)

But, the above problem is computationally expensive to solve. To solve this problem, we propose an iterative method. At each iteration, an objective function is approximated using $\tilde{y}'_c A y_c / 1' \tilde{y}_c$ that can substitute $y'_c A y_c / 1' y_c$, where $\tilde{y}_c$ is the solution for y_c inferred from the previous iteration. Therefore, we can solve the linear programming problem at each iteration, as follows:

$\max_{y \in \mathcal{Y}} \sum_{c=1}^{p}$ $b'_c y_c$, s.t. Y = [$y_1,\ldots,y_{(p+1)}$]′ (12)

where $b_c = A\tilde{y}_c / 1'\tilde{y}_c, \forall c = 1,\ldots,p$. f the faces may not annotated with their correct name.

The problem in Eq. (12) can be reformed by defining $B \in R^{(p+1)\times n}$ as B=[$b_1,\ldots,b_{p+1}$]. The reformulated form is as follows:

$\max_{y \in \mathcal{Y}}$ <B,Y> (13)

The viable set for Y is defined as $\mathcal{Y}$ = {Y = [$Y^1,\ldots,Y^m$]|$Y^i \in \mathcal{Y}^i$, ∀i =1,...,m}. Matrix B can be expressed as B=[$B^1,\ldots,B^m$] , where each $B^i \in R^{(p+1)\times n_i}$ correlates to Y^i. Then, the objective function in Eq. (13) can be conveyed as <B,Y> = $\sum_{i=1}^{m}$ <B^i,Y^i >. Therefore, Eq. (13) can be optimized by solving m sub-problems, with each sub-problem related to one image in the following form:

$\max_{y^i \in \mathcal{Y}^i}$ <B^i,Y^i> ∀i = 1,...,m (14)

The ith problem in Eq. (14) can be reformulated as a minimization problem as follows:

$$\min_{Y^i_{q,f} \in \{0,1\}} \sum_{q \in \mathcal{N}^i} \sum_{f=1}^{n_i} -B^i_{q,f} Y^i_{q,f}$$

S.T $\sum_{q \in \mathcal{N}^i} Y^i_{q,f} = 1 \quad \forall f = 1,\ldots,n_i$

$$\sum_{f=1}^{n_i} Y^i_{q,f} \le 1 \quad \forall q \in N^i$$

$$\sum_{f=1}^{n_i} Y^i_{(p+1),f} \le n_i \qquad (15)$$

in which the elements $\{Y^i_{q,f} |_{q \notin \mathcal{N}^i}\}$ are left out because these elements are zeros according to the feasibility constraint in Eq. (1). In this paper, the Hungarian algorithm is adopted to efficiently solve the problem in Eq. (15). Certainly, for an i^{th} image, the cost c(f, p +1) for assigning a face X^i_f to the corresponding null name is set to $-B^i_{(p+1),f}$ and the cost c(f , q) for assigning a face X^i_f to a real name q is set to $-B^i_{q,f}$.

The iterative face naming algorithm is as follows:

Algorithm 2: Face Naming Algorithm

Input: The feasible label sets $\{\mathcal{Y}^i|_{i=1}^m\}$, the affinity matrix A, the initial label matrix Y(1) and the parameters $\tilde{N}_{iter}, \theta$.

1: ***for*** t= $1 : \tilde{N}_{iter}$ ***do***

2: Update B by using B =$[b_1,\ldots,b_{p+1}]'$, where $b_c = \frac{A\tilde{y}_c}{1'\tilde{y}_c}, \forall c = 1,\ldots,p$ with $\tilde{y}_c$ being the c-th column of Y$(t)'$, and $b_{p+1} = \theta 1$.

3: Update Y(t +1) by solving m sub problems in Eq (14).

4: break if Y(t +1)=Y(t).

5: ***end for***

Output: the label matrix Y(t +1)

V. EXPERIMENTS

Analyzing the proposed schemes RLRR, and LMNN algorithms for face labeling using real-world datasets;

A. Real-world Datasets

1) Movie Face Database (MFD) - MFD is built from frames extracted from movies of different languages. MFD database consists of 4512 facial images corresponding to 430 actors collected from approximately 103 movies. MFD consists of 67 male and 33 female actors with at least 200 images for each actor.

VI. CONCLUSION

This paper investigated a promising search-based face annotation framework, in which the focus was on undertaking the critical problem of enhancing the label quality and proposed a RLRR algorithm.An approach was presented for face detection and naming which reduces computation time while attaining high detection accuracy. To effectively employ the face naming of the facial images we introduce RLRR and by using this scheme the evaluation of auto face annotation performance is increased. This proposed methods focus on tackling the critical problem of enhancing the label quality and accurately naming the facial images. Two challenging and interesting real-world datasets are analyzed from which it can be certified that this RLRR and LMNN overtakes ULR and kNN respectively and several other baseline algorithms. The future work will address and investigate other techniques to further improve the label refinement task.

TABLE I:
One Document Example With Naming Results LRR, ULR and RLRR, Shows The Maximum Number Of Accurately Named Faces In An Image.

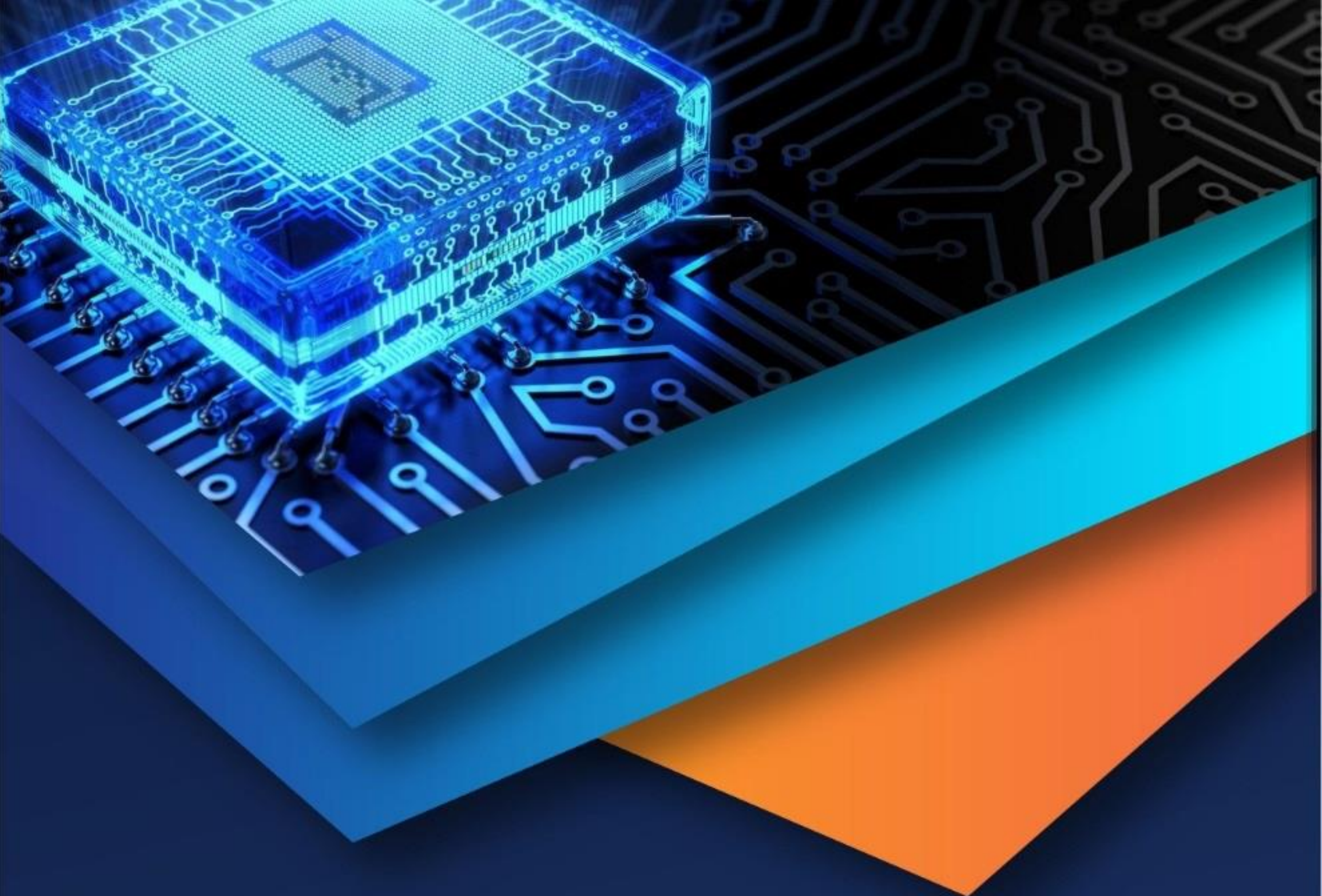
Images	LRR	ULR	RLRR
	Alia Bhat, Priyanka Chopra.	Alia Bhat, Priyanka Chopra, Parineeti Chopra, Tabbu	Tabbu, Parineeti Chopra, Priyanka Chopra, Alia Bhat,

REFERENCES

[1] Z. Lin, A review on low-rank models in data analysis, Big Data & Information Analytics 1 (2/3) (2017) 139–161.
[2] G. Liu, Z. Lin, and Y. Yu, "Robust subspace segmentation by low-rank representation," in Proc. 27th Int. Conf. Mach. Learn., Haifa, Israel, Jun. 2010, pp. 663–670.
[3] Dayong WANG, Chu Hong HOI, Ying HE, Jianke ZHU "Mining Weakly-Labeled Web Facial Images for Search-Based Face Annotation",2014.
[4] K. Q. Weinberger and L. K. Saul, "Distance metric learning for large margin nearest neighbor classification," J. Mach. Learn. Res., vol. 10, pp. 207–244, Feb. 2009.
[5] T. L. Berget al., "Names and faces in the news," in Proc. 17th IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., Washington, DC, USA, Jun./Jul. 2004, pp. II-848–II-854.
[6] M. Guillaumin, J. Verbeek, and C. Schmid, "Multiple instance metric learning from automatically labeled bags of faces," in Proc. 11th Eur. Conf. Comput. Vis., Heraklion, Crete, Sep. 2010, pp. 634–647.
[7] D. Ozkan and P. Duygulu, "A graph based approach for naming faces in news photos," in Proc. 19th IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., New York, NY, USA, Jun. 2006, pp. 1477–1482
[8] M. Guillaumin, T. Mensink, J. Verbeek, and C. Schmid, "Face recognition from caption-based supervision," Int. J. Comput. Vis., vol. 96, no. 1, pp. 64–82, 2012
[9] Q. Yao, J. T. Kwok, T. Wang, T.-Y. Liu, Large-scale low-rank matrix learning with nonconvex regularizers, IEEE Transactions on Pattern Analysis and Machine Intelligence 41 (11) (2018) 2628–2643.

Author Profile

Kavitha G L is an MTech (CSE) graduate, currently working as Asst. Professor in Atria Institute of Technology, Bangalore.
Her area of interests includes cloud computing and image processing.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY