Joint Patch Weighting and Moment Matching for Unsupervised Domain Adaptation in Micro-Expression Recognition

Jie ZHU\textsuperscript{1,††}, Yuan ZONG\textsuperscript{1,∗}, Hongli CHANG\textsuperscript{1,††}, Li ZHAO\textsuperscript{††}, Nonmembers, and Chuangao TANG\textsuperscript{1, Member}

SUMMARY Unsupervised domain adaptation (DA) is a challenging machine learning problem since the labeled training (source) and unlabeled testing (target) sets belong to different domains and then have different feature distributions, which has recently attracted wide attention in micro-expression recognition (MER). Although some well-performing unsupervised DA methods have been proposed, these methods cannot well solve the problem of unsupervised DA in MER, a. k. a., cross-domain MER. To deal with such a challenging problem, in this letter we propose a novel unsupervised DA method called Joint Patch weighting and Moment Matching (JPMM). JPMM bridges the source and target micro-expression feature sets by minimizing their probability distribution divergence with a multi-order moment matching operation. Meanwhile, it takes advantage of the contributive facial patches by the weight learning such that a domain-invariant feature representation involving micro-expression distinguishable information can be learned. Finally, we carry out extensive experiments to evaluate the proposed JPMM method is superior to recent state-of-the-art unsupervised DA methods in dealing with cross-domain MER.

key words: micro-expression recognition, domain adaptation, transfer learning, patch weighting, feature representation learning

1. Introduction

As a hidden emotional expression, the micro-expression (ME) is usually exposed on a person’s face when he is trying to conceal something [1]. Due to this important characteristic, accurately recognizing MEs have a lot of potential applications, e.g., lie detection and hence attracted wide attention among researchers from affective computing and computer vision. However, compared with the ordinary facial expression recognition [2]–[4], a ME has a much shorter duration whose boundary is often 500ms for the upper limit and about 170ms for the lower limit [5] which makes it difficult to be recognized. Moreover, the facial muscle movements caused by MEs are very lowly intensive. Consequently, it is barely perceptible for inexperienced individuals to recognize micro-expressions [6]. To help human beings better understand MEs, researchers have focused on developing effective automatic MER technique over past several years and proposed a lot of well-performing methods. For example, several spatiotemporal descriptors, e.g., local binary patterns from three orthogonal planes (LBP-TOP) [7], [8], histogram of image gradient from three orthogonal planes (HIGO-TOP) [9] and sparse main directional mean optical-flow (MMDMO) [10] have been designed for describing micro-expressions. In more recent years, deep learning methods have been also applied to MER and achieve more promising performance than conventional handcrafted methods such as three-stream convolutional neural networks (TSCNNs) [11], spatiotemporal recurrent neural networks (STRNNs) [12], and recurrent convolutional networks (RCNs) [13].

It should be pointed out that existing MER methods are still far away from the practical scenarios. One of the major reasons is that most of them are designed under an ideal assumption that the training and test data come from the same database. In this case, both ME datasets abide by the same or similar feature probability distributions. However, this assumption is very easily broken because the training and testing ME samples may be recorded by different cameras or in different environments. Hence, the performance of these methods may decrease sharply and bring us a more challenging MER task, i.e., unsupervised cross-domain MER. To tackle this problem in MER, several unsupervised domain adaptation (UDA) methods have been proposed in recent years. In the work of [14], Zong et al. firstly proposed a simple yet effective method called target sample re-generator (TSRG) and subsequently extend it to a domain regeneration framework to deal with unsupervised cross-domain MER. Zhang et al. [15] proposed a region selective transfer regression (RSTR) method and further designed a comprehensive unsupervised cross-domain MER evaluation protocol to evaluate the state-of-the-art unsupervised DA methods.

Inspired by the existing unsupervised cross-domain MER works, in this paper we also focus on this challenging but interesting topic and propose a novel unsupervised DA method called Joint Patch weighting and Moment Matching (JPMM). The basic idea of the proposed JPMM mainly includes two folds. First, JPMM takes into full consideration the facial patches contributing to distinguish different MEs in feature learning. Second, we also design a second-order moment matching loss for JPMM to alleviate the dissimilarity of the feature distributions between source and target ME datasets such that a domain-invariant feature to bet-

\textsuperscript{*}The authors with the Key Laboratory of Child Development and Learning Science of Ministry of Education, Southeast University, Nanjing 210096, China.
\textsuperscript{††}The authors are with the School of Information Science and Engineering, Southeast University, Nanjing 210096, China.
\textsuperscript{†}This work was supported in part by the National Key R & D Program of China under Grant 2018YFB1305200, in part by the NSFC under Grants 61902064 and 81971282, and in part by the Zhishan Young Scholarship of Southeast University.
\textsuperscript{a) E-mail: xhzongyuan@seu.edu.cn (Corresponding author)} DOI: 10.1587/transinf.2021EDL8045

Copyright © 2022 The Institute of Electronics, Information and Communication Engineers
ter describe MEs from both domains. Experimental results demonstrate the superior performance of JPMM over the recent state-of-the-art UDA methods in dealing with cross-domain MER.

2. Proposed Method

2.1 Problem Formulation

In this section, we will address the proposed JPMM and how to apply it to deal with unsupervised cross-domain MER in detail. Suppose we have $N_s$ labeled source micro-expression samples. Their corresponding features extracted based on the widely-used spatial division scheme and label information are denoted by $X^s \in \mathbb{R}^{M_{ds} \times N_s}$ and $Y^s \in \mathbb{R}^{N_{sc} \times N_s}$. We are also given $N_t$ unlabeled target micro-expression samples whose feature matrix is denoted by $X^t \in \mathbb{R}^{M_{dt} \times N_t}$. Note that in the above notations, $c$ is the class number of micro-expressions and each column of feature matrices $X^s$ and $X^t$ is a micro-expression sample’s feature vector with dimension of $Md$. It is obtained by concatenating the $d$ dimensional spatiotemporal descriptors, e.g., LBP-TOP, extracted from all the $M$ facial patches generated according to spatial division scheme, e.g., $8 \times 8$. Each column of $Y^s$ is a one-hot vector and its $j^{th}$ entry would be 1 if it belonged to the $j^{th}$ micro-expression. The task of unsupervised cross-domain MER requires us to learn new feature representations for both source and target micro-expression feature sets and then refer the micro-expression categories of $X^s$ given $X^t$ and $Y^t$.

2.2 Formulation of JPMM Model

To achieve the above goal, we raise the idea of JPMM model. The basic idea of JPMM is to learn a common projection matrix to bridge the gap between source and target micro-expression datasets by jointly picking out the facial patches having contributions to distinguish different micro-expressions and trying to eliminate the feature distribution difference of both domains. Under this consideration, we denote this projection matrix by $U \in \mathbb{R}^{M_{ds} \times M_{dt}}$ and enable it to model the relationship between the source micro-expression feature samples and their corresponding label information, which can be formulated as the following simple optimization problem:

$$
\min_{U} \|Y^s - \sum_{i=1}^{M} U^T X^s_i\|^2_F, 
$$

(1)

where $X^s_i$ and $U_i$ are the $i^{th}$ facial patch’s feature matrix and its corresponding sub-projection matrix satisfying $X^s = [X^s_1, \ldots, X^s_M]^T$ and $U = [U_1^T, \ldots, U_M^T]^T$.

Subsequently, we further enhance the above simple feature learning model following the basic idea. Specifically, we first endow it the ability to automatically pick out the typical facial patches from the entire face. This consideration for JPMM is mainly due to the theory of facial action coding system (FACS) [16], in which each facial/micro-expression can be thought to be encoded by a set of several facial action units (AU$s$). Hence, it would be benefit learning more discriminative feature representations for micro-expression if these important facial patches associated with micro-expression are selected and involved. To this end, we design a patching weighting regularization as follows:

$$
\mathcal{R}_1(U) = \sum_{i=1}^{M} \|U_i\|^2_F.
$$

(2)

According to the definitions of Frobenius norm and L2-norm, it can be found that $\|U_i\|_F^2$ equals $\|\tilde{U}_i\|_2^2$, where $\tilde{U}_i = [u_{i,1}^T, \ldots, u_{i,c}^T]^T$ and $u_{i,j}$ is the $j^{th}$ column of $U_i$ satisfying $U_i = [u_{i,1}, \ldots, u_{i,c}]$. Hence, we are able to obtain that $\sum_{i=1}^{M} \|U_i\|_F^2$ equals $\|\tilde{U}\|_{2,1} = \sum_{i=1}^{M} \|\tilde{U}_i\|_2$, where $\tilde{U} = [\tilde{U}_i, \ldots, \tilde{U}_M]$. Thus, minimizing $\|\tilde{U}\|_{2,1}$, which results in a column-sparse $\tilde{U}$, is actually equivalent to minimizing $\sum_{i=1}^{M} \|U_i\|_F^2$ leading to a set of sparse $U_i$. Consequently, it is clear to see that the projection matrix learned by JPMM including $\mathcal{R}_1$ would be able to ignore the useless facial patches while preserving contributive ones guided by its sparse sub-matrix.

Second, we design a multi-order moment matching regularization term jointly considering the mean and covariance information to alleviate the feature distribution mismatch between source and target micro-expression samples, for JPMM, which can be formulated as follows:

$$
\mathcal{R}_2(U) = \frac{1}{N_s} \|U^T X^s 1_s^T - \frac{1}{N_t} U^T X^t 1_t^T\|^2_F + \|\text{var}(U^T X^s) - \text{var}(U^T X^t)\|^2_F,
$$

(3)

where $1_s$ and $1_t$ are the column vectors with the sizes of source and target sample numbers and all the entries of 1, and $\text{var}(U^T X^s)$ and $\text{var}(U^T X^t)$ denote the covariance matrices of source and target micro-expression features projected by $U$. Hence, $\mathcal{R}_2(U)$ can be further rewritten as

$$
\mathcal{R}_2(U) = \|U^T \Sigma_{st} U\|^2_F, 
$$

(4)

where $\Sigma_{st} = \frac{1}{N_s} (X^s - \frac{1}{N_s} X^s 1_s^T) (X^s - \frac{1}{N_s} X^s 1_s^T)^T + \frac{1}{N_t} (X^t - \frac{1}{N_t} X^t 1_t^T) (X^t - \frac{1}{N_t} X^t 1_t^T)^T$. The main reason of designing such a regularization is inspired by the success of the recent DA works [17], [18], in which the authors aligned the high-order moments between source and target domains.

By jointly minimizing the above two well-designed regularization terms in Eqs. (2) and (4) and the original optimization problem in Eq. (1), we are able to reach the optimization problem of the proposed JPMM model, which can be written as:

$$
\min_{U} \|Y^s - \sum_{i=1}^{M} U^T X^s_i\|^2_F + \lambda \sum_{i=1}^{M} \|U_i\|_F^2 + \mu (\|\frac{1}{N_s} U^T X^s 1_s^T - \frac{1}{N_t} U^T X^t 1_t^T\|^2_F + \|U^T \Sigma_{st} U\|^2_F), 
$$

(5)

where $\lambda$ and $\mu$ are the trade-off parameters which control
the balance between the original objective function and the regularization terms.

2.3 Optimization

The optimization problem of JPMM in Eq. (4) can be efficiently solved by the Inexact Lagrangian Multiplier (IALM) method [19]. Specifically, we introduce two auxiliary matrices \( Q \) and \( P \) that equal \( U \). Then, the optimization problem of JPMM can be converted to a constrained one as follows:

\[
\begin{align*}
\min_{U, Q, P} & \quad ||y^t, 0||^2 - \sum_{i=1}^{M} p_i^T [X_i^t, \tilde{x}^t_i]_F^2 + \lambda \sum_{i=1}^{M} ||u_i||_F^2 \\
& + \mu ||Q^T \Sigma_u p||_F^2, \\
\text{s.t.} & \quad U_i = P_i, \quad U_i = Q_i,
\end{align*}
\]

where \( 0 \) is a zero vector and \( \tilde{x}^t_i = \frac{1}{N_i} U^T X_i^t - \frac{1}{N_i} U^T X_i^t' \). Subsequently, we are able to obtain its corresponding Lagrangian function, which can be expressed as follows:

\[
\begin{align*}
& L(U_i, P, Q, T_1, T_2, \kappa) = ||\tilde{y}^t - \sum_{i=1}^{M} P_i^T \tilde{x}^t_i||_F^2 \\
& + \lambda \sum_{i=1}^{M} ||u_i||_F + \mu ||Q^T \Sigma_u p||_F^2 + \tau r [T_1^T (U - P)] \\
& + \tau r [T_2 (U - Q)] + \kappa \left( \frac{1}{2} ||U - P||^2_F + ||U - Q||^2_F \right),
\end{align*}
\]

where \( \tilde{y}^t = [y^t, 0], \tilde{x}^t_i = [X_i^t, \tilde{x}^t_i], \) and \( T_1 \) and \( T_2 \) are the Lagrangian multiplier matrices, and \( \kappa \) is a regularization parameter. In order to obtain the optimal \( U \), we only need to iteratively minimize the Lagrangian function with respect to one of the variables while fixing the others. Specifically, repeating the steps in Algorithm 1 until obtaining convergence:

**Algorithm 1:** Algorithm for solving the optimization problem of Eq. (7).

1. Fix \( U_i, Q, T_1, T_2, \kappa, \) and update \( P \): In this case, the optimization problem is minimized by \( \min_{P} ||\tilde{y}^t - P^T \tilde{x}^t_i||_F^2 + \mu ||Q^T \Sigma_u p||_F^2 + \tau r [T_1^T (U - P)] + \frac{\kappa}{2} ||U - P||^2_F \) and has the close-form solution.

2. Fix \( U_i, P, T_1, T_2, \kappa, \) and update \( Q \): The optimization problem is then reformulated as \( \min_{Q} \mu ||Q^T \Sigma_u p||_F^2 + \tau r [T_1^T (U - Q)] + \frac{\kappa}{2} ||U - Q||^2_F \) and has the close-form solution.

3. Fix \( P, Q, T_1, T_2, \kappa, \) and update \( U \): The optimization problem can be rewritten as \( \min_{U} \frac{\lambda}{2} ||u_i||_F^2 + \frac{\mu}{2} ||U - (Q^T p + T_1^T p + T_2^T p)||^2_F \), where \( P, Q, \)

4. Update \( T_1, T_2, \) and \( \kappa \): \( T_1 = T_1 + \kappa (U - P), \) \( T_2 = T_2 + \kappa (U - Q), \) \( \kappa = \min(\kappa_r, \kappa_{max}), \) where \( \rho \) is a scale parameter and \( \kappa_{max} \) is preset maximal value of \( \kappa \).

5. Check Convergence: \( ||U - P||_F < \epsilon, ||U - Q||_F < \epsilon. \)

Once the optimal projection matrix of JPMM is learned, it is easy to predict the micro-expression labels of target samples. Suppose we have a target micro-expression sample and its feature vector \( x' \). The optimal projection matrix be denoted by \( U \). Then, we are able to calculate the label vector of such target sample by \( y' = U' x' \) and obtain its micro-expression category by \( \arg \max_i y'(i) \), where \( y'(i) \) is the \( i \)th entry of vector \( y' \).

### 3. Experiments

#### 3.1 Experiment Setting

In this section, extensive cross-domain MER experiments are carried out to evaluate the proposed JPMM. Following the experiment protocol designed by Zhang et al. [15], we adopt the selected CASME II [26] and three subset (HS, VIS, and NIR) of SMIC [27], which are relabeled by the same micro-expression labels, to serve as the source and target databases alternately. Among the selected CASMEII samples, Happy samples are given Positive labels, Disgust, Sad, and Fear samples are relabeled with Negative labels, and the label for Surprise samples remains unchanged. Hence, we have six experiments, i.e., \( C \rightarrow H, C \rightarrow V, C \rightarrow N, H \rightarrow C, V \rightarrow C, \) and \( N \rightarrow C \), where \( C, H, V, \) and \( N \) are the abbreviations of CASME II, SMIC-HS, SMIC-VIS, and SMIC-NIR, respectively. Table 1 shows the statistics of micro-expression samples in the above databases. In order to show the overall performance of the proposed JPMM, in the experiments we choose two different typical spatiotemporal descriptors, i.e., LBP-TOP [7] (the radius and number of neighboring points is set as \( R = 3, P = 4 \)) and LBP-SIP [28] (the radius is set as \( R = 3 \)) to describe micro-expressions, respectively. Before that, we first use TIM [29] method to normalize all the micro-expression image sequences into the same frame length (10 frames). Then, we resize each frame image into \( 112 \times 112 \) pixels. In addition, we also include recent state-of-the-art subspace learning based DA methods, i.e., transfer component analysis (TCA) [21], geodesic flow kernel (GFK) [22], transfer kernel learning (TKL) [23], subspace alignment (SA) [24], TSRG [14], domain regeneration in the Label space (DRLS) [25], and region selective transfer regression (RSTR) [15] in the comparison. For the sake of fair comparison, linear SVM with \( C = 1 \) is served as the classifier if needed and is implemented by LibSVM [30]. As for the evaluation performance metric, we adopt mean F1-score, which is less sensitive to class-imbalance existing in micro-expression databases.

Moreover, considering that the target domain data is unlabeled, it is impossible to determine the optimal param-

**Table 1** Sample statistics of the relabeled CASME II and SMIC databases used in the experiments.

| Database    | Negative | Positive | Surprise | Total |
|-------------|----------|----------|----------|-------|
| CASME II    | 73       | 32       | 25       | 130   |
| SMIC (HS)   | 70       | 51       | 43       | 164   |
| SMIC (VIS)  | 23       | 28       | 20       | 71    |
| SMIC (NIR)  | 23       | 28       | 20       | 71    |
3.2 Results and Analysis

The experimental results are given in Tables 2 and 3, where Table 2 shows the mean F1-scores of the DA methods using two different spatial-temporal descriptors in dealing with cross-domain MER. Note that we also calculate the average result of each method among the six cross-domain MER experiments. The best result in each experiment is highlighted in bold.

Table 2 The mean F1-scores of the DA methods using two different spatial-temporal descriptors in dealing with cross-domain MER. The best result in each experiment is highlighted in bold.

| Method | C → H | C → V | C → N | H → C | V → C | N → C | Average |
|--------|-------|-------|-------|-------|-------|-------|---------|
| SVM    | 0.3680| 0.4624| 0.3000| 0.3339| 0.5880| 0.1927| 0.3742  |
| TCA [21]| 0.4556| 0.7014| **0.6141**| 0.5196| 0.5870| 0.4114| 0.5482  |
| GFK [22]| 0.4008| 0.6598| 0.5259| 0.4508| 0.6078| 0.3437| 0.4981  |
| TKL [23]| 0.3722| 0.6178| 0.4811| 0.4488| 0.5250| 0.3684| 0.4689  |
| SA [24]| 0.4021| 0.6460| 0.4875| 0.4875| 0.5979| 0.4472| 0.5253  |
| TSRG [14]| 0.4968| **0.7101**| 0.4591| 0.4872| 0.6062| 0.4297| 0.5315  |
| DRLS [25]| 0.5318| 0.6712| 0.4950| 0.5311| **0.6737**| 0.4658| 0.5554  |
| RSTR [15]| 0.5382| 0.6955| 0.5174| 0.4841| 0.6941| 0.4889| **0.5697** |
| JPMM (Ours)| 0.5296| 0.6152| 0.5291| **0.5784**| 0.6142| 0.5305| 0.5662  |

| Method | C → H | C → V | C → N | H → C | V → C | N → C | Average |
|--------|-------|-------|-------|-------|-------|-------|---------|
| SVM    | 0.3772| 0.5846| 0.3469| 0.3742| 0.6065| 0.2790| 0.4281  |
| TCA [21]| 0.4596| **0.6980**| 0.5409| 0.5431| 0.7120| 0.5046| 0.5764  |
| GFK [22]| 0.4597| 0.6754| 0.5493| 0.5455| 0.5912| 0.3523| 0.5289  |
| TKL [23]| 0.4554| 0.6611| **0.5567**| 0.5009| 0.5815| 0.3247| 0.5134  |
| SA [24]| 0.4518| 0.6765| 0.5244| **0.5909**| 0.5782| 0.3984| 0.5367  |
| TSRG [14]| 0.5283| 0.6748| 0.5354| 0.5401| 0.6641| 0.4400| 0.5638  |
| DRLS [25]| 0.4895| 0.6560| 0.5274| 0.5770| 0.6300| 0.5066| 0.5644  |
| RSTR [15]| 0.5745| 0.6774| 0.5547| 0.4989| 0.6495| **0.5217**| 0.5795  |
| JPMM (Ours)| 0.5527| 0.6767| 0.5242| 0.5907| **0.6664**| 0.5070| **0.5863**|

The meta average mean F1-score for all the DA methods is shown in Table 3.

Table 3 The meta average mean F1-score for all the DA methods.

| Method | Mean F1-score |
|--------|---------------|
| SVM    | 0.4011±0.0477 |
| TCA [21]| 0.5623±0.0380 |
| GFK [22]| 0.5135±0.0218 |
| TKL [23]| 0.4911±0.0345 |
| SA [24]| 0.5331±0.0082 |
| TSRG [14]| 0.5476±0.0228 |
| DRLS [25]| 0.5599±0.0064 |
| RSTR [15]| 0.5746±0.0069 |
| JPMM (Ours)| **0.5763±0.0140** |

In this letter, we have investigated the problem of UDA in MER by proposing a novel subspace learning based on DA method called JPMM model. JPMM aims at learning a shared projection matrix guided by a joint of the first and second-order moment matching regularization for both source and target micro-expression samples such that the feature difference existing in original feature space can be alleviated after feature transformation. Moreover, we also take into consideration of the facial local region contribution during the feature transformation. Extensive cross-domain MER experiments are conducted to evaluate the performance of the DA methods using two different spatial-temporal descriptors in dealing with cross-domain MER. The best result in each experiment is highlighted in bold.

4. Conclusion

In this letter, we have investigated the problem of UDA in MER by proposing a novel subspace learning based on DA method called JPMM model. JPMM aims at learning a shared projection matrix guided by a joint of the first and second-order moment matching regularization for both source and target micro-expression samples such that the feature difference existing in original feature space can be alleviated after feature transformation. Moreover, we also take into consideration of the facial local region contribution during the feature transformation. Extensive cross-domain MER experiments are conducted to evaluate the performance of the DA methods using two different spatial-temporal descriptors in dealing with cross-domain MER. The best result in each experiment is highlighted in bold.
formance of the proposed JPMM, which demonstrates its effectiveness and superior performance over recent subspace learning based on DA methods.

References

[1] P. Ekman, “Darwin, deception, and facial expression,” Annals of the New York Academy of Sciences, vol.1000, no.1, pp.205–221, 2003.
[2] D. Jeong, B.-G. Kim, and S.-Y. Dong, “Deep joint spatiotemporal network (djsn) for efficient facial expression recognition,” Sensors, vol.20, no.7, p.1936, 2020.
[3] I.P. Adegun and H.B. Vadapalli, “Facial micro-expression recognition: a machine learning approach,” Scientific African, vol.8, p.e00465, 2020.
[4] J.-H. Kim, B.-G. Kim, P.P. Roy, and D.-M. Jeong, “Efficient facial expression recognition algorithm based on hierarchical deep neural network structure,” IEEE access, vol.7, pp.41273–41285, 2019.
[5] W.-J. Yan, Q. Wu, J. Liang, Y.-H. Chen, and X. Fu, “How fast are the leaked facial expressions: The duration of micro-expressions,” Journal of Nonverbal Behavior, vol.37, no.4, pp.217–230, 2013.
[6] M. Frank, M. Herbasz, K. Sinuk, A. Keller, and C. Nolan, “I see how you feel: Training laypeople and professionals to recognize fleeting emotions,” The Annual Meeting of the International Communication Association, New York City, 2009.
[7] G. Zhao and M. Pietikainen, “Dynamic texture recognition using local binary patterns with an application to facial expressions,” IEEE Trans. Pattern Anal. Mach. Intell., vol.29, no.6, pp.915–928, 2007.
[8] T. Pfister, X. Li, G. Zhao, and M. Pietikainen, “Recognising spontaneous facial micro-expressions,” 2011 international conference on computer vision, pp.1449–1456, IEEE, 2011.
[9] X. Li, X. Hong, A. Moilanen, X. Huang, T. Pfister, G. Zhao, and M. Pietikainen, “Towards reading hidden emotions: A comparative study of spontaneous micro-expression spotting and recognition methods,” IEEE transactions on affective computing, vol.9, no.4, pp.563–577, 2017.
[10] Y.J. Liu, B. J. Li, and Y.K. Lai, “Sparse MDMO: Learning a discriminative feature for micro-expression recognition,” IEEE Transactions on Affective Computing, vol.12, no.1, pp.254–261, 2018.
[11] B. Song, K. Li, Y. Zong, J. Zhu, W. Zheng, J. Shi, and L. Zhao, “Recognizing spontaneous micro-expression using a three-stream convolutional neural network,” IEEE Access, vol.7, pp.184537–184551, 2019.
[12] Z. Xia, X. Hong, X. Gao, X. Feng, and G. Zhao, “Spatiotemporal recurrent convolutional networks for recognizing spontaneous micro-expressions,” IEEE Trans. Multimedia, vol.22, no.3, pp.626–640, 2019.
[13] Z. Xia, W. Peng, H.-Q. Khor, X. Feng, and G. Zhao, “Revealing the invisible with model and data shrinking for composite-database micro-expression recognition,” IEEE Trans. Image Process., vol.29, pp.8590–8605, 2020.
[14] Y. Zong, X. Huang, W. Zheng, Z. Cui, and G. Zhao, “Learning a target sample re-generator for cross-database micro-expression recognition,” Proc. 25th ACM international conference on Multimedia, pp.872–880, ACM, 2017.
[15] T. Zhang, Y. Zong, W. Zheng, C.P. Chen, X. Hong, C. Tang, Z. Cui, and G. Zhao, “Cross-database micro-expression recognition: A benchmark,” IEEE Trans. Knowl. Data Eng., 2020.
[16] R. Ekman, What the Face Reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS), Oxford University Press, USA, 1997.
[17] Z. Zhang, M. Wang, Y. Huang, and A. Nehorai, “Aligning infinite-dimensional covariance matrices in reproducing kernel hilbert spaces for domain adaptation,” Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp.3437–3445, 2018.
[18] X. Peng, Q. Bai, X. Xia, Z. Huang, K. Saenko, and B. Wang, “Moment matching for multi-source domain adaptation,” Proc. IEEE/CVF International Conference on Computer Vision, pp.1406–1415, 2019.
[19] Z. Lin, R. Liu, and Z. Su, “Linearized alternating direction method with adaptive penalty for low-rank representation,” Proc. 24th International Conference on Neural Information Processing Systems, pp.612–620, 2011.
[20] W. Zheng, Y. Zong, X. Zhou, and M. Xin, “Cross-domain color facial expression recognition using transductive transfer subspace learning,” IEEE transactions on Affective Computing, vol.9, no.1, pp.21–37, 2018.
[21] S.J. Pan, I.W. Tsang, J.T. Kwok, and Q. Yang, “Domain adaptation via transfer component analysis,” IEEE Trans. Neural Netw., vol.22, no.2, pp.199–210, 2010.
[22] B. Gong, Y. Shi, F. Sha, and K. Grauman, “Geodesic flow kernel for unsupervised domain adaptation,” 2012 IEEE Conference on Computer Vision and Pattern Recognition, pp.2066–2073, IEEE, 2012.
[23] M. Long, J. Wang, J. Sun, and S.Y. Philip, “Domain invariant transfer kernel learning,” IEEE Trans. Knowl. Data Eng., vol.27, no.6, pp.1519–1532, 2014.
[24] B. Fernando, A. Habrard, M. Sebban, and T. Tuytelaars, “Unsupervised visual domain adaptation using subspace alignment,” Proc. IEEE international conference on computer vision, pp.2960–2967, 2013.
[25] Y. Zong, W. Zheng, X. Huang, J. Shi, Z. Cui, and G. Zhao, “Domain regeneration for cross-database micro-expression recognition,” IEEE Trans. Image Process., vol.27, no.5, pp.2484–2498, 2018.
[26] W.-J. Yan, X. Li, S.-J. Wang, G. Zhao, Y.-J. Liu, Y.-H. Chen, and X. Fu, “Casme ii: An improved spontaneous micro-expression database and the baseline evaluation,” PloS one, vol.9, no.1, p.e86041, 2014.
[27] X. Li, T. Pfister, X. Huang, G. Zhao, and M. Pietikäinen, “A spontaneous micro-expression database: Inducement, collection and baseline,” 2013 10th IEEE International Conference and Workshops on Automatic face and gesture recognition (fg), pp.1–6, IEEE, 2013.
[28] Y. Wang, J. See, R.C.-W. Phan, and Y.-H. Oh, “Efficient spatio-temporal local binary patterns for spontaneous facial micro-expression recognition,” PloS one, vol.10, no.5, p.e0124674, 2015.
[29] Z. Zhou, G. Zhao, and M. Pietikäinen, “Towards a practical lipreading system,” CVPR 2011, pp.137–144, IEEE, 2011.
[30] C.-C. Chang and C.-J. Lin, “Libsvm: A library for support vector machines,” ACM transactions on intelligent systems and technology (TIST), vol.2, no.3, p.27, 2011.