

Article

TKRM: Learning a Transfer Kernel Regression Model for Cross-Database Micro-Expression Recognition

Zixuan Chen ¹, Cheng Lu ², Feng Zhou ³  and Yuan Zong ^{2,*} ¹ School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China² School of Biological Science and Medical Engineering, Southeast University, Nanjing 210096, China³ College of Information Engineering, Yancheng Institute of Technology, Yancheng 224051, China

* Correspondence: xhzyuan@seu.edu.cn

Abstract: Cross-database micro-expression recognition (MER) is a more challenging task than the conventional one because its labeled training (source) and unlabeled testing (target) micro-expression (ME) samples are from different databases. In this circumstance, a large feature-distribution gap may exist between the source and target ME samples due to the different sample sources, which decreases the recognition performance of existing MER methods. In this paper, we focus on this challenging task by proposing a simple yet effective method called the transfer kernel regression model (TKRM). The basic idea of TKRM is to find an ME-discriminative, database-invariant and common reproduced kernel Hilbert space (RKHS) to bridge MEs belonging to different databases. For this purpose, TKRM has the ME discriminative ability of learning a kernel mapping operator to generate an RKHS and build the relationship between the kernelized ME features and labels in such RKHS. Meanwhile, an additional novel regularization term called target sample reconstruction (TSR) is also designed to benefit kernel mapping operator learning by improving the database-invariant ability of TKRM while preserving the ME-discriminative one. To evaluate the proposed TKRM method, we carried out extensive cross-database MER experiments on widely used micro-expression databases, including CASME II and SMIC. Experimental results obtained proved that the proposed TKRM method is indeed superior to recent state-of-the-art domain adaptation methods for cross-database MER.



Citation: Chen, Z.; Lu, C.; Zhou, F.; Zong, Y. TKRM: Learning a Transfer Kernel Regression Model for Cross-Database Micro-Expression Recognition. *Mathematics* **2023**, *11*, 918. <https://doi.org/10.3390/math11040918>

Keywords: cross-database micro-expression recognition; micro-expression recognition; probability distribution gap; kernel learning; domain adaptation

MSC: 68T10; 68T09

Academic Editors: Andrea Prati, Luis Javier García Villalba and Vincent A. Cicirello

Received: 21 January 2023

Revised: 9 February 2023

Accepted: 9 February 2023

Published: 11 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Micro-expression (ME) is defined as a set of brief, subtle and unconscious facial movements, whose duration boundary is often within half a second [1–3]. It has been demonstrated to be an important cue to reveal people's hidden intentions and, hence, it can be widely applied in police interrogation [4], national and public security [5], psychology and healthcare [6], and deceit analysis [7]. However, unlike conventional facial expressions, it is very difficult for individuals to recognize MEs even once they have gained professional training due to the brief and subtle characteristics of MEs. Therefore, it is of value to investigate the problem of automatic ME recognition (MER) to enable the computer to understand the emotion states from micro-expression video clips, helping people better understand MEs.

Recently, researchers have made great progress in the research of MER. For example, lots of publicly available ME databases, e.g., CASME series [8–11], SMIC series [12,13], and SAMM [14], have been released, which support the advancement of MER research. In order to effectively describe MEs, researchers have also designed a number of hand-crafted spatiotemporal descriptors, e.g., local binary pattern from three orthogonal planes

(LBP-TOP) [12,15], spatial temporal LBP with integral projection (STLBP-IP) [16], and a histogram of image gradient-TOP (HIGO-TOP) [17], as well as feature-enhancement methods [18]. Recently, inspired by the success of deep neural networks in computer vision tasks, deep learning methods [19–21] have also been used to learn discriminative features for recognizing MEs.

Although MER research has experienced remarkable progress, it is also worth mentioning that most existing MER methods that perform well were designed and evaluated under an ideal assumption that the training and test ME samples both belong to the same database. In other words, these ME samples were often recorded by the same camera and under the same environment; hence, their features can be viewed as abiding by the same probability distributions. In practical applications, however, this assumption may easily be broken, which leads to the feature distribution mismatch between the training and testing ME samples. Hence, the performance of existing MER methods would be significantly degraded. To address this issue, several researchers have shifted their focus to a more challenging but interesting MER task, i.e., cross-database MER [22], in which the training and testing ME samples derive from different databases. Recently, numerous methods have been proposed to manage this MER task and achieve promising performance, e.g., data augmentation [23], transfer subspace learning [22], and deep transfer learning methods [24].

In this paper, we also focus on this challenging but interesting topic and try to approach it from the angle of kernel learning. By utilizing the nonlinear mapping ability of kernel learning, we propose a novel transfer subspace learning method called the transfer kernel regression model (TKRM) to cope with the cross-database MER tasks, an overview of which is shown in Figure 1. In TKRM, we first build a simple and straightforward kernelized regression model to seek a reproduced kernel Hilbert space (RKHS) to build the relationship between the source ME features and label vectors. Meanwhile, a novel regularization term called target sample reconstruction (TSR) is designed to further enhance the TKRM model so that it is robust to the database variance existing between the source and target ME databases and also preserves the ME's discriminative ability. Hence, the proposed TKRM model is also applicable in the task of accurately predicting the ME labels of target samples.

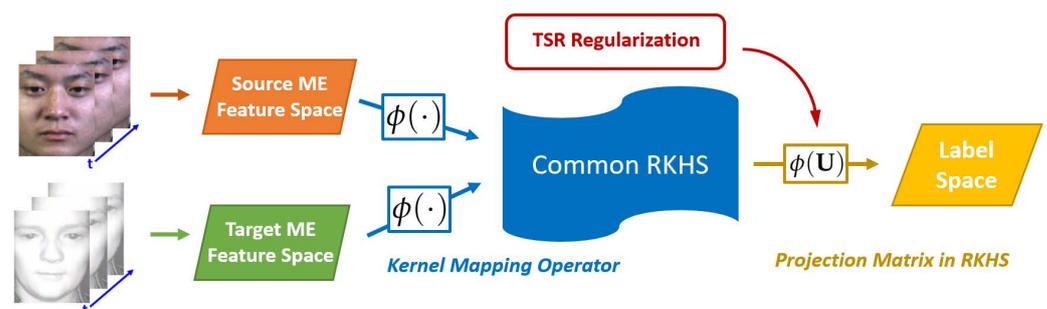


Figure 1. Basic Idea of the Proposed TKRM Method for Cross-Database MER Problem.

In summary, the major contributions of this paper are threefold:

1. We propose a novel transfer subspace learning model called TKRM to deal with the cross-database MER problem. The major advantage of the proposed TKRM is seeking a common reproduced kernel Hilbert space for the distribution alignment between the ME samples from different databases.
2. So that TKRM possesses the database-invariant ability, a well-designed regularization term called TSR is designed, which allows any target ME sample to be reconstructed by a few source ones.
3. A large number of comprehensive experiments are conducted to compare the proposed TKRM method and recent state-of-the-art transfer subspace learning methods. The results verify the effectiveness and superiority of TKRM in coping with the cross-database MER problem.

The rest of the manuscript is organized as follows: Section 2 reviews the recent progress in cross-database MER methods, including both transfer subspace learning and deep transfer learning methods. Section 3 provides the details of the proposed TKRM model, including how to build, optimize, and use the TKRM for cross-database MER tasks. In Section 4, we describe the extensive cross-database MER experiments that were conducted to evaluate the proposed TKRM model and we also provide in-depth discussion. Finally, the conclusions and directions for future work are provided in Section 5.

2. Related Work

The earliest work of cross-database MER may be traced to [22], in which Zong et al. designed extensive cross-database MER tasks based on two widely used spontaneous ME databases, including CASME II [9] and SMIC [13]. They also proposed a straightforward transfer subspace learning method called target sample re-generator (TSRG) to remove the feature distribution mismatch between the training (source) and testing (target) ME samples. Subsequently, Zong et al. [25] further extended TSRG to a more generalized framework called domain regeneration (DR). Recently, in order to advance the research of cross-database MER and provide a standard platform to evaluate emerging methods, a cross-database MER benchmark was presented in the works of [23,26]. Subsequently, researchers have shifted their focus to achieving better performance in the tasks of cross-database MER. To achieve this goal, some researchers continue to develop the transfer subspace learning methods by considering the ME-aware discriminative clues. For example, in the works of [26–28], the authors attempted to learn more discriminative features from the facial local region cues to better describe MEs, while considering the database-invariant performance.

On the other hand, several researchers have tried to manage the problem of cross-database MER using deep transfer learning methods. For example, Xia et al. [24] designed a style aggregated and attention transfer framework (SA-AT) consisting of two stages enabling the ResNet model learned from the source ME samples to also be applicable to target ME. Li et al. [29] proposed a novel deep domain adaptation method called the discriminative region transfer network (DRTN), which uses an adversarial-based adaptation structure and attention mechanism to learn the database-invariant ME features from discriminative facial regions. More recently, Song et al. [30] designed a dual-stream convolutional neural network (DSCNN) to cope with cross-database MER tasks. The major novelty of DSCNN is that the temporal dynamic and salient cues in ME samples are simultaneously considered to ensure that the network can learn more discriminative and robust features.

3. TKRM for Cross-Database MER

In this section, we first formulate the TKRM model in detail. Then, we design an efficient algorithm to learn the optimal parameters of the proposed TKRM model. Finally, we present how to use the TKRM model to predict the target ME samples that are different from the source ones.

3.1. Formulation of TKRM

Assume that we are given a labeled source ME database and their corresponding feature matrix is denoted by $\mathbf{X}_s \in \mathbb{R}^{d \times n_s}$, where d is the dimension of the feature vector used for describing ME, and n_s is the source ME sample number. Let $\mathbf{L}_s \in \mathbb{R}^{c \times n_s}$ be the ME label matrix, where c is the number of micro-expression categories. The i th column of \mathbf{L}_s , denoted by \mathbf{l}_s^i , is a c -dimensional vector whose j th entry is set as 1, while the others are fixed at 0 if this ME sample's label is j th ME. Accordingly, we are also given an unlabeled target ME database, whose feature matrix is denoted by $\mathbf{X}_t \in \mathbb{R}^{d \times n_t}$, where n_t is the target sample number.

As described previously, the basic idea of the proposed TKRM is seeking a common RKHS for both source and target ME samples by resorting to a kernel mapping operator to bridge them such that their feature distribution mismatch can be removed, which is shown in Figure 1. Following this idea, we first build the relationship between the source ME

features and their corresponding label vectors via a simple kernel regression model, whose optimization problem can be formulated as follows:

$$\min_{\phi(\mathbf{U})} \|\mathbf{L}_s - \phi(\mathbf{U})^T \phi(\mathbf{X}_s)\|_F^2, \tag{1}$$

where $\phi(\cdot)$ is a kernel mapping operator, and $\phi(\mathbf{U})$ and $\phi(\mathbf{X}_s)$ denote the projection matrix and source ME features in the reproduced kernel Hilbert space (RKHS) generated by $\phi(\cdot)$, respectively. From Equation (1), it is clear to see that the proposed TKRM can own the ME discriminative ability and we can easily predict the ME category of the source ME feature vector based on the projection matrix $\phi(\mathbf{U})$ in the RKHS.

Meanwhile, according to Figure 1, the $\phi(\mathbf{U})$ learned by our TKRM model should also be applicable to predicting the target ME samples' labels. In other words, the RKHS should be a common space for target ME samples as well as the source ones. Unfortunately, in the case of cross-database MER, a large feature distribution gap exists between the source and target ME databases. Therefore, we hope to find a common RKHS, in which the feature distributions of the source and target domains are close, i.e., samples of two domains can be bridged by a linear combination. Since the source data have emotion labels, reconstructing each target sample with source data can not only accurately supervise the emotion prediction of the target samples, but also effectively align the feature distributions between two domains. To this end, we also map the target ME features onto the same RKHS, which can be denoted by $\phi(\mathbf{X}_t)$. Then, we design a novel regularization term called TSR for the TKRM model to help modify the RKHS so that it is also suitable for the target ME samples. The objective function of TSR can be formulated as follows:

$$f(\phi(\mathbf{U})) = \|\phi(\mathbf{U})^T \phi(\mathbf{X}_t) - \phi(\mathbf{U})^T \phi(\mathbf{X}_s) \mathbf{W}\|_F^2 + \lambda \sum_{i=1}^{n_t} \|\mathbf{w}_i\|_1, \tag{2}$$

where $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_{n_t}] \in n_s \times n_t$ is a coefficient matrix, $\|\cdot\|_1$ denotes the L_1 norm of a vector and λ is a trade-off parameter to control the balance between the L_1 norm and reconstruction terms. Note that, by minimizing the objective function of TSR in Equation (2), the learned projection matrix $\phi(\mathbf{U})$ in the common RKHS forces the target ME samples to be a linear combination of several source ME ones, which indicates the projected target ME samples will be assimilated into the source ones. Furthermore, the sparse property of the L_1 regularization term also effectively reduces the dimension of the parameter space to avoid over-fitting in the optimization process. Therefore, it is a good choice to jointly minimize the objective function in Equation (1) and the one shown in Equation (2); in this way, we will arrive at the optimization problem of the proposed TKRM, which is expressed as follows:

$$\min_{\phi(\mathbf{U}), \mathbf{W}} \|\mathbf{L}_s - \phi(\mathbf{U})^T \phi(\mathbf{X}_s)\|_F^2 + \lambda_1 \|\phi(\mathbf{U})^T \phi(\mathbf{X}_t) - \phi(\mathbf{U})^T \phi(\mathbf{X}_s) \mathbf{W}\|_F^2 + \lambda_2 \sum_{i=1}^{n_t} \|\mathbf{w}_i\|_1, \tag{3}$$

where λ_1 and $\lambda_2 = \lambda_1 \times \lambda$ are the trade-off parameters controlling the balance among three terms in the objective function of TKRM.

However, it should be pointed out that it is difficult to solve the current version of the optimization problem for TKRM because the row dimension of the kernel mapping operator $\phi(\mathbf{U})$ is infinite. Using the kernel trick [31], the kernel function can be used to convert the optimization problem to be a solvable one. Specifically, let $\phi(\mathbf{U}) = [\phi(\mathbf{X}_s), \phi(\mathbf{X}_t)]\mathbf{P}$, where $\mathbf{P} \in \mathbb{R}^{(n_s+n_t) \times c}$ is a reconstruction of the coefficient matrix used to enforce $\phi(\mathbf{U})$ to be linearly combined by the source ME features $\phi(\mathbf{X}_s)$ and target ME features $\phi(\mathbf{X}_t)$ in the RKHS. By substituting this equation into Equation (3), we are able to arrive at the new formulation of the optimization problem for TKRM, as follows:

$$\min_{\mathbf{P}, \mathbf{W}} \|\mathbf{L}_s - \mathbf{P}^T \mathbf{K}_s\|_F^2 + \lambda_1 \|\mathbf{P}^T \mathbf{K}_t - \mathbf{P}^T \mathbf{K}_s \mathbf{W}\|_F^2 + \lambda_2 \sum_{i=1}^{n_t} \|\mathbf{w}_i\|_1, \tag{4}$$

where $\mathbf{K}_s = [\mathbf{K}_{ss}^T, \mathbf{K}_{ts}^T]^T$ and $\mathbf{K}_t = [\mathbf{K}_{st}^T, \mathbf{K}_{tt}^T]^T$. Among them, $\mathbf{K}_{ss} = \phi(\mathbf{X}_s)^T \phi(\mathbf{X}_s)$, $\mathbf{K}_{ts} = \phi(\mathbf{X}_t)^T \phi(\mathbf{X}_s)$, $\mathbf{K}_{st} = \phi(\mathbf{X}_s)^T \phi(\mathbf{X}_t)$, and $\mathbf{K}_{tt} = \phi(\mathbf{X}_t)^T \phi(\mathbf{X}_t)$, and can be directly calculated by the kernel function, e.g., Gaussian, polynomial, and linear ones.

In addition, we would like to further regularize the model parameter matrix \mathbf{P} of TKRM by adding an L_1 norm term with respect to \mathbf{P} to the objective function, which results in a sparse column vector for \mathbf{P} . This was inspired by the works of [18,23], in which a sparse coefficient matrix led to a more effective reconstruction in the RKHS and helped to avoid overfitting for optimization. Consequently, the eventual optimization problem of the proposed TKRM model can be written as follows:

$$\min_{\mathbf{P}, \mathbf{W}} \|\mathbf{L}_s - \mathbf{P}^T \mathbf{K}_s\|_F^2 + \lambda_1 \|\mathbf{P}^T \mathbf{K}_t - \mathbf{P}^T \mathbf{K}_s \mathbf{W}\|_F^2 + \lambda_2 \sum_{i=1}^{n_t} \|\mathbf{w}_i\|_1 + \lambda_3 \|\mathbf{P}\|_1. \tag{5}$$

Note that similar to λ_1 and λ_2 , λ_3 is a trade-off parameter as well. $\|\mathbf{P}\|_1 = \sum_{i=1}^c \|\mathbf{p}_i\|_1$ and \mathbf{p}_i is the i th column of \mathbf{P} .

3.2. Optimization of TKRM

We would like to use the alternated direction method (ADM) [32] to learn the optimal solution of the TKRM model, i.e., alternatively updating \mathbf{P} and \mathbf{W} until convergence. Specifically, we initialize \mathbf{W} and repeat the following three major steps:

1. Fix \mathbf{W} and update \mathbf{P} . In this step, the original optimization problem will be reduced to:

$$\min_{\mathbf{P}} \|\mathbf{L}_s - \mathbf{P}^T \mathbf{K}_s\|_F^2 + \lambda_1 \|\mathbf{P}^T \Delta \mathbf{K}_{st}\|_F^2 + \lambda_3 \|\mathbf{P}\|_1, \tag{6}$$

where $\Delta \mathbf{K}_{st} = \mathbf{K}_t - \mathbf{K}_s \mathbf{W}$. The inexact augmented Lagrangian multiplier (IALM) [33] method can be adopted to efficiently learn the optimal solution. Concretely, we first introduce an additional variable \mathbf{Q} satisfying $\mathbf{P} = \mathbf{Q}$. Then, the original optimization problem of TKRM, which is an unconstrained one, can be converted to a constrained one as follows:

$$\min_{\mathbf{P}, \mathbf{Q}} \|\mathbf{L}_s - \mathbf{Q}^T \mathbf{K}_s\|_F^2 + \lambda_1 \|\mathbf{Q}^T \Delta \mathbf{K}_{st}\|_F^2 + \lambda_3 \|\mathbf{P}\|_1, \text{ s.t. } \mathbf{P} = \mathbf{Q}. \tag{7}$$

Subsequently, the Lagrangian function for the above optimization problem can be written as follows:

$$\mathcal{L}(\mathbf{P}, \mathbf{Q}, \mathbf{W}, \mathbf{T}, \mu) = \|\mathbf{L}_s - \mathbf{Q}^T \mathbf{K}_s\|_F^2 + \lambda_1 \|\mathbf{Q}^T \Delta \mathbf{K}_{st}\|_F^2 + \lambda_3 \|\mathbf{P}\|_1 + \mathbf{T}^T (\mathbf{P} - \mathbf{Q}) + \frac{\mu}{2} \|\mathbf{P} - \mathbf{Q}\|_F^2, \tag{8}$$

where \mathbf{T} is the Lagrangian multiplier matrix and μ is the trade-off parameter. By alternatively minimizing the Lagrangian function with respect to the variables until convergence, the optimal \mathbf{P} will be obtained. The detailed updating procedures are as follows:

- (a) Fix \mathbf{P} , \mathbf{T} , and μ , and update \mathbf{Q} ; the optimization problem with respect to \mathbf{Q} can be reformulated as:

$$\min_{\mathbf{Q}} \|\mathbf{L}_s - \mathbf{Q}^T \mathbf{K}_s\|_F^2 + \lambda_1 \|\mathbf{Q}^T \Delta \mathbf{K}_{st}\|_F^2 + \mathbf{T}^T (\mathbf{P} - \mathbf{Q}) + \frac{\mu}{2} \|\mathbf{P} - \mathbf{Q}\|_F^2, \tag{9}$$

which has a closed-form solution, as follows:

$$\mathbf{Q} = (2\mathbf{K}_s \mathbf{K}_s^T + \lambda_1 \Delta \mathbf{K}_{st} \Delta \mathbf{K}_{st}^T + \mu \mathbf{I})^{-1} (2\mathbf{K}_s \mathbf{L}_s^T + \mathbf{T} + \mu \mathbf{P}), \tag{10}$$

where \mathbf{I} is an $(n_s + n_t) \times (n_s + n_t)$ identity matrix.

- (b) Fix \mathbf{Q} , \mathbf{T} , and μ , and update \mathbf{P} . In this step, the optimization problem with respect to \mathbf{P} is:

$$\min_{\mathbf{P}} \lambda_3 \|\mathbf{P}\|_1 + \mathbf{T}^T (\mathbf{P} - \mathbf{Q}) + \frac{\mu}{2} \|\mathbf{P} - \mathbf{Q}\|_F^2, \tag{11}$$

which can be rewritten as the following formulation:

$$\min_{\mathbf{P}} \frac{\lambda_3}{\mu} \|\mathbf{P}\|_1 + \frac{1}{2} \|\mathbf{P} - (\mathbf{Q} - \frac{\mathbf{T}}{\mu})\|_F^2. \tag{12}$$

Referring to Step 2 of optimization procedures in [22], the solution of Equation (12) can be determined according to the following criterion, i.e.,

- (1) $P_{ij} = (P_{ij} - \frac{Q_{ij}}{\mu}) - \frac{\lambda_3}{\mu}$, if $P_{ij} - \frac{Q_{ij}}{\mu} > \frac{\lambda_3}{\mu}$,
- (2) $P_{ij} = (P_{ij} - \frac{Q_{ij}}{\mu}) + \frac{\lambda_3}{\mu}$, if $P_{ij} - \frac{Q_{ij}}{\mu} < \frac{\lambda_3}{\mu}$,
- (3) $P_{ij} = 0$, otherwise,

where P_{ij} , Q_{ij} , and T_{ij} are the entry in the i th row and j th column of matrices \mathbf{P} , \mathbf{Q} , and \mathbf{T} , respectively.

- (c) Update \mathbf{T} and μ : $\mathbf{T} = \mathbf{T} + \mu(\mathbf{P} - \mathbf{Q})$ and $\mu = \min\{\rho\mu, \mu_{max}\}$, where ρ is a scaled parameter greater than 1 and μ_{max} is a preset maximal value for μ .
 - (d) Check convergence: $\|\mathbf{P} - \mathbf{Q}\|_F < \epsilon$ or the interaction reaches maximal value $Iter_{max}$.
2. Fix \mathbf{P} and update \mathbf{W} : In this step, the original optimization problem is:

$$\min_{\mathbf{w}_i} \|\mathbf{P}^T \mathbf{k}_t^i - \mathbf{P}^T \mathbf{K}_s \mathbf{w}_i\|_F^2 + \frac{\lambda_2}{\lambda_1} \|\mathbf{w}_i\|_1, \tag{13}$$

where \mathbf{k}_t^i is the i th column of \mathbf{K}_t corresponding to the kernel function of the i th target ME sample. This is a standard Lasso problem and can be efficiently solved by a number of typical algorithms, e.g., coordinate descent.

- 3. Check convergence or reach the preset maximal iterations.

3.3. Convergence Analysis for the Optimization of TKRM

As Section 3.2 shows, we proposed an efficient iterative algorithm to learn the optimal solutions of the proposed TKRM. In this section, we would like to analyze the convergence of this algorithm from a mathematical angle. Specifically, it is clear to see that the optimization problem for TKRM in Equation (5), whose objective function is denoted by $f(\mathbf{P}, \mathbf{W})$, is divided into two minimization subproblems, i.e., Equations (6) and (13), whose objective functions are denoted by $f(\mathbf{P}, \hat{\mathbf{W}})$ and $f(\hat{\mathbf{P}}, \mathbf{W})$, in the proposed optimization algorithm, where $\hat{\mathbf{W}}$ and $\hat{\mathbf{P}}$ are constants learned by their corresponding minimization subproblems in a previous round of optimization. Then, since all the terms in the objective function of TKRM are L_1 and Frobenius norms, which are continuous and lower-bounded, the objective function value of TKRM in Equation (5), $f(\mathbf{P}, \mathbf{W})$, would sequentially decrease if these two minimization subproblems, Equations (6) and (13), were iteratively optimized. Consequently, the convergence of the proposed iterative algorithm for learning the optimal solutions of the proposed TKRM can be guaranteed by the following inequation:

$$f(\hat{\mathbf{P}}_t, \hat{\mathbf{W}}_t) < f(\hat{\mathbf{P}}_t, \hat{\mathbf{W}}_{t-1}) < f(\hat{\mathbf{P}}_{t-1}, \hat{\mathbf{W}}_{t-1}), \tag{14}$$

where t denotes the updating step index in the optimization algorithm.

This inequation holds obviously in the proposed iterative algorithm, which can be supported by the following two conclusions drawn from the divided minimization subproblem: (1) given a constant $\hat{\mathbf{W}}$, the minimization subproblem in Equation (6) would undoubtedly result in $f(\hat{\mathbf{P}}_t, \hat{\mathbf{W}}) < f(\hat{\mathbf{P}}_{t-1}, \hat{\mathbf{W}})$ at the t th step of the proposed iterative

algorithm; (2) similarly, based on the minimization subproblem in Equation (13), we can also obtain that $f(\hat{\mathbf{P}}, \hat{\mathbf{W}}_t) < f(\hat{\mathbf{P}}, \hat{\mathbf{W}}_{t-1})$ for a constant $\hat{\mathbf{P}}$ at the t th step.

3.4. Target Sample's ME Label Prediction

Suppose the optimal solution of the TKRM model is denoted by $\hat{\mathbf{P}}$. Then, we can conveniently predict the ME labels of the target samples, even though the source of these samples is quite different from that of the training samples. Let \mathbf{x}_t^i be a target ME sample's feature vector. Thus, we can first calculate its ME label vector, which is obtained with the following optimization problem:

$$\min_{\mathbf{I}_t^i} \|\mathbf{1}_t^i - \hat{\mathbf{P}}^T \mathbf{k}_t^i\|^2, \tag{15}$$

where $\mathbf{k}_t^i = [\phi(\mathbf{x}_t^i)^T \phi(\mathbf{X}_s), \phi(\mathbf{x}_t^i)^T \phi(\mathbf{X}_t)]^T$. It is noted that the inner product in the RKHS associated with \mathbf{k}_t^i must be calculated using the same kernel function as the one chosen in the training stage. Equation (15) is a quadratic programming (QP) problem and can be easily solved with a number of QP algorithms, including the interior point method.

Finally, according to the optimal label vector \mathbf{I}_t^i , the eventual ME label would be determined as the j th ME label if the j th entry of \mathbf{I}_t^i has a maximal value among all the entries.

4. Experiments

4.1. Micro-Expression Databases and Experiment Setting

In this section, we adopt the CASME II [9] and SMIC [13] micro-expression databases to design cross-database MER tasks to evaluate the proposed TKRM model. The SMIC database includes recordings with high speed (HS), normal visual (VIS), and near-infrared (NIR) cameras, resulting in three subsets, i.e., HS, VIS, and NIR. It is the first publicly available ME database containing diverse ME samples. In all the subsets, the ME samples are divided into three different ME types including *Positive*, *Negative*, and *Surprise*. Similarly to the HS subset in SMIC, CASME II is also recorded using a high-speed camera. Each sample of CASME II is assigned one of five ME labels, i.e., *Happy*, *Disgust*, *Surprise*, *Regression*, and *Others*. Figure 2 provides an example of the above four ME sample sets, which demonstrates the sample difference existing among these databases.

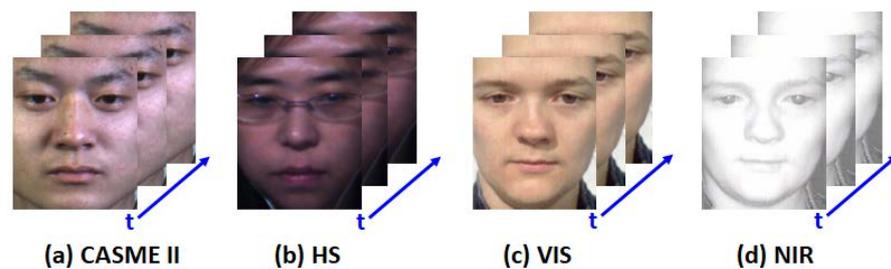


Figure 2. Sample Examples of Four ME Sample Sets Used in the Evaluation Experiments.

By using any two of the above four types of ME samples to alternatively serve as the source and target, we are able to obtain six different cross-database MER tasks. We denote these tasks as $C \rightarrow H, H \rightarrow C, C \rightarrow V, V \rightarrow C, C \rightarrow N$, and $N \rightarrow C$, respectively, where C, H, V , and N are abbreviations for CASME II, HS, VIS, and NIR. Since the source and target samples in cross-database MER require consistent ME labels, we select and relabel parts of samples from CASME II. Specifically, the *Happy* samples are given the *Positive* label, the *Disgust* and *Regression* samples are relabeled as *Negative*, and the label for *Surprise* samples remains unchanged. Table 1 shows the new statistics of the relabeled CASME II and SMIC databases used in our experiments.

Table 1. Sample Statistics of the CASME II and SMIC databases used in the designed cross-database MER experiments.

ME Database	Negative	Positive	Surprise	Summation
CASME II (C)	73	32	25	130
HS (H)	70	51	43	164
VIS (V)	28	23	20	71
NIR (N)	28	23	20	71

4.2. Comparison Methods and Performance Metrics

To see whether the proposed TKRM method performs well in coping with cross-database MER tasks, we choose seven state-of-the-art transfer subspace learning methods, including transfer component analysis (TCA) [34], geodesic flow kernel (GFK) [35], subspace alignment (SA) [36], transfer kernel learning (TKL) [37], target sample re-generator (TSRG) [22], domain regeneration in the original feature space with unchanged target domain (DRFS-T) [25], domain regeneration in the label space (DRLS) [25], and region selective transfer regression (RSTR) [26]. To analyze the experimental results, *mean F1-score* and *Accuracy* are used as the performance metrics. Among them, *mean F1-score* is the primary metric and calculated by $\frac{1}{c} \sum_{i=1}^c F_1^{(i)}$, where $F_1^{(i)}$ is the F1 score associated with *i*th predicted ME samples. As for the other metric, *Accuracy*, it is defined as the correct predictions divided by the total testing sample number.

Note that unlike end-to-end deep learning methods, subspace learning methods require the features to describe MEs; hence, we choose uniform LBP-TOP with the parameters $R = 3$ and $P = 8$ to serve as the spatiotemporal descriptors to extract the features of MEs in the experiments. In addition, a multi-scale spatial division scheme, including four types of grids, i.e., 1×1 , 2×2 , 4×4 , and 8×8 , is used to divide the facial image into 85 facial local blocks before extracting the LBP-TOP. Hence, a supervector consisting of 85 LBP-TOP feature vectors is eventually served to describe each ME sample. As for the parameter setting of transfer subspace learning methods, we follow the benchmark evaluation work in [23,26] and directly take their results for comparison because our experiment setting is identical to theirs. Regarding our TKRM, the kernel function is set as linear. Three trade-off parameters need also to be set for our TKRM. Still following the experiment setting of [23,26], we search for them from a preset parameter interval, i.e., [0.1:0.1:1,2:1:10], and then report the best results when it reaches the best *mean F1-score* for each experiment.

4.3. Experimental Results and Discussion

Tables 2 and 3 summarize the *mean F1-score* and *Accuracy* for our KDAL and seven comparison transfer subspace learning methods. Apart from these, it can be seen that we also directly use the SVM classifier to perform the same cross-database MER experiments to serve as the baselines. As Tables 2 and 3 show, it is easily found that nearly all the transfer subspace learning methods outperformed the baseline method, i.e., SVM without any knowledge transfer operation, which indicates that transfer learning is indeed an effective method to improve the model robustness to the database variance and provides a possible direction to advance the research of MER.

Second, it is noted that the proposed TKRM achieves the best performance in terms of both *mean F1-score* and *Accuracy* averaged among all six experiments. By further observing the detailed results of six experiments, it can also be obtained that our TKRM also achieved the best *mean F1-score* and *Accuracy* in three tasks including $H \rightarrow C$, $C \rightarrow V$, and $N \rightarrow C$. These observations indicates that our TKRM can achieve a more promising and satisfactory performance than recent state-of-the art transfer subspace learning methods for cross-database MER tasks.

Finally, it can also be noted that nearly all the transfer subspace learning methods performed better in the experiments between CASME II and VIS than the others. For example, we can find that seven methods, including our TKRM, achieved a *mean F1-score*

of over 60.00% in at least one task ($C \rightarrow V, V \rightarrow C$, or both). However, they do not all have improved performance in the four remaining tasks. Therefore, it can be concluded that the experiments between CASME II and HS, or between CASME II and NIR, are more challenging cross-database MER tasks. Nevertheless, our TKRM can still achieve a *mean F1-score* of over 50.00% for all four tasks, which demonstrates the superior performance of TKRM over all comparison methods.

Table 2. The results of cross-database MER experiments in terms of the mean F1-score (%) among CASME II, HS, VIS, and NIR, respectively. The best result in each experiment is underlined.

Method	C→H	H→C	C→V	V→C	C→N	N→C	Average
SVM	36.97	32.45	47.01	53.67	52.95	23.68	41.12
TCA	46.37	48.70	68.34	57.89	49.92	39.37	51.77
GFK	41.26	47.76	63.61	60.56	51.80	44.69	51.61
SA	43.02	54.47	59.39	52.43	<u>57.38</u>	35.92	48.77
TKL	38.29	46.61	60.42	53.78	53.92	42.48	49.25
TSRG	50.42	51.71	59.35	62.08	56.24	41.05	53.48
DRLS-T	45.24	54.60	62.17	67.62	53.69	46.53	54.98
DRLS	49.24	52.67	57.57	59.42	48.85	38.38	51.02
RSTR	<u>52.97</u>	56.22	58.82	<u>70.21</u>	50.09	46.93	55.87
TKRM	52.53	<u>59.41</u>	<u>70.20</u>	65.31	50.29	<u>51.35</u>	<u>58.18</u>

Table 3. The results of cross-database MER experiments in terms of accuracy (%) among CASME II, HS, VIS, and NIR, respectively. The best result in each experiment is underlined.

Method	C→H	H→C	C→V	V→C	C→N	N→C	Average
SVM	45.12	48.46	50.70	53.08	42.11	23.85	45.55
TCA	46.34	53.08	69.01	59.23	50.70	42.31	50.73
GFK	46.95	50.77	66.20	61.50	53.52	46.92	54.31
SA	47.56	62.31	59.15	51.54	47.89	36.92	50.90
TKL	44.51	54.62	60.56	53.08	54.93	43.85	51.93
TSRG	51.83	60.77	59.15	63.08	<u>56.34</u>	46.15	56.22
DRLS-T	46.95	60.00	63.38	68.46	<u>56.34</u>	50.77	57.65
DRLS	53.05	59.23	57.75	60.00	49.83	42.37	53.71
RSTR	<u>54.27</u>	60.77	59.15	<u>70.77</u>	50.70	50.77	57.74
TKRM	52.44	<u>66.15</u>	<u>70.42</u>	65.38	50.70	<u>52.31</u>	<u>59.57</u>

4.4. Performance of TKRM with Different Kernel Functions

In the above experiments, our TKRM only uses the linear kernel function and achieved promising performance in cross-database MER tasks. It should be pointed out that the nonlinear kernel function is also suitable for TKRM. To see how the kernel function affects the performance of TKRM, we further carry out several additional experiments by choosing different nonlinear kernel functions for TKRM. Specifically, we choose three tasks as the representatives and set the kernel function as the polynomial, i.e., $f(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^T \mathbf{y})^d$, PolyPlus, i.e., $f(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^T \mathbf{y} + 1)^d$, and Gaussian, i.e., $f(\mathbf{x}, \mathbf{y}) = e^{-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{2t}}$, where $f(\cdot)$ is the kernel function, \mathbf{x} and \mathbf{y} are the feature vectors, and d is the dimension of feature vectors. For polynomial and PolyPlus functions, their parameter d is set to 1.01, 1.05 and 1.1, respectively. For the Gaussian, its parameter t is fixed at 1, 5 and 10, respectively. Table 4 depicts the results of the experiments, which enables us to obtain several interesting observations and conclusions. It is clearly observed that, for most tasks, the linear-kernel-based TKRM obtained the best results in terms of both *mean F1-score* and *Accuracy*. However, it is still noted that the polynomial kernel with the parameter $d = 1.01$ achieved the highest *mean F1-score*, reaching 60.02. This indicates that the performance of the proposed TKRM with a suitable nonlinear kernel would achieve a further increase. Nevertheless, it should be pointed out that the performance of the nonlinear version of TKRM varies sensitively

with respect to the change in the kernel function's type and parameter, which can be demonstrated by the comparison between the results achieved by the polynomial and Gaussian functions. Consequently, how to choose the kernel function and search for its optimal parameter in the KDAL model to improve its performance when coping with cross-database MER tasks is still an open question. This point should be investigated in the future.

Table 4. The results of cross-database MER experiments by using TKRM with different kernel functions. The best result in each experiment is underlined.

Kernel Function	C→H	H→C	C→V
Linear	<u>52.53/52.44</u>	59.41/66.15	<u>70.20/70.42</u>
Polynomial ($d = 1.01$)	<u>52.53/52.44</u>	60.02/65.38	58.89/69.01
Polynomial ($d = 1.05$)	51.99/51.83	58.46/60.77	67.31/67.61
Polynomial ($d = 1.1$)	<u>52.53/52.44</u>	59.39/62.31	63.89/63.38
PolyPlus ($d = 1.01$)	52.53/52.44	60.02/65.38	68.89/69.01
PolyPlus ($d = 1.05$)	51.99/51.83	58.46/60.77	67.35/67.61
PolyPlus ($d = 1.1$)	51.85/51.83	59.39/62.31	<u>70.20/70.42</u>
Gaussian ($t = 1$)	37.93/40.24	47.14/57.69	51.25/50.70
Gaussian ($t = 5$)	50.61/50.61	57.22/57.60	67.65/67.61
Gaussian ($t = 10$)	49.45/49.35	51.74/52.31	66.05/66.20

5. Conclusions and Future Work

In this paper, we propose a new transfer kernel regression model (TKRM), which makes full use of the kernel learning's powerful nonlinear mapping ability, to deal with the cross-database MER problem. The major advantage of TKRM is minimizing the probability distributions divergence between source and target ME samples in the kernelized feature space with an implicit method of feature reconstruction. Moreover, extensive experiments were carried out to evaluate the effectiveness and superiority of the proposed TKRM. The experimental results proved that our TKRM method can achieved improved performance compared with most recent state-of-the-art transfer subspace learning methods for the cross-database MER problem. In addition, it should be noted that the proposed TKRM method is still not satisfied because its performance in challenging tasks, e.g., tasks between SMIC (NIS) and CASME II, only reaches a low-level mean F1-score of about 50%. Consequently, the ME discriminative ability of TKRM needs to be improved so that the cross-database MER problem can be better addressed. In the future, we will further enhance the proposed TKRM model by absorbing discriminative information learned from the ME-aware cues.

Author Contributions: Conceptualization, methodology, and writing—original draft preparation, Z.C. and C.L.; funding acquisition, software and validation, formal analysis, investigation, resources, data curation, and writing—review and editing, F.Z. and Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by the Major Basic Research Project of the Natural Science Foundation of the Jiangsu Higher Education Institutions under the Grant 19KJA110002, in part by the Natural Science Foundation of China under the Grants 61902064 and 61673108, in part by the Zhishan Young Scholarship of Southeast University, and in part by the Yancheng Institute of Technology High-level Talent Research Initiation Project under the Grant XJR2022001.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: [CASME II] [<http://casme.psych.ac.cn/casme/e2>] (accessed on 12 September 2022) and [SMIC] [<http://www.cse.oulu.fi/SMICDatabase>] (accessed on 12 September 2022).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Oh, Y.H.; See, J.; Le Ngo, A.C.; Phan, R.C.W.; Baskaran, V.M. A survey of automatic facial micro-expression analysis: Databases, methods, and challenges. *Front. Psychol.* **2018**, *9*, 1128. [[CrossRef](#)] [[PubMed](#)]
2. Ben, X.; Ren, Y.; Zhang, J.; Wang, S.J.; Kpalma, K.; Meng, W.; Liu, Y.J. Video-based facial micro-expression analysis: A survey of datasets, features and algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 5826–5846. [[CrossRef](#)]
3. Ekman, P.; Friesen, W.V. *Facial Action Coding System*; Consulting Psychologist Press: Palo Alto, CA, USA, 1978.
4. O'Sullivan, M.; Frank, M.G.; Hurley, C.M.; Tiwana, J. Police lie detection accuracy: The effect of lie scenario. *Law Hum. Behav.* **2009**, *33*, 530. [[CrossRef](#)] [[PubMed](#)]
5. Frank, M.G.; Maccario, C.J.; Govindaraju, V. Behavior and security. In *Protecting Airline Passengers in the Age of Terrorism*; ABC-CLIO: Santa Barbara, CA, USA, 2009; pp. 86–106.
6. Frank, M.; Herbasz, M.; Sinuk, K.; Keller, A.; Nolan, C. I see how you feel: Training laypeople and professionals to recognize fleeting emotions. In *Annual Meeting of the International Communication Association*; Sheraton New York: New York, NY, USA, 2009; pp. 1–35.
7. Ekman, P. Lie catching and microexpressions. *Philos. Decept.* **2009**, *1*, 5.
8. Yan, W.J.; Wu, Q.; Liu, Y.J.; Wang, S.J.; Fu, X. CASME database: A dataset of spontaneous micro-expressions collected from neutralized faces. In Proceedings of the 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), Shanghai, China, 22–26 April 2013.
9. Yan, W.J.; Li, X.; Wang, S.J.; Zhao, G.; Liu, Y.J.; Chen, Y.H.; Fu, X. CASME II: An improved spontaneous micro-expression database and the baseline evaluation. *PLoS ONE* **2014**, *9*, e86041. [[CrossRef](#)] [[PubMed](#)]
10. Qu, F.; Wang, S.J.; Yan, W.J.; Li, H.; Wu, S.; Fu, X. CAS (ME) 2: A database for spontaneous macro-expression and micro-expression spotting and recognition. *IEEE Trans. Affect. Comput.* **2017**, *9*, 424–436. [[CrossRef](#)]
11. Li, J.; Dong, Z.; Lu, S.; Wang, S.J.; Yan, W.J.; Ma, Y.; Liu, Y.; Huang, C.; Fu, X. CAS (ME) 3: A Third Generation Facial Spontaneous Micro-Expression Database with Depth Information and High Ecological Validity. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 2782–2800. [[CrossRef](#)] [[PubMed](#)]
12. Pfister, T.; Li, X.; Zhao, G.; Pietikäinen, M. Recognising spontaneous facial micro-expressions. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 1449–1456.
13. Li, X.; Pfister, T.; Huang, X.; Zhao, G.; Pietikäinen, M. A spontaneous micro-expression database: Inducement, collection and baseline. In Proceedings of the 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), Shanghai, China, 22–26 April 2013.
14. Davison, A.K.; Lansley, C.; Costen, N.; Tan, K.; Yap, M.H. Sann: A spontaneous micro-facial movement dataset. *IEEE Trans. Affect. Comput.* **2016**, *9*, 116–129. [[CrossRef](#)]
15. Zhao, G.; Pietikäinen, M. Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE Trans. Pattern Anal. Mach. Intell.* **2007**, *29*, 915–928. [[CrossRef](#)]
16. Huang, X.; Wang, S.J.; Zhao, G.; Pietikäinen, M. Facial micro-expression recognition using spatiotemporal local binary pattern with integral projection. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Santiago, Chile, 7–13 December 2015; pp. 1–9.
17. Li, X.; Hong, X.; Moilanen, A.; Huang, X.; Pfister, T.; Zhao, G.; Pietikäinen, M. Towards reading hidden emotions: A comparative study of spontaneous micro-expression spotting and recognition methods. *IEEE Trans. Affect. Comput.* **2017**, *9*, 563–577. [[CrossRef](#)]
18. Zong, Y.; Huang, X.; Zheng, W.; Cui, Z.; Zhao, G. Learning from hierarchical spatiotemporal descriptors for micro-expression recognition. *IEEE Trans. Multimed.* **2018**, *20*, 3160–3172. [[CrossRef](#)]
19. Xia, Z.; Hong, X.; Gao, X.; Feng, X.; Zhao, G. Spatiotemporal recurrent convolutional networks for recognizing spontaneous micro-expressions. *IEEE Trans. Multimed.* **2019**, *22*, 626–640. [[CrossRef](#)]
20. Wei, M.; Zheng, W.; Jiang, X.; Zong, Y.; Lu, C.; Liu, J. A Novel Magnification-Robust Network with Sparse Self-Attention for Micro-expression Recognition. In Proceedings of the 2022 26th International Conference on Pattern Recognition (ICPR), Montreal, QC, Canada, 21–25 August 2022; pp. 1120–1126.
21. Zhu, J.; Zong, Y.; Chang, H.; Xiao, Y.; Zhao, L. A Sparse-Based Transformer Network with Associated Spatiotemporal Feature for Micro-Expression Recognition. *IEEE Signal Process. Lett.* **2022**, *29*, 2073–2077. [[CrossRef](#)]
22. Zong, Y.; Huang, X.; Zheng, W.; Cui, Z.; Zhao, G. Learning a target sample re-generator for cross-database micro-expression recognition. In Proceedings of the 25th ACM International Conference on Multimedia, Mountain View, CA, USA, 23–27 October 2017; pp. 872–880.
23. Zong, Y.; Zheng, W.; Hong, X.; Tang, C.; Cui, Z.; Zhao, G. Cross-database micro-expression recognition: A benchmark. In Proceedings of the 2019 International Conference on Multimedia Retrieval, Ottawa, ON, Canada, 10–13 June 2019; pp. 354–363.
24. Xia, Z.; Liang, H.; Hong, X.; Feng, X. Cross-database micro-expression recognition with deep convolutional networks. In Proceedings of the 2019 3rd International Conference on Biometric Engineering and Applications, Stockholm, Sweden, 29–31 May 2019; pp. 56–60.
25. Zong, Y.; Zheng, W.; Huang, X.; Shi, J.; Cui, Z.; Zhao, G. Domain regeneration for cross-database micro-expression recognition. *IEEE Trans. Image Process.* **2018**, *27*, 2484–2498. [[CrossRef](#)]
26. Zhang, T.; Zong, Y.; Zheng, W.; Chen, C.P.; Hong, X.; Tang, C.; Cui, Z.; Zhao, G. Cross-Database Micro-Expression Recognition: A Benchmark. *IEEE Trans. Knowl. Data Eng.* **2022**, *34*, 544–559. [[CrossRef](#)]

27. Zong, Y.; Zheng, W.; Cui, Z.; Zhao, G.; Hu, B. Toward bridging microexpressions from different domains. *IEEE Trans. Cybern.* **2020**, *50*, 5047–5060. [[CrossRef](#)]
28. Jiang, X.; Zong, Y.; Zheng, W.; Liu, J.; Wei, M. Seeking Salient Facial Regions for Cross-Database Micro-Expression Recognition. In Proceedings of the 2022 26th International Conference on Pattern Recognition (ICPR), Montreal, QC, Canada, 21–25 August 2022; pp. 1019–1025.
29. Li, J.; Hu, R.; Mukherjee, M. Discriminative Region Transfer Network for Cross-Database Micro-Expression Recognition. In Proceedings of the ICC 2022-IEEE International Conference on Communications, Seoul, Republic of Korea, 16–20 May 2022; pp. 5082–5087.
30. Song, B.; Zong, Y.; Li, K.; Zhu, J.; Shi, J.; Zhao, L. Cross-Database Micro-Expression Recognition Based on a Dual-Stream Convolutional Neural Network. *IEEE Access* **2022**, *10*, 66227–66237. [[CrossRef](#)]
31. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: Berlin, Germany, 1999.
32. Zheng, W. Multi-view facial expression recognition based on group sparse reduced-rank regression. *IEEE Trans. Affect. Comput.* **2014**, *5*, 71–85. [[CrossRef](#)]
33. Lin, Z.; Liu, R.; Su, Z. Linearized alternating direction method with adaptive penalty for low-rank representation. *Adv. Neural Inf. Process. Syst.* **2011**, *24*, 1–9.
34. Pan, S.J.; Tsang, I.W.; Kwok, J.T.; Yang, Q. Domain adaptation via transfer component analysis. *IEEE Trans. Neural Netw.* **2010**, *22*, 199–210. [[CrossRef](#)]
35. Gong, B.; Shi, Y.; Sha, F.; Grauman, K. Geodesic flow kernel for unsupervised domain adaptation. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 2066–2073.
36. Fernando, B.; Habrard, A.; Sebban, M.; Tuytelaars, T. Unsupervised visual domain adaptation using subspace alignment. In Proceedings of the IEEE International Conference on Computer Vision, Sydney, Australia, 1–8 December 2013; pp. 2960–2967.
37. Long, M.; Wang, J.; Sun, J.; Philip, S.Y. Domain invariant transfer kernel learning. *IEEE Trans. Knowl. Data Eng.* **2014**, *27*, 1519–1532. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.