

Article

Collaborative Representation Using Non-Negative Samples for Image Classification

Jianhang Zhou and Bob Zhang * 

PAMI Research Group, Department of Computer and Information Science, Faculty of Science and Technology, University of Macau, Taipa, Macau 999078, China; mb85405@um.edu.mo

* Correspondence: bobzhang@um.edu.mo; Tel.: +853-8822-4425

Received: 6 May 2019; Accepted: 4 June 2019; Published: 8 June 2019



Abstract: Collaborative representation based classification (CRC) is an efficient classifier in image classification. By using l_2 regularization, the collaborative representation based classifier holds competitive performances compared with the sparse representation based classifier using less computational time. However, each of the elements calculated from the training samples are utilized for representation without selection, which can lead to poor performances in some classification tasks. To resolve this issue, in this paper, we propose a novel collaborative representation by directly using non-negative representations to represent a test sample collaboratively, termed Non-negative Collaborative Representation-based Classifier (NCRC). To collect all non-negative collaborative representations, we introduce a Rectified Linear Unit (ReLU) function to perform filtering on the coefficients obtained by l_2 minimization according to CRC's objective function. Next, we represent the test sample by using a linear combination of these representations. Lastly, the nearest subspace classifier is used to perform classification on the test samples. The experiments performed on four different databases including face and palmprint showed the promising results of the proposed method. Accuracy comparisons with other state-of-art sparse representation-based classifiers demonstrated the effectiveness of NCRC at image classification. In addition, the proposed NCRC consumes less computational time, further illustrating the efficiency of NCRC.

Keywords: collaborative representation-based classification; non-negative samples; image classification

1. Introduction

Image classification techniques have been extensively researched in computer vision [1–5]. Among them, sparse representation based classification methods [6] and its variants are frequently proposed and refined due to their effectiveness and efficiency, especially in face recognition [7–9]. Rather than using sparse representation (SR), the collaborative representation based classifier was proposed by using collaborative representation (CR), which achieved competitive performances with higher efficiency. The main difference between sparse representation and collaborative representation is the usage of different regularization terms in the minimization formulation, which is l_1 norm for sparse representation and l_2 norm for collaborative representation. Many applications have shown that both methods provide good results in image classification [5,10–13], where they can be further improved for a better recognition performance.

To improve the recognition rate, many works focus on weighting the training samples in different ways. For example, Xu et al. proposed a two-phase test sample sparse representation method [7] by using sparse representation in the first phase, followed by representation based on previously exploited neighbors of the test sample in the second phase. Timofte et al. imposed weights on the coefficients of collaborative representation [14] and achieved better performances in face recognition. Similarly, Fan et al. [8] provided another weights-imposing method, which derives weights of each coefficient

from the corresponding training sample by calculating the Gaussian kernel distance between each training sample and test sample. In [7,8,14], the authors intended to seek a weight that can truly help in the classification, indicating that using different training samples may influence the discriminative pattern. Moreover, the negative coefficient implies a negative correlation between the training sample and test sample. Inspired by this idea, and considering the relationship between negative coefficients in the training sample and test samples, we intend to represent the test sample using non-negative representation. Ultimately, this produces a more effective classification [15–17].

The CRC uses l_2 norm to perform classification. As representation coefficients can be derived from an analytic solution to a least square problem, CRC is much more efficient than SRC. Furthermore, the collaborative representation is interpretive as well [18]. Due to its intrinsic property, CRC has been extensively refined. For instance, Dong et al. used sparse subspace on weighted CRC [14] to improve the recognition rate for face recognition [19]. Zeng et al. proposed S**CRC* to achieve promising performances by fusing coefficients from sparse representation with the coefficients from collaborative representation [20]. However, in this work, each test sample is represented twice using SRC and CRC simultaneously, which is time consuming. Zheng et al. selected k candidate classes before representing a test sample collaboratively, while the p in the objective function should be defined in advance to obtain optimal results [21]. In [14,18–21], the authors ignored the relationship between the collaborative coefficient and the test sample. However, we strongly believe that, by considering this relationship, we are able to obtain a higher recognition rate.

The Restricted Linear Unit (ReLU) function [22] is widely used in deep learning to improve the performance of a deep neural network. Until now, many popular deep neural networks use ReLU as the activation function (e.g., VGG19 [3]). By using the ReLU function, there is sparse activation, making the network sparser. Similarly, we can apply the ReLU function to enhance the sparsity of CRC, which should improve its performance. Since the sparsity can help the CRC model to perform robust classification [23], representations mapped by the ReLU function will also achieve promising classification results.

Bringing everything together, this paper proposes a novel collaborative representation based classification method, named Non-negative CRC (NCRC). The system architecture of NCRC can be found in Figure 1. In the first step, NCRC performs normalization on all samples from a dataset based on l_2 norm. Next, NCRC uses all training samples to calculate the collaborative representation coefficients by representing the test samples collaboratively via l_2 regularization. Afterwards, the ReLU function is utilized to filter these collaborative coefficients and map the negative ones to zero. Then, we use the newly mapped collaborative coefficients to represent the test sample. Finally, to classify each test sample, the nearest subspace classification is performed. Specifically, the residuals between the representation of each class and the test sample are calculated, and we select the class label associated with the minimum residual as the result. The main contributes are:

1. We propose a novel image classification algorithm using non-negative samples based on the collaborative representation based classifier.
2. The proposed method enhances the sparsity of CRC by introducing the Restricted Linear Unit (ReLU) function, which increases the sparsity of the coefficients and improves the recognition rate.

The remainder of this paper is organized as follows. In Section 2, we give a brief overview of the collaborative representation based classifier before introducing our proposed NCRC. In Section 3, we show extensive experimental results to demonstrate the effectiveness and efficiency of NCRC, while at the same time discussing the superiority of our method. Finally, in Section 4, we give a conclusion.

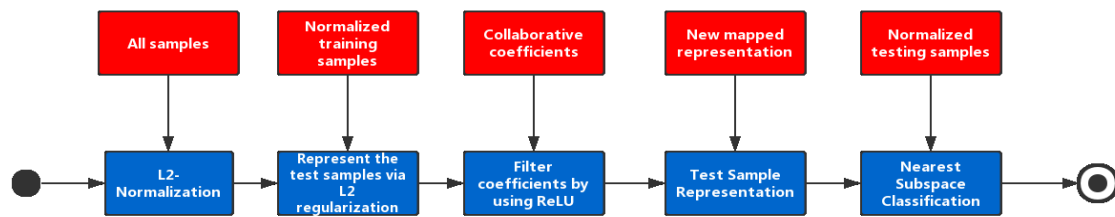


Figure 1. The pipeline of NCRC. In the first step, l_2 normalization is performed on all samples. Next, we calculate the collaborative representation coefficients by representing the test samples collaboratively via l_2 regularization using all training samples. In the third step, the ReLU function is utilized to filter collaborative coefficients and to map the negative ones to zero. Afterwards, we use the newly mapped collaborative coefficients to represent the test sample. In the last step, the nearest subspace classification is performed to classify each test sample.

2. CRC and Non-Negative CRC

2.1. Collaborative Representation-Based Classifier

Firstly, we review the background knowledge of the collaborative representation-based classifier. The collaborative representation based classifier [24] has been widely used in image classification, especially for face recognition [25–27]. Given M classes of samples and denote $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N]$ as the dataset, we calculate the collaborative coefficients as follows:

$$\hat{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha}} \|\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_2^2 \quad (1)$$

where $\boldsymbol{\alpha}$ is the collaborative coefficient vector, \mathbf{y} is the test sample, and λ is the regularization term. According to Equation (1), we can obtain an analytic solution of $\hat{\boldsymbol{\alpha}}$ as follows:

$$\hat{\boldsymbol{\alpha}} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T \mathbf{y} \quad (2)$$

Lastly, the identity of test sample is determined by the minimum distance with a specific class subspace:

$$\text{identity}(\mathbf{y}) = \arg \min_i \|\mathbf{y} - \mathbf{X}_i \boldsymbol{\alpha}_i\|_2 \quad (3)$$

We summarize CRC in Algorithm 1.

Algorithm 1 Collaborative representation based classifier.

- 1: Normalize \mathbf{X} to have l_2 norm
- 2: Calculate the collaborative representation coefficients vector $\hat{\boldsymbol{\alpha}}$ using

$$\hat{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha}} \|\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_2^2$$

- 3: Calculate the the residuals between test sample \mathbf{y} and the representation of each class using formulation and obtain the identity of \mathbf{y} using Equation (3):

$$\text{identity}(\mathbf{y}) = \arg \min_i \|\mathbf{y} - \mathbf{X}_i \boldsymbol{\alpha}_i\|_2$$

2.2. Non-Negative Collaborative Representation Classifier

Based on CRC described in Section 2.1, we propose our methodology here. Since the collaborative representation based classifier uses positive as well as negative coefficients simultaneously to represent

a test sample. The distribution of coefficients is not sparse enough, which makes the residuals of each class less discriminative. The negative coefficient indicates negative correlation between the test sample and its representation. Accordingly, we propose a novel collaborative representation, which only uses the non-negative representation, namely Non-negative Collaborative Representation based Classifier (NCRC). The proposed NCRC first represents the test sample using l_2 norm as follows:

$$\mathbf{y} = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \dots + \alpha_n x_n \quad (4)$$

where $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]$ is obtained from Equation (1)

Then, we introduce the Rectified Linear Unit (ReLU) function to filter the coefficient vector α :

$$f(x) = \begin{cases} x & , x \geq 0 \\ 0 & , x < 0 \end{cases} \quad (5)$$

Afterwards, we represent the test sample using representation filtered by the ReLU function as follows:

$$\mathbf{y} = \tilde{\alpha}_1 x_1 + \tilde{\alpha}_2 x_2 + \tilde{\alpha}_3 x_3 + \dots + \tilde{\alpha}_n x_n \quad (6)$$

Finally, the identity of a test sample is determined by calculating the residual between the test sample and each class:

$$\text{identity}(\mathbf{y}) = \arg \min_i \|\mathbf{y} - \mathbf{X}_i \tilde{\alpha}_i\|_2 \quad (7)$$

We summarize the NCRC classification procedure in Algorithm 2.

Algorithm 2 Non-negative collaborative representation based classifier.

- 1: Normalize \mathbf{X} to have l_2 norm
- 2: Calculate the collaborative representation coefficients vector $\hat{\alpha}$ using formulation (1)

$$\hat{\alpha} = \arg \min_{\alpha} \|\mathbf{y} - \mathbf{X}\alpha\|_2^2 + \lambda \|\alpha\|_2^2$$

- 3: Use ReLU function described in Equation (5) to map each collaborative representations to non-negative representation

$$f(x) = \begin{cases} x & , x \geq 0 \\ 0 & , x < 0 \end{cases}$$

$$\tilde{\alpha} = f(\hat{\alpha})$$

- 4: Represent the test sample using non-negative representations described as Equation (6)

$$\mathbf{y} = \tilde{\alpha}_1 x_1 + \tilde{\alpha}_2 x_2 + \tilde{\alpha}_3 x_3 + \dots + \tilde{\alpha}_n x_n$$

- 5: Calculate the the residuals between test sample \mathbf{y} and representation of each class using Equation (7)

$$\text{identity}(\mathbf{y}) = \arg \min_i \|\mathbf{y} - \mathbf{X}_i \tilde{\alpha}_i\|_2$$

Figure 2 depicts the coefficients of each training sample (corresponding to the AR database; please refer to Section 3.1 for more information) from CRC (left) and NCRC (right), respectively. Since the test sample from the AR database belongs to the first class, it is obvious that training samples from the first class have a high positive value in this figure, indicating the high positive correlation with the test sample. It can also be observed from the CRC coefficient figure (left) that several coefficients still contain many negative values, which are samples having negative correlation with the test sample. Clearly, by using the ReLU filter, coefficients from NCRC (right) are much sparser than that

of CRC. Obviously, several regions are sparser in NCRC compared with CRC (237 coefficients from NCRC are 0). As the sparsity of collaborative representation can help the classifier to perform more robust classification [20,23], the recognition performance of NCRC is better than the original CRC (see Sections 3.2–3.5).

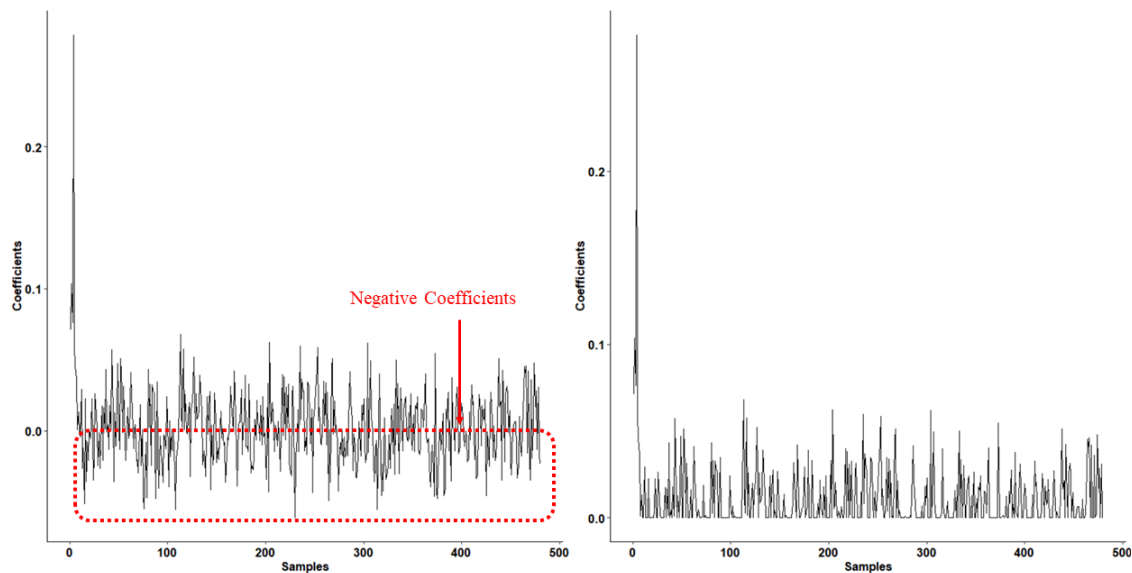


Figure 2. Coefficients of CRC (**left**) and NCRC (**right**) on an image from the AR database. Values on the horizontal axis represent the index of a sample in the training set. Values on the vertical axis represent the coefficient value. Samples (**left**) with negative (<0) coefficients are enclosed in a dotted red rectangle, indicating several coefficients still contain many negative values. However, these negative coefficients are non-existent on the right.

3. Experiments

We performed experiments on AR [28], LFW [29], MUCT [30] and the PolyU palmprint [31] datasets to verify the effectiveness of our method. To show our proposed method's capability in different image classification tasks, the recognition tasks ranged from human face recognition to palmprint recognition, where the performances were validated by the hand-out method [32]. For each dataset, we divided them into training and testing to evaluate the result in each iteration. Besides this, we increased the size of the training samples in each iteration. At last, we computed the average accuracy achieved by the proposed method and compared it with other state-of-the-art sparse representation based classifiers (S**CRC* [20] and Pro*CRC* [18]), the original SRC [4] and CRC [24] as well as traditional classifiers such as SVM and KNN. To achieve the optimal results for all classifiers, various parameters were tested. For SRC, CRC, S**CRC*, and Pro*CRC*, we used $\lambda = 1 \times e^{-3}$, $1 \times e^{-2}$, $1 \times e^{-1}$, $2 \times e^{-1}$, $3 \times e^{-1}$, and $4 \times e^{-1}$. In SVM, we tried different kernel functions, as for KNN ($K = 7$). All of the experiments were performed on a PC running Windows 10 with a 3.40 GHz CPU and 16 GB RAM running MATLAB R2018a. Below, for the comparison methods, we report the accuracy that was achieved using the optimal parameter. To guarantee the stability of our final results on each dataset, we repeated each experiment 30 times and took the average value as the final result.

3.1. Dataset Description

The AR face database [28] contains 4000 color images of 126 human faces, where each image is 768×576 pixels. Images of the same person are captured in two sessions separated by 14 days. In the experiments, we varied the number of training samples from 4 to 20 images per class, and took

remaining samples in each class as the testing samples. Examples of images from this database can be found in Figure 3.

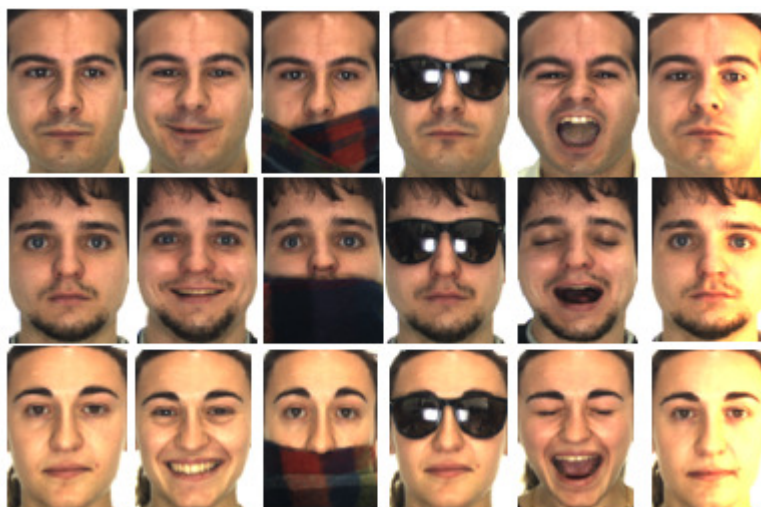


Figure 3. Image samples from the AR face database.

The LFW face database [29] contains 13,233 images of 5749 human faces. Among them, 1680 people had more than two images. Figure 4 depicts some examples from this database. For the experiments, each image consisted of 32×32 pixels. We used the number of training samples from 5 to 35 images per class, and employed the remaining samples in each class as the testing samples. In addition, we applied FaceNet to extract the features from the database before feeding it into the classifiers.

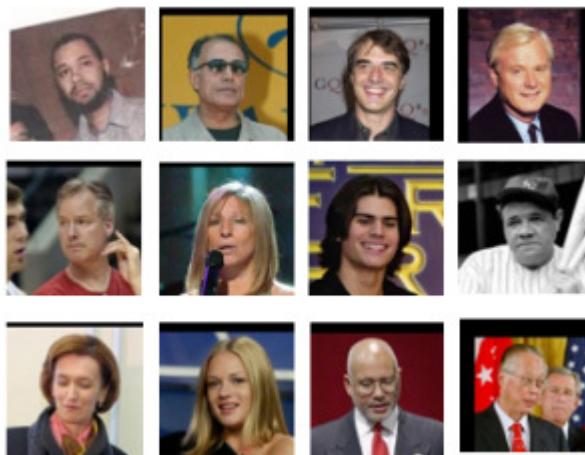


Figure 4. Image samples from the LFW face database.

The MUCT face [30] database contains in total 3755 face images collected from 276 people. Each image is given a size of 640×480 pixels. These face images were captured by a CCD camera and stored in 24-bit RGB format. Samples from this database are shown in Figure 5. In the experiments, we adjusted the number of training samples from 1 to 7 in each class and took the remaining samples in each class as the testing samples.

The PolyU palmprint database was created by the Biometric Research Centre, Hong Kong Polytechnic University [31]. There are 7752 gray-scale images of 386 different palms stored in BMP format. Figure 6 illustrates some examples from this database. For an individual's palmprint, there are approximately 20 samples collected in two sessions, where each palm image is 384×284 pixels.

For our experiments, the number of training samples ranged from 1 to 5 images per class, and we used the remaining samples in each class as the testing samples.



Figure 5. Image samples from the MUCT face database.

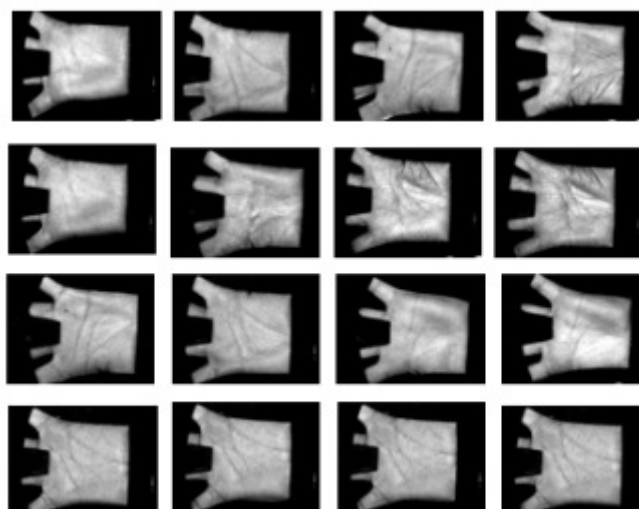


Figure 6. Image samples from the PolyU palmprint database.

The general information of the databases in the experiments is summarized in Table 1.

Table 1. Information of the databases.

Database	Classes	Samples	Image Size	Dimension
AR	120	3120	40 × 32	2-D
LFW	86	1251	32 × 32	2-D
MUCT	276	3755	640 × 480	2-D
PolyU Palmprint	386	7752	384 × 284	2-D

3.2. Experiments on the AR Face Database

The experimental results on the AR face database can be found in Table 2 and Figure 7, where it can be seen that NCRC outperformed other classifiers. The highest recognition rate is 93.06% when using 20 samples per class for training. Compared with CRC, the improvement of NCRC ranges from 0.84% to 2.02%, which shows NCRC enhanced the recognition ability of the original CRC classifier on

the AR database. Besides CRC, NCRC also achieved a better result than SRC (82.50%). Furthermore, the proposed method outperformed other variants of SRC and CRC such as S**CRC* (82.50%, $\lambda = 0.01$) and Pro*CRC* (91.81%, $\lambda = 0.01$), as well as traditional classifications including KNN ($K = 7$) and SVM (polynomial kernel function). For the parameter selection of NCRC, we experimented with $\lambda = 1 \times e^{-3}, 1 \times e^{-2}, 1 \times e^{-1}, 2 \times e^{-1}, 3 \times e^{-1}$, and $4 \times e^{-1}$, respectively (refer to Table 3) by fixing the number of training samples at 20 (which was used to achieve the highest recognition). According to Table 3, NCRC obtained the best accuracy when $\lambda = 0.01$. In Table 2, the highest accuracy achieved using each number of samples is marked in bold. In Table 3, highest accuracy achieved in each parameter λ is marked in bold.

Table 2. Recognition performance of NCRC and the comparison methods on the AR database.

Samples	SRC	CRC	S* <i>CRC</i>	Pro <i>CRC</i>	SVM	KNN	NCRC
4	63.94	66.59	68.40	67.72	56.47	45.57	67.73
8	62.36	71.06	70.28	70.19	53.33	50.83	72.78
12	72.44	75.18	74.34	77.62	45.95	53.45	77.20
16	86.83	91.83	85.91	90.08	51.25	64.50	92.67
20	82.50	91.94	82.50	91.81	54.86	65.00	93.06

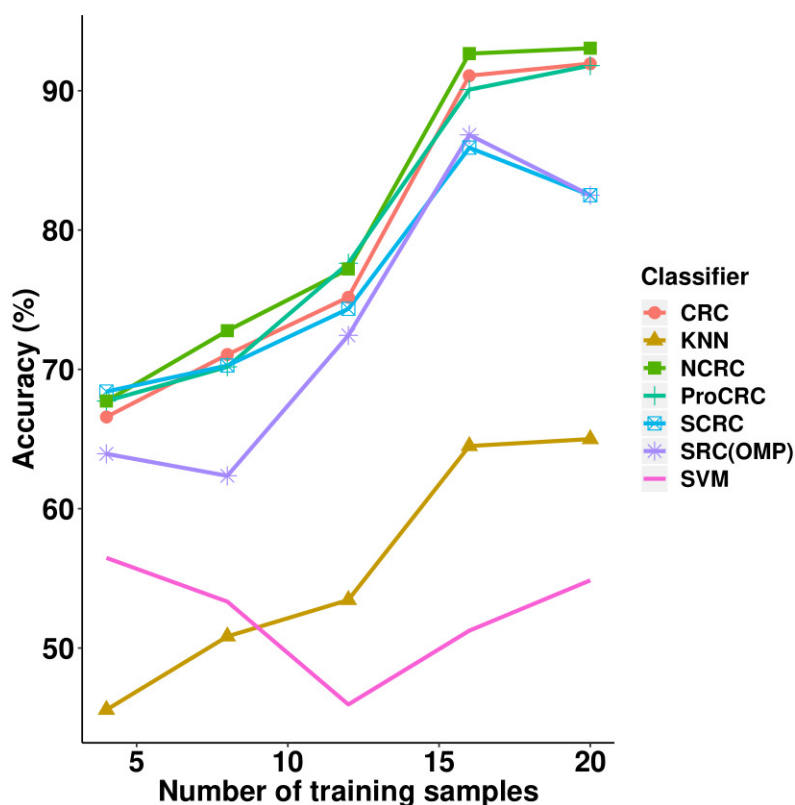


Figure 7. Accuracy vs. increasing the number of training samples on the AR database.

Table 3. Parameter selection of NCRC for the AR database.

λ	0.001	0.01	0.1	0.2	0.3	0.4
Accuracy	92.50	93.06	91.94	91.67	90.56	90.42

3.3. Experiments on the LFW Face Database

Next, we performed experiments on the LFW face database and compared its results with other classifiers including SRC, CRC, S**CRC*, Pro*CRC*, SVM and KNN. These results are illustrated in Table 4 and Figure 8, where it can be observed that NCRC performed best using 10–35 training samples.

When using five training samples, ProCRC slightly outperformed NCRC by 0.39%. For the parameter selection of NCRC, we experimented with $\lambda = 1 \times e^{-3}$, $1 \times e^{-2}$, $1 \times e^{-1}$, $2 \times e^{-1}$, $3 \times e^{-1}$, and $4 \times e^{-1}$ (refer to Table 5) by fixing the number of training samples at 35. The optimal result was obtained when $\lambda = 0.1$. Compared with CRC, the improvement of NCRC ranges from 0.17% to 3.62%, which shows NCRC again enhanced the recognition ability of the original CRC. The proposed method using 35 training samples also achieved a better result than SRC (38.02%). Furthermore, NCRC outperformed other variants of SRC and CRC such as S**CRC* (50.25%, $\lambda = 0.1$), ProCRC (51.66%, $\lambda = 0.01$), SVM (28.12%, polynomial kernel function), and KNN (27.29%, $K = 7$). In Table 4, the highest accuracy achieved using each number of samples is marked in bold. In Table 5, highest accuracy achieved in each parameter λ is marked in bold.

Table 4. Recognition performance of NCRC and the comparison methods on the LFW database.

Samples	SRC	CRC	S* <i>CRC</i>	ProCRC	SVM	KNN	NCRC
5	24.32	27.37	26.02	27.93	21.83	15.80	27.54
10	30.41	35.24	34.88	35.78	23.20	19.02	37.15
15	32.36	43.36	41.53	41.53	25.85	23.32	44.62
20	34.03	47.61	45.33	46.07	28.04	27.17	50.64
25	35.63	49.35	46.98	51.66	26.94	27.51	52.44
30	36.78	49.58	49.58	51.66	28.14	26.91	53.20
35	38.02	52.00	50.25	51.66	28.12	27.29	55.32

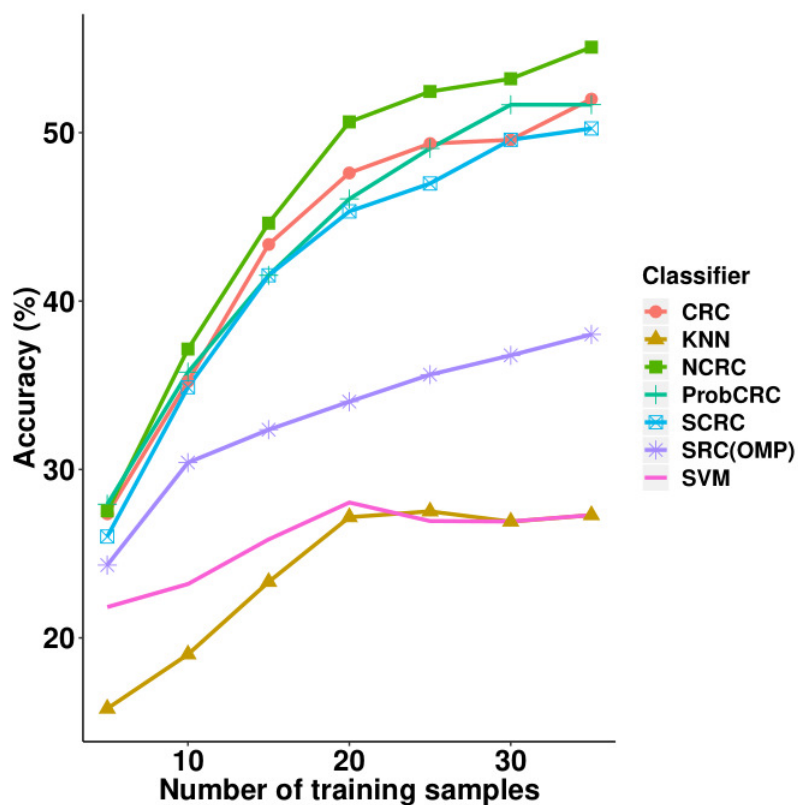


Figure 8. Accuracy vs. increasing the number of training samples on the LFW database.

Table 5. Parameter selection of NCRC on the LFW database.

λ	0.001	0.01	0.1	0.2	0.3	0.4
Accuracy	42.1	55.07	55.32	51.33	48.25	46.67

3.4. Experiments on the MUCT Face Database

Table 6 and Figure 9 present the experimental results on the MUCT face database. From the seven different training sample sizes, NCRC attained the highest accuracy using 2–7 samples compared with the other classifiers. Using the same parameter selection progress as presented in Sections 3.2 and 3.3 for NCRC, the best λ value using 7 training samples was 0.01 (refer to Table 7). The highest recognition rate is 77.78% when using seven samples per class for training. Compared with CRC, the improvement of NCRC ranges from 0.46% to 2.68%. Furthermore, NCRC also achieved a better result than SRC (77.07%), S**CRC* (76.85%, $\lambda = 0.01$) and Pro*CRC* (74.32%, $\lambda = 0.01$), as well as traditional classifications including KNN ($K = 7$) and SVM (polynomial kernel function). As for the use of one training sample, S**CRC* outperformed NCRC by only 0.58%. In Table 6, the highest accuracy achieved using each number of samples is marked in bold. In Table 7, highest accuracy achieved in each parameter λ is marked in bold.

Table 6. Recognition performance of NCRC and the comparison methods on the MUCT database.

Samples	SRC	CRC	S* <i>CRC</i>	Pro <i>CRC</i>	SVM	KNN	NCRC
1	26.58	29.26	30.30	29.26	20.06	5.92	29.72
2	36.84	41.21	42.00	41.21	21.91	10.90	43.65
3	42.77	45.75	46.67	45.75	20.53	12.70	46.84
4	54.55	55.60	58.66	55.60	20.11	15.50	58.28
5	67.58	68.76	70.99	66.95	20.59	18.65	71.16
6	74.23	73.89	74.46	72.84	24.54	24.58	75.99
7	77.07	75.86	76.85	74.32	25.56	27.87	77.78

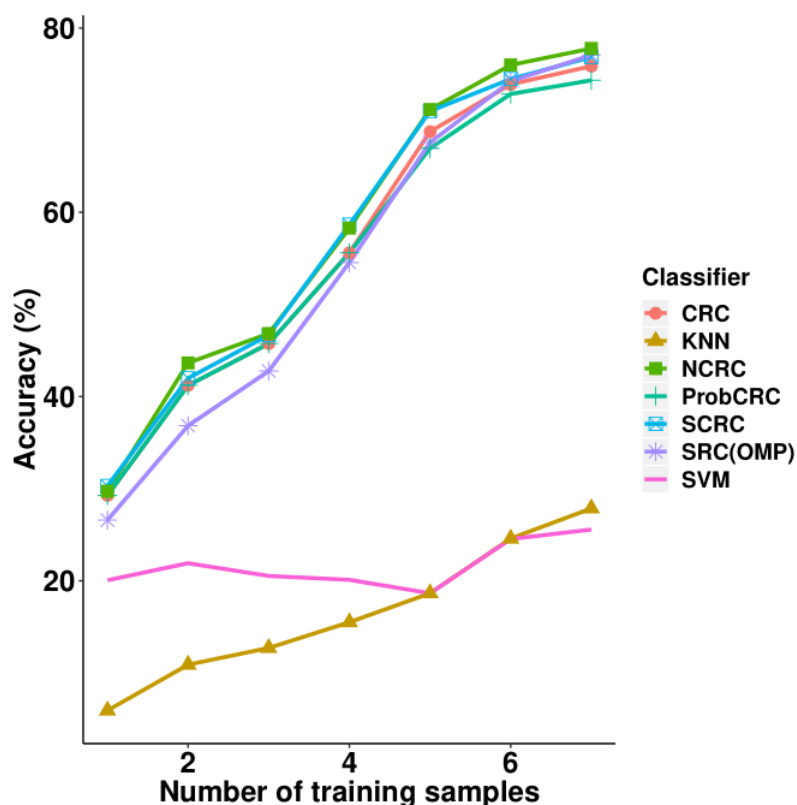


Figure 9. Accuracy vs. increasing the number of training samples on the MUCT database.

Table 7. Parameter selection of NCRC on the MUCT database.

λ	0.001	0.01	0.1	0.2	0.3	0.4
Accuracy	74.49	77.78	74.60	71.80	69.12	67.2

3.5. Experiments on the PolyU Palmprint Database

Finally, the experimental results on the PolyU palmprint database can be found in Table 8. According to this table, the highest average recognition rate with 1–5 training samples on average is 95.04% when using our proposed NCRC classifier, where various λ values were tested similar to the other experiments (refer to Table 9). Compared with CRC, the improvement of NCRC is 0.13% on average, which shows NCRC enhanced the recognition ability of the original CRC classifier on the PolyU palmprint database. When compared to the other classifiers, NCRC also achieved a better result than SRC (95.03%), Furthermore, the proposed method outperformed other variants of SRC and CRC such as S**CRC* (94.92%), ProCRC (93.54%), KNN (57.99%), and SVM (86.91%). In Table 8, the highest average accuracy achieved is marked in bold. In Table 9, highest accuracy achieved in each parameter λ is marked in bold.

Table 8. Average accuracy of NCRC and the comparison methods on the PolyU palmprint database.

SRC	CRC	S* <i>CRC</i>	ProCRC	SVM	KNN	NCRC
95.03	94.91	94.92	93.54	86.91	57.99	95.04

Table 9. Parameter selection of NCRC on the PolyU palmprint database.

λ	0.001	0.01	0.1	0.2	0.3	0.4
Accuracy	94.9	95.04	93.4	92.28	91.26	97

3.6. Comparison of Classification Time

To demonstrate the classification efficiency of the proposed NCRC, we further made comparisons between NCRC and other classifiers in terms of classification time. Figure 10 shows the classification time of SRC, CRC, S**CRC* and NCRC on the MUCT (left) and PolyU (right) databases with an increasing number of training samples. As shown in Figure 10, NCRC required less classification time compared with the other classifiers, which indicates a higher efficiency for image classification. Even though the classification accuracy of MUCT using S**CRC* with one training sample is slightly higher than the result of NCRC (refer to Table 6), in terms of classification time, NCRC performed (over three times) faster than S**CRC*. As a variant of CRC, the ProCRC used less classification time than our proposed method, while its performance on the datasets show inferiority compared with our proposed method. These can be seen in Tables 2, 4, 6 and 8, with NCRC outperforming ProCRC every time except on the LFW database using five training samples, where the difference is only 0.39%.

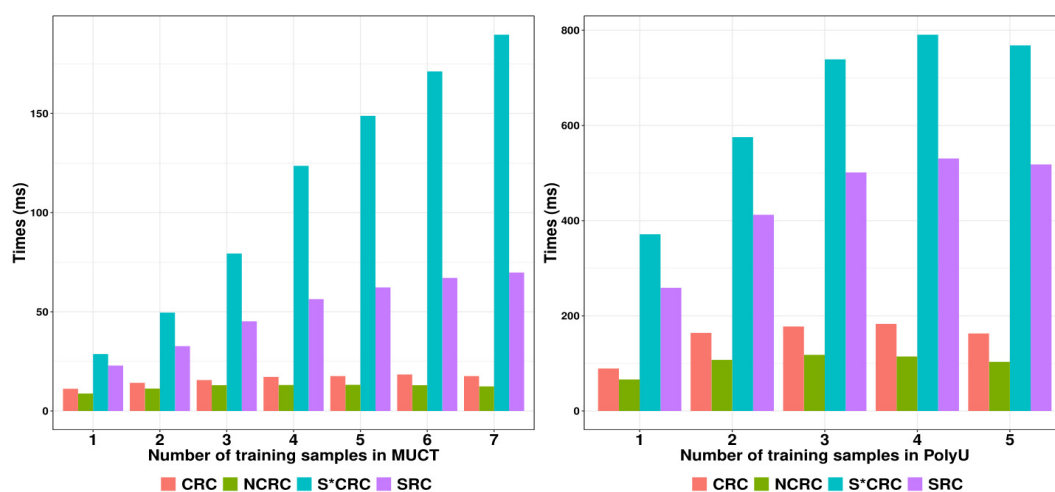


Figure 10. Comparison of the classification times using different databases with: MUCT (left); and PolyU (right).

3.7. Discussion

We conducted experiments ranging from face to palmprint recognition, where it has been well proven that our proposed method achieved promising performances as well as efficient classification times. Here, we are able to reach the following inferences from the experiments:

1. For face recognition, NCRC tends to achieve better results when the number of training samples is increased compared with SRC and CRC. Here, the highest improvement reaches 17.3% on the LFW database. Furthermore, NCRC (AR: 93.06%; LFW: 55.32%; and MUCT: 77.78%) is even more effective in terms of accuracy than S*CRC (AR: 82.50%; LFW: 50.25%; and MUCT: 76.85%) and ProCRC (AR: 91.81%; LFW: 51.66%; and MUCT: 74.32%), which are refined classifiers based on SRC (AR: 82.50%; LFW: 38.02%; and MUCT: 77.07%) and CRC (AR: 91.94%; LFW: 52.00%; and MUCT: 75.86%).
2. When it comes to palmprint recognition, NCRC (95.04%) also shows competitive recognition rates on average, reaching its highest improvement of 1.5% over other state-of-the-art sparse representation methods such as SRC (95.03%), CRC (94.91%), S*CRC (94.92%), and ProCRC (93.54%). This indicates that our proposed method is not only effective in face recognition, but also other image classification tasks.
3. Besides the recognition rate, NCRC (MUCT: 12.1 ms) consumed less time in classification compared with SRC (MUCT: 50.9 ms), CRC (MUCT: 16.0 ms) and S*CRC (MUCT: 113 ms) implying its efficiency in image classification. Although ProCRC performed faster than NCRC, in terms of the recognition rate, the proposed method outperformed all classifiers on average.

4. Conclusions

In this paper, we propose a novel collaborative representation classifier termed as NCRC to perform image classification. When performing test sample representation, NCRC uses a ReLU function to enhance the sparsity of the CRC coefficients. Afterwards, the proposed method represents the test sample collaboratively using new non-negative coefficients. Finally, in the classification stage, the nearest subspace classifier is applied. According to extensive experiments, the proposed NCRC on average outperformed other popular classifiers in different image classification tasks ranging from face recognition to palmprint recognition using both recognition rate and classification time as performance measurements. This proves our new representation is effective as well as efficient. Therefore, the novel classifier has the potential to be applied to real recognition tasks that require a higher accuracy and faster recognition speeds. As part of our future work, we will extend this method using non-negative representation for classification to other classifiers to eventually develop a deep learning based methodology.

Author Contributions: Conceptualization, J.Z.; methodology, J.Z.; validation, J.Z.; formal analysis, J.Z.; investigation, J.Z.; resources, J.Z. and B.Z.; writing—original draft preparation, J.Z.; supervision, B.Z.; and funding acquisition, B.Z.

Funding: This research was funded by the National Natural Science Foundation of China (No. 61602540).

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

Abbreviations

The following abbreviations are used in this manuscript:

SRC	Sparse representation-based classifier
CRC	Collaborative representation-based classifier
ProCRC	Probabilistic collaborative representation-based classifier
NCRC	Non-negative collaborative representation-based classifier
KNN	K nearest neighbor classifier
SVM	Support vector machine
ReLU	Restricted Linear Unit

References

1. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*; Neural Information Processing Systems Foundation: Lake Tahoe, CA, USA, 2012; pp. 1097–1105.
2. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
3. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
4. Wright, J.; Ma, Y.; Mairal, J.; Sapiro, G.; Huang, T.S.; Yan, S. Sparse representation for computer vision and pattern recognition. *Proc. IEEE* **2010**, *98*, 1031–1044. [[CrossRef](#)]
5. Huang, K.; Aviyente, S. Sparse representation for signal classification. In *Advances in Neural Information Processing Systems*; Neural Information Processing Systems Foundation: Vancouver, BC, Canada, 2007; pp. 609–616.
6. Wright, J.; Yang, A.Y.; Ganesh, A.; Sastry, S.S.; Ma, Y. Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 210–227. [[CrossRef](#)] [[PubMed](#)]
7. Xu, Y.; Zhang, D.; Yang, J.; Yang, J.Y. A two-phase test sample sparse representation method for use with face recognition. *IEEE Trans. Circuits Syst. Video Technol.* **2011**, *21*, 1255–1262.
8. Fan, Z.; Ni, M.; Zhu, Q.; Liu, E. Weighted sparse representation for face recognition. *Neurocomputing* **2015**, *151*, 304–309. [[CrossRef](#)]
9. Chang, L.; Yang, J.; Li, S.; Xu, H.; Liu, K.; Huang, C. Face recognition based on stacked convolutional autoencoder and sparse representation. In Proceedings of the 2018 IEEE 23rd International Conference on Digital Signal Processing (DSP), Shanghai, China, 19–21 November 2018; pp. 1–4. [[CrossRef](#)]
10. Shu, T.; Zhang, B.; Tang, Y. Novel noninvasive brain disease detection system using a facial image sensor. *Sensors* **2017**, *17*, 2843. [[CrossRef](#)] [[PubMed](#)]
11. Zhao, S.; Zhang, B.; Chen, C.P. Joint deep convolutional feature representation for hyperspectral palmprint recognition. *Inf. Sci.* **2019**, *489*, 167–181. [[CrossRef](#)]
12. Jin, W.; Gong, F.; Zeng, X.; Fu, R. Classification of clouds in satellite imagery using adaptive fuzzy sparse representation. *Sensors* **2016**, *16*, 2153. [[CrossRef](#)] [[PubMed](#)]
13. Shi, L.; Song, X.; Zhang, T.; Zhu, Y. Histogram-based CRC for 3D-aided pose-invariant face recognition. *Sensors* **2019**, *19*, 759. [[CrossRef](#)] [[PubMed](#)]
14. Timofte, R.; Van Gool, L. Weighted collaborative representation and classification of images. In Proceedings of the IEEE 21st International Conference on Pattern Recognition (ICPR2012), Tsukuba, Japan, 11–15 November 2012; pp. 1606–1610.
15. Akhtar, N.; Shafait, F.; Mian, A. Efficient classification with sparsity augmented collaborative representation. *Pattern Recognit.* **2017**, *65*, 136–145. [[CrossRef](#)]
16. Zeng, S.; Gou, J.; Yang, X. Improving sparsity of coefficients for robust sparse and collaborative representation-based image classification. *Neural Comput. Appl.* **2018**, *30*, 2965–2978. [[CrossRef](#)]
17. Deng, W.; Hu, J.; Guo, J. In defense of sparsity based face recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 25–27 June 2013, pp. 399–406.
18. Cai, S.; Zhang, L.; Zuo, W.; Feng, X. A probabilistic collaborative representation based approach for pattern classification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2950–2959.
19. Dong, X.; Zhang, H.; Zhu, L.; Wan, W.; Wang, Z.; Wang, Q.; Guo, P.; Ji, H.; Sun, J. Weighted locality collaborative representation based on sparse subspace. *J. Vis. Commun. Image Represent.* **2019**, *58*, 187–194. [[CrossRef](#)]
20. Zeng, S.; Yang, X.; Gou, J. Multiplication fusion of sparse and collaborative representation for robust face recognition. *Multimed. Tools Appl.* **2017**, *76*, 20889–20907. [[CrossRef](#)]
21. Zheng, C.; Wang, N. Collaborative representation with k-nearest classes for classification. *Pattern Recognit. Lett.* **2019**, *117*, 30–36. [[CrossRef](#)]
22. Nair, V.; Hinton, G.E. Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th International Conference on Machine Learning (ICML-10), Haifa, Israel, 21–24 June 2010; pp. 807–814.

23. Zeng, S.; Zhang, B.; Zhang, Y.; Gou, J. Collaboratively weighting deep and classic representation via l_2 regularization for image classification. In Proceedings of the Asian Conference on Machine Learning, Beijing, China, 14–16 November 2018; pp. 502–517.
24. Zhang, L.; Yang, M.; Feng, X. Sparse representation or collaborative representation: Which helps face recognition? In Proceedings of the IEEE 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 471–478.
25. Zhu, P.; Zhang, L.; Hu, Q.; Shiu, S.C. Multi-scale patch based collaborative representation for face recognition with margin distribution optimization. In *European Conference on Computer Vision*; Springer: Berlin, Germany, 2012; pp. 822–835.
26. Zhu, P.; Zuo, W.; Zhang, L.; Shiu, S.C.K.; Zhang, D. Image set-based collaborative representation for face recognition. *IEEE Trans. Inf. Forensics Secur.* **2014**, *9*, 1120–1132.
27. Song, X.; Chen, Y.; Feng, Z.H.; Hu, G.; Zhang, T.; Wu, X.J. Collaborative representation based face classification exploiting block weighted LBP and analysis dictionary learning. *Pattern Recognit.* **2019**, *88*, 127–138. [[CrossRef](#)]
28. Martinez, A.M.; Benavente, R. The AR face database. In *CVC Technical Report24*; June 1998. Available online: <http://www.cat.uab.cat/Public/Publications/1998/MaB1998/CVCReport24.pdf> (accessed on 6 June 2019).
29. Learned-Miller, E.; Huang, G.B.; RoyChowdhury, A.; Li, H.; Hua, G. Labeled faces in the wild: A survey. In *Advances in Face Detection and Facial Image Analysis*; Springer: Berlin, Germany, 2016; pp. 189–248.
30. Milborrow, S.; Morkel, J.; Nicolls, F. The MUCT landmarked face database. *Pattern Recognit. Assoc. S. Afr.* **2010**, 179–184.
31. PolyU. *PolyU Multispectral Palmprint Database*; The Hong Kong Polytechnic University: Hong Kong, China, 2003.
32. Aggarwal, C.C. *Data Mining: The Textbook*; Springer: Berlin, Germany, 2015.



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).