Research Article

New Robust Part-Based Model with Affine Transformations for Facial Landmark Localization and Detection in Big Data

Chentao Zhang,1,2 Habte Tadesse Likassa,3 Peidong Liang,2 and Jielong Guo4

1Fujian (Quanzhou)-HIT Research Institute of Engineering and Technology, Quanzhou 362000, China
2Department of Instrumental and Electrical Engineering, Xiamen University, Xiamen, China
3Department of Statistics, College of Natural and Computational Sciences, Addis Ababa University, Addis Ababa, Ethiopia
4Fujian Institute of Research on the Structure of Matter Fuzhou, Chinese Academy of Sciences, Fuzhou, China

Correspondence should be addressed to Peidong Liang; 420374727@qq.com

Received 31 March 2021; Revised 7 October 2021; Accepted 9 October 2021; Published 3 November 2021

Academic Editor: Angelos Markopoulos

Copyright © 2021 Chentao Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we developed a new robust part-based model for facial landmark localization and detection via affine transformations. In contrast to the existing works, the new algorithm incorporates affine transformations with the robust regression to tackle the potential effects of outliers and heavy sparse noises, occlusions and illuminations. As such, the distorted or misaligned objects can be rectified by affine transformations and the patterns of occlusions and outliers can be explicitly separated from the true underlying objects in big data. Moreover, the search of the optimal parameters and affine transformations is cast as a constrained optimization programming. To mitigate the computations, a new set of equations is derived to update the parameters involved and the affine transformations iteratively in a round-robin manner. Our way to update the parameters compared to the state of the art of the works is relatively better, as we employ a fast alternating direction method for multiplier (ADMM) algorithm that solves the parameters separately. Simulations show that the proposed method outperforms the state-of-the-art works on facial landmark localization and detection on the COFW, HELEN, and LFPW datasets.

1. Introduction

Localization of detailed facial features arises in a variety of applications including occlusion coherence for prediction in high-dimensional images [1], landmark localization [2–4], head pose estimation [5, 6], face image alignment [7], low-rank estimation [8–10], and object detection [11–14], etc. However, high-dimensional data are easily impacted by outliers and occlusions. It is thus of importance to develop a deep robust part-based models combining the robust regression algorithm with affine transformation for separating the landmark localization from the occlusions as in [15] and detection [16]. Thus, developing a new algorithm that is resilient with adverse effects while processing big data particularly for face detection and landmark localization is highly indispensable.

A number of algorithms have been suggested for robust facial landmark localization and object detection [17–24]. For instance, [18, 25, 26] addressed these problems using a regression approach. However, these approaches are mostly an approximate and, additionally, they are no longer pose-invariant. To improve the performance of the algorithm and overcome the aforementioned setback, deformable part models (DPMs) based on high-dimensional images [27–30] were taken into account. However, they can not deal with a huge amount of outliers and occlusions. Zhang et al. [31] proposed a tasks-constrained deep convolutional network (TCDCN) through multitask learning which directly conducted the simulations based on high-dimensional images. However, it ignores the potential annoying effects and the correlation between images which jeopardize the robustness of the algorithm. Plus, the issue of the
2. Characteristics of the State-of-the-Art Methods

In this section, we illustrate the characteristics of the five state-of-the-art works such as TCDCN [31], RFLD [35], Fast-SIC [41], HPM [42], and UF-DD [45] with the proposed approach, which further adds affine transformations and multiple subspaces. The merits of the proposed method are quite different from the baselines in terms of its novelty and updating the parameters. For instance, TCDCN [31] addressed the idea of facial landmark detection and occlusion and pose variation via deep multitask learning, but it yet requires to incorporate affine transformation to explore deep multitask learning in dense landmark detection and high-dimensional image recovery. This method considered a few regularization parameters with about six different labels that fail to penalize the complexity of the weights. To tackle the dilemmas of gross errors and variation in images, RFLD [35] proposed a novel regression method that substitutes the commonly used least squares regression developed by [31] via the use of the $L_{2,1}$ norm, but it lacks performance when there are a large number of landmarks. Moreover, there are seven numbers of parameters taken into account, which increases the computational complexity of the method. To alleviate the problems in facial deformable models for the unconstrained images, Fast-SIC was developed by [41], where six different parameters including the vector appearance, localization, and warping parameters involved in problem formulation, which makes the method difficult to optimize the parameters. To solve this setback, [42] proposed HPM for facial landmark localization and occlusion estimation to boost the performance, with five parameters and a number of network hierarchical structure of layers which is increasing the computational complexity of the algorithm. To tackle the setback of outliers and heavy sparse noises, occlusions and illuminations, [45] proposed a novel UF-DD method for face detection where five numbers of parameters are involved; however, the impacts of large outliers and heavy sparse noises are not alleviated. As a new annotated unconstrained face detector is addressed in [45], the time complexity is relatively better over RFLD [35], Fast-SIC [41], and HPM [42]. In the latest work, there are about seven number of parameters involved including affine transformation. As affine transformation is newly added in the new model, it shares the advantage of pruning out the potential impacts of occlusions and illuminations, outliers, and heavy sparse noises in landmark localization and facial detection from which we note that the time complexity of the proposed method is highly reduced and the performance is enhanced. Another advantage of the new part-based model over the state-of-the-art works is it considers multiple subspace parameters to constrain each parameters and shares the advantages of positive semidefinite. However, the new technique seeks to establish the spatial dependency between different images that can be captured incorporating the spatial weight matrix, as such the performance of the new part-based model be boosted.

In this paper, we present a new robust regression algorithm for facial landmark localization and face detection. Invoked by the novel idea of affine transformations in [8], and to be more resilient to occlusions and outliers, the new algorithm incorporates affine transformations with the robust regression based on the hierarchical part-based models [1, 42]. As such, the distorted or misaligned images can be rectified by affine transformations and the patterns of occlusions and outliers can be explicitly separated. The proposed method via affine transformation contributes in clearly separating the facial landmark localization and occlusion estimation. The search of the optimal parameters and the affine transformations is cast as a constrained
optimization programming. To mitigate the computational overhead, a new set of equations are derived to update the parameters involved and the affine transformations iteratively using ADMM approach in a round-robin manner. Conducted simulations show that the proposed method outperforms the state-of-the-art works on face detection and landmark localization on some common on the COFW, HELEN, and LFPW datasets. The major contributions of this work include the following:

1. The affine transformations are incorporated in the robust regression based on the part-based models to take advantages of both schemes in the learning process.
2. The affine transformations are aggregated with the low-rank-sparse representation, where the low-rank component lies in a union of subspaces instead of one single subspace. These transformations can fix the distortion or misalignment in a batch of corrupted images to render more faithful image decomposition, thereby being more robust against heavy sparse errors and outliers.
3. The ADMM method is employed to solve the new convex optimization problem, and a set of updating equations is derived to iteratively update the optimization variables and the affine transformations.
4. New set of updating equations is established to iteratively solve the constrained optimization problem.
5. We conduct experiments on several benchmark datasets, and the experimental results demonstrate the effectiveness of our new method.

3. Problem Formulation

Given $n$ images $\{I_i^0\} \in \mathbb{R}^{w \times h}$, $i = 1, \ldots, n$, all of which contain the same object and are linearly correlated, where $w$ and $h$ denote the weight and height of the images, respectively. Based on the part-based models [1], we approximate $A = \{\text{vec}(I_i^0)\} \cdots \{\text{vec}(I_n^0)\} \in \mathbb{R}^{bn}$ by a shape parameter, where $\text{vec}(-)$ is the vector stacking operator [46] and $l$ is the number of landmarks. However, in practice, the images $\{I_i^0\}$ are usually contaminated by occlusions and outliers, so the images $M \in \mathbb{R}^{bn}$ can be represented as $M = AB + E$, where $M \in \mathbb{R}^{bn}$ is a corrupted high-dimensional images, $E \in \mathbb{R}^{bn}$ is a term accounting for the occlusions and outliers, and $B \in \mathbb{R}^{m \times n}$ is an appearance parameter.

To solve the misalignment problem incurred by the outliers and occlusions in $M$, the original images need to be linearized, which can be done by affine transformations. Applying a set of affine transformations, denoted by $\tau_1, \ldots, \tau_n$ to $M$ [33, 47], then we have $M_i' = M_i + \sum_{j=1}^n I_j \Delta \tau_i v_j^T$, $\Delta \tau \in \mathbb{R}^{q \times n}$, in which $q$ denotes the number of parameters, $I_j = \{\partial \text{vec}(I_j^i) / \partial \tau_i\} \in \mathbb{R}^{n \times p}$ denotes the Jacobian of the $i$th training images with respect to $\tau_i$, and $v_i$ denotes the standard basis for $\mathbb{R}^n$. Thus, consequently, the main objective is to minimize the localization error via incorporating a novel idea of affine transformation for better performance, so the overall problem can thus be posed as the following constrained optimization problem

$$\begin{align*}
\min_{A, L, \beta, E, B, \Delta \tau} & \quad \frac{\gamma}{2} \|U\|_F^2 + \|A\|_* + \|B\|_* + \lambda_1 \|G\|_1 + \lambda_2 \|E\|_1, \\
\text{s.t} & \quad Q = AB + E, \quad \tilde{\beta} = \begin{bmatrix} AB ; 1^T \end{bmatrix}, B = G, G \geq 0,
\end{align*}$$

(1)

where $U = W(R - L \tilde{\beta})$, $Q = M_i^0 + \sum_{j=1}^n I_j \Delta \tau_i v_j^T$, $W \in \mathbb{R}^{d \times d}$, is a weight matrix, $L \in \mathbb{R}^{d \times (l+1)}$ is the regression matrix mapping $M_i^0 \in \mathbb{R}^{b \times n}$ to $R \in \mathbb{R}^{d \times n}$, $\beta \in \mathbb{R}^{(l+1) \times n}$ is a regression coefficient that controls the appearance variation, $I_j \in \mathbb{R}^n$ denotes as a vector of all ones, $E \in \mathbb{R}^{b \times n}$, $\lambda_1, \gamma, \lambda_2$ and $\lambda_2$ are the regularization parameters, $\|A\|_1 = \sum_{i=1}^{n} \|A_i\|_1 \sigma_i(A)$ with $\sigma_i(A)$ being the singular values of $A$, $\|A\|_F = \text{Trace}(A^T A)$, and $\|E\|_1 = \sum_{i=1}^{n} \|E_i\|_1$.

3.1. Proposed Approach. To solve a constrained optimization problem (2), we consider an ADMM given by

$$\begin{align*}
\min_{L, \beta, A, E, B, \Delta \tau} & \quad \frac{\gamma}{2} \|U\|_F^2 + \|A\|_* + \|B\|_* + \lambda_1 \|G\|_1 + \lambda_2 \|E\|_1 + \langle Y_1, C \rangle + \frac{\mu_1}{2} \|C\|_F^2 + \langle Y_2, \beta - [AB ; 1^T] \rangle \\
& + \frac{\mu_2}{2} \| \beta - [AB ; 1^T] \|_F^2 + \langle Y_3, B - G \rangle + \frac{\mu_3}{2} \| B - G \|_F^2, \\
\end{align*}$$

(3)

where $Y_1 \in \mathbb{R}^{b \times n}$, $Y_2 \in \mathbb{R}^{(l+1) \times n}$ and $Y_3 \in \mathbb{R}^{m \times n}$ are the Lagrangian multipliers, $C = (Q - AB - E)$, $\mu_1$, $\mu_2$, and $\mu_3$ are the penalty parameters, and $\langle X, Y \rangle = \text{Trace}(X^T Y)$. This can be solved by using ADMM. The major necessary condition of ADMM is the convexity issue; we give sufficient conditions under which the algorithm asymptotically reaches the standard first-order necessary conditions for local optimality.

Then, based on a linearized alternating direction method [48], the augmented Lagrangian multiplier in (3) can be reexpressed as

$$\begin{align*}
\min_{L, \beta, A, E, B, G, \Delta \tau} & \quad \frac{\gamma}{2} \|U\|_F^2 + \|A\|_* + \|B\|_* + \lambda_1 \|G\|_1 + \lambda_2 \|E\|_1 + \\
& + \frac{\mu_1}{2} \|C + \frac{Y_1}{\mu_1}\|_F^2 + \frac{\mu_2}{2} \|V + \frac{Y_2}{\mu_2}\|_F^2 + \frac{\mu_3}{2} \|B - G + \frac{Y_3}{\mu_3}\|_F^2, \\
\end{align*}$$

(4)

where $V = (\tilde{\beta} - [AB ; 1^T])$. Directly solving (4) is computationally expensive, so to follow, we consider a set of equations to iteratively update the parameters in (4). So typically,
we can choose to minimize the Lagrangian function to update all involved parameters involved alternatively.

Firstly, the updates of $L$ and $\hat{\beta}$ are determined, respectively, by:

$$L^{k+1} = \arg\min_L \left( L^k, \beta \wedge, A^k, \lambda^k, E^k, B^k, G^k, \Delta \tau^k \right)$$

$$\beta^{k+1} = \arg\min_{\beta} \left( L^k, \beta \wedge, A^k, \lambda^k, E^k, B^k, G^k, \Delta \tau^k \right)$$

where $k$ is the index of iterations, and by ignoring all irrelevant terms and applying the ordinary least square regression procedures on $L$ and $\hat{\beta}$, finally, the updates of $L$ and $\hat{\beta}$ are given by:

$$L^{k+1} = \left( \beta \wedge \beta \wedge^T + \gamma \left( I_{d_{m+1}} \right) \right)^{-1} R(\beta \wedge)^T$$

$$\beta^{k+1} = \left[ NL + \mu_2 L \right]^{-1} \left[ N R - Z \right]$$

where $N = (\gamma L^T W^T W)$ and $Z = (Y_2 + \mu_2 [AB ; 1]^T)$ and $I_{d_{m+1}}$ is an identity matrix.

Secondly, to update the shape parameter $\lambda$, we fix $E, L, \hat{\beta}, G, B$, and $\Delta \tau$ remain as constants, and $\lambda^{k+1}$ can be determined by:

$$\lambda^{k+1} = \arg\min_{\lambda} \left( L^{k+1}, \beta \wedge, A^{k+1}, E^{k+1}, B^k, G^k, \Delta \tau^k \right)$$

Using the soft shrinkage operator, we can get the update of $A$ as:

$$A^{k+1} = \Omega_{1/k\lambda} \left( A^k - \frac{p_k}{k\lambda} \right)$$

where $\Omega$ is the singular value threshold, $P_k = (\mu_1 + \mu_2)A^T B - \mu_1 (Q - E) - (Y_2 - \mu_2 (\hat{\beta} + Y_2) B^T$, $\kappa_k = (\mu_1 + \mu_2) / 2$, $\eta_k > \sigma(B^T B)$ is the proximal parameter, in which $\sigma(B^T B)$ denotes the spectral radius of $B^T B$.

Thirdly, to update $E$, we fix $\hat{\beta}, A, L, G, B$, and $\Delta \tau$, and $E^{k+1}$ can be determined by:

$$E^{k+1} = \arg\min_E \left( L^{k+1}, \beta \wedge, A^{k+1}, E^{k+1}, B^k, G^k, \Delta \tau^k \right)$$

By invoking a linearized alternating direction, $E$ is updated by:

$$E^{k+1} = \Gamma / \mu \left( Q - A^T B + \frac{Y_1}{\mu} \right)$$

where $\Gamma$ is the soft shrinkage operator [49, 50].

Next, to get the optimal update of $B$, we again keep all $A, E, L, \hat{\beta}, G, \Delta \tau$ as constants, then the updating parameter $B$ can be determined by:

$$B^{k+1} = \arg\min_B \left( L^{k+1}, \beta \wedge, A^{k+1}, E^{k+1}, B^{k+1}, G^k, \Delta \tau^k \right)$$

Then, by employing a singular value threshold operators, the update of $B$ is given by:

$$B^{k+1} = \Omega_{2/k\beta} \left( B^k - \frac{p_k^B}{k\beta} \right)$$

where $\kappa_B = (\mu_1 + \mu_2 + \mu_3)(\eta_B / 2), \eta_B > \sigma(AA^T)$ and $P_k^B = A^T (\mu_1 + \mu_3) A^{k+1} - \mu_1 (Q - E) - Y_1 - (\mu_2 (\beta + Y_2) + \mu_3 (B^k + G^k)) + Y_3.$

To find the optimal update of $G$, we fix $A, E, L, \hat{\beta}, B$, and $\Delta \tau$ as constants, then $G^{k+1}$ can be obtained by:

$$G^{k+1} = \arg\min_G \left( L^{k+1}, \beta \wedge, A^{k+1}, E^{k+1}, B^{k+1}, G^{k+1}, \Delta \tau^k \right)$$

Similarly, employing the soft threshold operator and an augmented Lagrangian multiplier to subproblem (14), then finally, $G$ is updated by:

$$G^{k+1} = \Gamma / \mu \left( B^{k+1} + \frac{Y_3}{\mu} \right)$$

Invoked with an affine transformations, we obtained an
additional optimal parameter. Therefore, we need to derive an additional update for $\Delta \tau$.

To do so, we update $\Delta \tau$, we fix $A$, $E$, $L$, $\hat{\beta}$, $B$, and $G$, and $\Delta \tau^{k+1}$ can be obtained by

$$\Delta \tau^{k+1} = \arg\min_{\Delta \tau} \mathcal{L} \left( L^{k+1}, \hat{\beta}^{k+1}, A^{k+1}, E^{k+1}, B^{k+1}, G^{k+1}, \Delta \tau \right).$$

Along the same line, the affine transformations $\Delta \tau$ can be updated by

$$\Delta \tau^{k+1} = \sum_{i=1}^{n} \chi_i^T \left( A^{k+1} B^{k+1} + E^{k+1} - M_{\tau} - \frac{Y_i}{\mu_i} \right) v_i v_i^T,$$

where $\chi_i^T$ is a More-Penrose-pseudoinverse of $\chi_i$.

Figure 1: Precision-recall curves of the proposed method and baselines methods on the (a) COFW dataset and (b) LFPW dataset.
Similarly, the updates of $Y_1$, $Y_2$, and $Y_3$ are given, respectively, by

$$Y_1^{k+1} = Y_1^k + \mu_1^{k+1} \left( Q - A^{k+1} B^{k+1} - E^{k+1} \right)$$

$$Y_2^{k+1} = Y_2^k + \mu_2^{k+1} \left( \beta - \left[ A^{k+1} B^{k+1} \Delta \right] \right)$$

$$Y_3^{k+1} = Y_3^k + \mu_3^{k+1} \left( B^{k+1} - G^{k+1} \right)$$

The updates of the regularization parameters $\mu_1$, $\mu_2$, and $\mu_3$ are given by

$$\mu_1^{k+1} = \min \left( \mu_{\max}, \rho \mu_1^k \right)$$

$$\mu_2^{k+1} = \min \left( \mu_{\max}, \rho \mu_2^k \right)$$

$$\mu_3^{k+1} = \min \left( \mu_{\max}, \rho \mu_3^k \right)$$

where $\rho$ is a properly chosen constant and $\mu_{\max}$ are tunable parameters adjusting the convergence of the proposed algorithm. We sequentially update $L$, $\beta$, $A$, $E$, $G$, and $\Delta$ independently by keeping all of the other parameters unchanged. First, the regression matrix $L$ and $\beta$ are updated by (6). Next, we can update $A$, $E$, $G$, and the optimal affine transformation parameters, $\Delta \tau$, by (8), (10), (12), (14), and (16), respectively. Finally, the Lagrangian multipliers $Y_1$, $Y_2$, and $Y_3$ are updated by (17) and the regularization parameters $\mu_1$, $\mu_2$, and $\mu_3$ are updated by (18). These new updating equations proceed in a round-robin manner until convergence. Since a monotonically decreasing and bounded below sequence will convergence, the above algorithm is ensured to converge. It is noteworthy that based on the augmented Lagrangian multiplier, ordinary least squares procedures, and soft-threshold operators, all the updating parameters are achieved. The overall summary of the proposed method compared with other related works is summarized in Table 1.

**4. Experimental Results**

In this section, some simulations are conducted to assess the effectiveness of the proposed algorithm. Three datasets are considered in the simulations, including the Labeled Face Parts in the Wild (LFPW) [65], the HELEN68 [21], and the more challenging Caltech Occluded Faces in the Wild (COFW) [66] datasets. First, we consider the novel ideas of affine transformation that tackle the adverse effects of outliers, heavy sparse noises, occlusions, and illuminations. After an intensive mathematical derivation, we experimentally supported the effectiveness of the proposed method through some numerical simulations.

**4.1. Comparison with the State-of-the-Art Methods.** In this subsection, we compare the proposed approach, which adds affine transformations with some recently reported works in terms of precision-recall curves, landmark localization error, and time complexity on the aforementioned three databases.

<table>
<thead>
<tr>
<th>Methods</th>
<th>HELEN68</th>
<th>COFW</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCDCN [31]</td>
<td>0.0910</td>
<td>0.0892</td>
</tr>
<tr>
<td>RFLD [35]</td>
<td>0.0799</td>
<td>0.0687</td>
</tr>
<tr>
<td>Fast-SIC [41]</td>
<td>0.0422</td>
<td>0.0382</td>
</tr>
<tr>
<td>HPM [42]</td>
<td>0.0138</td>
<td>0.0178</td>
</tr>
<tr>
<td>UF-DD [45]</td>
<td>0.0122</td>
<td>0.0142</td>
</tr>
<tr>
<td>Ours</td>
<td>0.0067</td>
<td>0.0063</td>
</tr>
</tbody>
</table>

In this work, five different state-of-the-art methods, including TCDCN [31], RFLD [35], Fast-SIC [41], HPM [42], UF-DD [45], and the proposed algorithm are conducted based on the performance evaluation and time complexity. The results from these baselines are to reimplement of the publicly available codes.

**4.2. Face Detection.** First, we assess the proposed approach for face detection based on the LFPW and COFW databases in terms of the precision-recall curve. The comparisons of the proposed approach with UF-DD [45] and HPM [42] are shown in Figure 1, from which we can see that UF-DD outperforms HPM, as it downsamples the images with high resolution using an unconstrained face detector to tackle the impact of large variations. As noted in Figure 1, the HPM without occlusion is relatively better compared to the HPM with occlusion, Multiresolution HPM with rotation outperforming to the HPM with occlusion and multiresolution HPM without rotation is relatively outperforming to all other versions of the HPM in face detecting. We can also notice from Figure 1 that the proposed algorithm is superior than all HPM versions and UF-DD by achieving better recall rate on both datasets. The superiority of the new approach is due to an incorporation of an affine transformation and the multiple subspaces and constraining the parameters taking the advantages of positive semidefinite, which makes it to be more robust with a various adverse effects of outliers, illuminations, occlusions, and heavy sparse noises. This is because the new algorithm combines affine transformations with the part-based model so that it is more robust against the impact of occlusions and outliers. To justify the effectiveness of the proposed method, comparison of the localization error is taken into account, from which we note that the proposed method is outperforming to five different baselines. The performance of the proposed method is superior to all the baselines. This is due to affine transformation; first, it aligns the misaligned images and it prunes out the potential impact of annoying effects which boosts the performance of the proposed method.

**4.3. Facial Landmark Localization.** In this subsection, we compare the proposed algorithm with some state-of-the-art for the facial landmark localization on the HELEN68 and COFW datasets. The comparison with five baselines, TCDCN [31], RFLD [35], Fast-SIC [41], HPM [42], and UF-DD [45] in terms of the root mean square error is as shown in Table 2, from which we can see that TCDCN has
Figure 2: Performance evaluation of the proposed method via facial landmark localization on the HELEN68 dataset: (a) TCDCN, (b) RFLD, (c) Fast-SIC, (d) HPM, (e) UF-DD, and (f) ours. The red points indicate the key points which are predicted as being occluded.
the largest error compared with the other baselines, as it ignores the multicollinearity during multitasks to localize the objects. Also, RFLD [35] and Fast-SIC [41] can not provide satisfactory landmark localization accuracy, as their localization outputs are distorted and the patterns of occlusions can not be explicitly separated. We can also notice that HPM outperforms [31, 35, 41], as it utilizes a hierarchical-based structures to explicitly model the occlusions to attain accurate landmark localization. UF-DD outperforms [31, 35, 41], since UF-DD further downsampled the images with high resolution using a face detector. We can also find that the proposed method outperforms all baselines on both datasets. Thus, the results of Table 2 justify the effectiveness of the proposed approach based on the localization error for the facial landmark localization as compared with the main-state-of-the art works. This is because it aggregates robust regression using affine transformations with part-based models to accurately localize facial landmarks even under outliers and occlusions.

As an illustration, we also show three images for facial landmark localization based on the aforementioned algorithms as shown in Figure 2, from which we can see that TCDCN, RFLD, and Fast-SIC provide poor localization. This is because the approaches fail to clearly separate the landmark localization from the occluded ones as TCDCN [31] considers combined multitask learning and ignores multicollinearity while RFLD [35] considers a large number of concatenating histogram of oriented gradient (HOG) descriptors with a multiple initializations which influences the separation, which still makes the separation challenging, when the images are distorted by the presence of occlusions as shown in Figures 2(a)–2(c). Moreover, it is also observed that the Fast-SIC [41] is relatively better compared to TCDCN [31] and RFLD [35] in separating landmark localization from the annoying effects such as occlusion, illuminations, outliers, and heavy sparse noises, at it considers the project out optimization framework. UF-DD [45] which incorporates the new annotated unconstrained face detection and downsampled strategy the images with high resolution provides better localization than HPM [31, 35, 41, 42], as it better separates the landmark localization (blue) and occluded landmarks (red) explicitly as depicted in Figure 2(e). The proposed approach works better under the impact of outliers and occlusion and explicitly separates the landmark localization’s from the occlusions as shown in Figure 2(f). This once again justifies that effectiveness of the combination of the affine transformations and part-based models, which makes the proposed method to be more robust with outliers and images which are highly influenced by occlusions.

In summary, the proposed method is relatively better in landmark localization and facial detection; this is due to affine transformation considered detects the potential impact of outliers, heavy sparse noises, occlusions and illuminations. However, the major drawback of the proposed method is it needs to account the spatial dependency between images through incorporating a spatial weight matrix in the mathematical formulation.

5. Computational Complexity

The time complexity of the proposed method as compared to the state-of-the-art works is described in this section. On a very standard desktop computer, the computational load of the baselines mainly TCDCN [31], RFLD [35], Fast-SIC [41], HPM [42], and UF-DD [45] along with the proposed method is given in Table 3, from which we note that the proposed method has small number of running time due to in the proposed method there are less number of parameters involved in updating the parameters. Additionally, our algorithm can handle batches of one hundred images in a few minutes on a standard PC as the number of parameters involved is small as compared to the state-of-the-art works. The new algorithm is guaranteed faster time complexity compared to the state-of-the-art algorithms as shown in Table 3. This is due to affine transformation removing the extreme values, and various regularization parameters are also involved to make the situation of the proposed model stable and gross errors including outliers, occlusions, and illuminations.

6. Conclusions

In this paper, we developed a new model for facial landmark localization and detection, comprising an affine transformation and ADMM methods. The new algorithm combines the efficacious affine transformation-assisted robust regression with the part-based models to enjoy the advantages of both schemes. The problem is formulated as a constrained optimization programming, and a set of equations are established to iteratively update the parameters and the affine transformations. The search of the affine transformations and the optimization variables is formulated as a constrained convex optimization problem. The ADMM approach is then employed, and a new set of equations is established to iteratively update the optimization variables and the affine transformations. Simulations justify the effectiveness of the new algorithm on some common on the COFW, HELEN, and LFPW datasets compared to the state-of-the-art works. To minimize the time complexity, the ADMM approach is further considered in a round-robin manner. In this work, the affine transformation taken into account is to correct the alignments of the individual images, but not consider the issue of spatial dependency between images. This can be captured as one of our future work through incorporating the spatial weight matrix between images. Another future work is to excel this new part-based model in high-dimensional tensor datasets.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Time complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCDCN [31]</td>
<td>4.34</td>
</tr>
<tr>
<td>RFLD [35]</td>
<td>4.56</td>
</tr>
<tr>
<td>Fast-SIC [41]</td>
<td>4.78</td>
</tr>
<tr>
<td>HPM [42]</td>
<td>5.67</td>
</tr>
<tr>
<td>UF-DD [45]</td>
<td>4.20</td>
</tr>
<tr>
<td>Ours</td>
<td>3.56</td>
</tr>
</tbody>
</table>
Data Availability
The data used in this article are freely available for the user.

Conflicts of Interest
The authors declare that there are no conflicts of interest.

Acknowledgments
This work was supported by the National Key Research and Development Program of China under Grant no. 2018YFB1305700 and Scientific and Technological Program of Quanzhou City under Grant no. 2019CT009. We thank Addis Ababa University, Ethiopia.

References


