Mitigating Bias in Face Recognition using Skewness-Aware Reinforcement Learning

Mei Wang, Weihong Deng*
Beijing University of Posts and Telecommunications
{wangmei1, whdeng}@bupt.edu.cn

Abstract

Racial equality is an important theme of international human rights law, but it has been largely obscured when the overall face recognition accuracy is pursued blindly. More facts indicate racial bias indeed degrades the fairness of recognition system and the error rates on non-Caucasians are usually much higher than Caucasians. To encourage fairness, we introduce the idea of adaptive margin to learn balanced performance for different races based on large margin losses. A reinforcement learning based race balance network (RL-RBN) is proposed. We formulate the process of finding the optimal margins for non-Caucasians as a Markov decision process and employ deep Q-learning to learn policies for an agent to select appropriate margin by approximating the Q-value function. Guided by the agent, the skewness of feature scatter between races can be reduced. Besides, we provide two ethnicity aware training datasets, called BUPT-Globalface and BUPT-Balancedface dataset, which can be utilized to study racial bias from both data and algorithm aspects. Extensive experiments on RFW database show that RL-RBN successfully mitigates racial bias and learns more balanced performance.

1. Introduction

Recently, with the emergence of deep convolutional neural networks (CNN) [26, 43, 46, 19, 20], the performance of face recognition (FR) [49, 45, 42] is dramatically boosted. However, as its wider and wider application, its potential for unfairness is raising alarm [8, 5, 1, 2]. For instance, according to [14], a year-long research investigation across 100 police departments revealed that African-American individuals are more likely to be stopped by law enforcement. As stated in the Universal Declaration Human Rights [7], all are equal before the law and are entitled without any discrimination to equal protection. Obviously, the development and deployment of fair and unbiased FR systems is crucial to prevent any unintended side effects and to ensure the long-term acceptance of these FR algorithms. Previous studies [50, 54] have shown that racial bias comes on both data and algorithm aspects. Unfortunately, there are still no sufficient research efforts on the fairness of face recognition algorithms [38, 39, 13], as well as building balanced databases [50], in the literature.

A major driver of bias in face recognition, as well as other AI tasks, is the training data. Deep face recognition networks are often trained on large-scale training datasets, such as CASIA-WebFace [52], VGGFace2 [10] and MS-Celeb-1M [17], which are typically constructed by scraping websites like Google Images. Such data collecting methods can unintentionally produce data that encode gender, ethnic and culture biases. Thus, social awareness must be brought to the building of datasets for training. In this work, we take steps to ensure such datasets are diverse and do not under represent particular groups by constructing two new training datasets, i.e. BUPT-Globalface and BUPT-Balancedface dataset. One is built up according to the population ratio of ethnicity in the world, and the other strictly balances the number of samples in ethnicity.

Another source of bias can be traced to the algorithms themselves. The state-of-the-art (SOTA) face recognition methods, such as Sphereface [30], Cosface [48] and Arcface [12], apply a margin between classes to maximize overall prediction accuracy for the training data. If a specific group of faces appears more frequently than others in the training data, the method will optimize for those individuals because...
testing on different races results in severe performance drop. Dakik et al. [25] collected mug shot face images of White, Black and Hispanic from the Pinellas County Sheriff’s Office (PCSO) and concluded that the Black cohorts are more difficult to recognize for all matchers. However, few efforts were made to study racial bias in deep face recognition. Recently, Wang et al. [50] contributed a test dataset called Racial Faces in-the-Wild (RFW) database, on which they validated the racial bias of four commercial APIs and four SOTA face recognition algorithms, and presented the solution using deep unsupervised domain adaptation to alleviate this bias. But there is still a vacuum in training datasets which can be used to study racial bias.

2.2. Debiased algorithms

In many computer vision applications, there are some works that seek to introduce fairness into networks and mitigate data bias. These are respectively classified as unbalanced-training [41, 9, 36, 53], attribute suppression [5, 33, 37, 32] and domain adaptation [23, 22, 50, 44]. By learning the underlying latent variables in an entirely unsupervised manner, Debiasing Variational Autoencoder (DB-VAE) [6] re-weighted the importance of certain data points while training. Calmon et al. [9] transformed the given dataset into a fair dataset by constraining the conditional probability of network prediction to be similar for any two values of demographic information. SensitiveNets [35] proposed to introduce sensitive information into triplet loss. They minimized the sensitive information, while maintaining distances between positive and negative embeddings.

2.3. Deep reinforcement learning

Mimicking humans’ decision making process, RL aims to enable the agent to decide the behavior from its experiences using a Markov decision process (MDP) [28]. Mnih et al. [34] combined RL learning with CNN and bridged the divide between high-dimensional sensory inputs and actions, resulting in human-level performance in Atari Games. Recently, RL has been successfully applied in compute vision. Rao et al. [40] used RL to discard the confounding frames and found the focuses of attentions for video recognition. Haque et al. [18] utilized RL to identify small, discriminative regions indicative of human identity in person identification. Liu et al. [29] used RL to learn optimal margin to improve overall FR performance. In this paper, we apply deep RL to address race balance problem in FR.

3. Ethnicity aware training datasets

A major driver of bias in face recognition is the training data. Frequently, some race groups are over-represented and others are under-represented. For example, East Asia and India together contribute just 8% of commonly-used training datasets, even though these countries represent 44% of the world’s population. In order to remove this source of
bias and represent people of different regions equally, we construct two ethnicity-aware training datasets, i.e., BUPT-Globalface and BUPT-Balancedface dataset. The identities in these two datasets are grouped into 4 categories, i.e. Caucasian, Indian, Asian and African, according to their races. As shown in Fig. 2, BUPT-Globalface contains 2M images from 38K celebrities in total and its racial distribution is approximately the same as real distribution of world’s population. BUPT-Balancedface dataset contains 1.3M images from 28K celebrities and is approximately race-balanced with 7K identities per race.

![Figure 2](image-url) The percentage of different races in commonly-used training datasets, BUPT-Globalface and BUPT-Balancedface.

Similar to RFW [50], we select images of different races from MS-Celeb-1M [3] with help of the “Nationality” attribute of FreeBase celebrities [15] and Face++ API. However, due to unbalanced distribution of MS-Celeb-1M [3], we can only obtain 2K Indians and 5K Asians which are not enough to construct a large scale dataset. As we know, MS-Celeb-1M only selected the top 100K entities from one-million FreeBase celebrity list [15]. Therefore, we download the remaining images of Asians and Indians by Google according to FreeBase celebrity list, and then clean them both automatically and manually in the similar way as other FR training datasets, such as VGGface2 [10] and Megface [24]. After obtaining enough images, we select images to construct our ethnicity-aware training datasets, and remove their overlapping subjects in RFW [50].

4. Our method

4.1. Investigation and observation

To make learned features potentially separable and enhance the discrimination power, some methods proposed to incorporate a margin between classes based on Softmax. For example, Arcface [12] used an additive angular margin $m$:

$$L_{arc} = - \frac{1}{N} \sum_{j=1}^{N} \log \frac{e^{\gamma \cos(\theta_{yj} + m)}}{\sum_{i \neq y} e^{\gamma \cos(\theta_{ij})}}$$

and Cosface [48] used an additive cosine margin $m$:

$$L_{cos} = - \frac{1}{N} \sum_{j=1}^{N} \log \frac{e^{\gamma \cos(\theta_{yj} + m)}}{\sum_{i \neq y} e^{\gamma \cos(\theta_{ij})}}$$

where $\theta_{ij}$ is the angle between the weight $W_i$ and the feature $x_j$, $x_j \in \mathbb{R}^d$ denotes the deep feature of the $j$-th sample, belonging to the $y$-th class, and $W_i \in \mathbb{R}^d$ denotes the $i$-th column of the weight $W \in \mathbb{R}^{d \times N}$. $N$ is the batch size, $n$ is the number of classes, and $s_i$ is the scale factor. Although large margin losses successfully improve feature discrimination, and get better performance on a series of face recognition benchmarks, Wang et al. [50] experimentally proved that they still fail to obtain balanced representations on different races under a uniform margin.

To get a better understanding, we train two ResNet-34 [19] models with the guidance of Arcface [12] and Softmax loss on the CASIA-Webface [52], and give the detailed angle statistics of different races in Table 1. The intra-class and inter-class angle are computed on set-1 and RFW [50]. The set-1 contains 500 identities per race randomly selected from our BUPT-Globalface dataset. Intra-class angle refers to the mean of angles between feature and the feature centre, which can be formulated as: $\Theta_{intra} = \frac{1}{N_i} \sum_{i \in I_c} \sum_{x_j \in I_c} \theta_{x_j,c_i}$, where $N_i$ is the number of identities belonged to one race group, $I_c$ is the set of all images in $i$-th identity, and $c_i$ is the feature centre of $i$-th identity computed by the mean vector of embeddings. Inter-class angle refers to the mean of minimum angles between embedding feature centres, which can be formulated as: $\Theta_{inter} = \frac{1}{N_y} \sum_{y \neq i} \min_{k=1:N_y, k \neq i} \theta_{c_k,c_i}$. In Table 1, we find that non-Caucasians can not obtain as good intra-class compactness and inter-class discrepancy as Caucasians do, especially the inter-class angle. That is to say, the performance of non-Caucasians is inferior to that of Caucasians on test set, even if a uniform margin is performed on different races when training.

<table>
<thead>
<tr>
<th>Loss functions</th>
<th>Caucasian</th>
<th>Indian</th>
<th>Asian</th>
<th>African</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Softmax</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intra1</td>
<td>35.55</td>
<td>36.00</td>
<td>38.83</td>
<td>37.29</td>
</tr>
<tr>
<td>Inter1</td>
<td>62.67</td>
<td>50.89</td>
<td>46.17</td>
<td>54.26</td>
</tr>
<tr>
<td>Intra2</td>
<td>33.81</td>
<td>31.49</td>
<td>32.09</td>
<td>32.39</td>
</tr>
<tr>
<td>Inter2</td>
<td>59.09</td>
<td>50.15</td>
<td>46.41</td>
<td>48.42</td>
</tr>
<tr>
<td><strong>Arcface [12]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intra1</td>
<td>36.70</td>
<td>40.00</td>
<td>42.78</td>
<td>41.50</td>
</tr>
<tr>
<td>Inter1</td>
<td>67.72</td>
<td>59.73</td>
<td>55.34</td>
<td>62.90</td>
</tr>
<tr>
<td>Intra2</td>
<td>36.22</td>
<td>34.87</td>
<td>36.15</td>
<td>36.18</td>
</tr>
<tr>
<td>Inter2</td>
<td>65.76</td>
<td>59.72</td>
<td>56.46</td>
<td>58.54</td>
</tr>
</tbody>
</table>

Table 1. The angle statistics of different races ([CASIA [52], ResNet34, loss*]). “Intra1” and “Intra2” refer to intra-class angle on set-1 and RFW, respectively. “Inter1” and “Inter2” refer to inter-class angle on set-1 and RFW, respectively.

Considering of poorer generalization ability of non-Caucasians, we should take special care of these difficult colored faces, and thus prefer an appropriate margin with them to improve their generalization ability. Therefore, we introduce the idea of adaptive margin into race balance problem. The margin of Caucasians remains unchanged, while optimal margins are selected adaptively for each col-

9324
Ethnicity aware training datasets

![CNN](image)

Adaptive margin loss guided by agent

![Agent](image)

Deep Q-learning

Offline sampling

**Figure 3.** An illustration of our method. **Offline sampling:** We vary margin for each race group to collect some training samples, i.e. \((s^t, a^t, r^t, s^{t+1})\), before training DQN. Details of \((s^t, a^t, r^t, s^{t+1})\) refer to Section 4.2. **Deep Q-learning network:** With these samples, DQN is trained to approximate the Q-value function, and the reward is determined by the skewness of inter/intra-class distance between races. Then, adaptive margin policy for agent can be generated according to Q-value. **Adaptive margin:** We train a race balanced network with a fixed margin for Caucasians and adaptive margins for each colored-face which changes at each training step guided by agent.

ordered race in order to minimize the skewness of angles between races and learn balanced performance for different races. We replace the fixed margin \(m\) in Eqn. 1 and Eqn. 2 by a race-related and training step related parameter \(m_j(t)\), where \(t\) represents the stage of the training. In fact, different races have different demands for the margins, and the demands may change during the training. The proposed adaptive margin loss function can be formulated as follows:

\[
L_{RBN} = \frac{1}{N} \sum_{j=1}^{N} \log \frac{e^{s_j \cos(y_{j}(s_j) + \alpha_j(t))}}{\sum_{j=1}^{N} e^{s_j \cos(y_{j}(s_j) + \alpha_j(t))}}
\]

where, \(\alpha_j(t) = \begin{cases} m_j(t), & \text{if } j \in \text{Caucasian} \\ m, & \text{otherwise} \end{cases} \)

(3)

The similar modification can be made for Cosface [48]. So the key problem is to find an optimal adaptive margin policy \(m_j(t)\) for each non-Caucasian group to minimize the skewness of angles between Caucasians and non-Caucasians. We make Caucasians as the benchmark (anchor) here because white subjects are overwhelmingly dominant in numbers in existing FR datasets. According to different racial distribution, our RL-RBN can change the benchmark and replace Caucasians by the most representative ethnic group.

### 4.2. Adaptive margin policy learning

In our method showed in Fig. 3, the problem of finding \(m_j(t)\) can be formulate as a Markov Decision Process (MDP) [28], based on which we use deep Q-learning to adjust the margin at each iteration. Deep Q-learning aims to enable the agent to decide the behavior from its experiences using a MDP. At each time step \(t \in [1, T]\), the agent takes an action \(a^t\) from action space \(A\) according to the Q-value \(Q(s^t, a)\) estimated by the deep Q-learning network (DQN) with the state \(s^t\) as input. The environment gives a reward \(r(s^t, a^t)\) for this action and then the agent updates its states with \(s^{t+1}\). According to this new state, the agent will go to next time step and choose a new action. The goal of deep Q-learning is to train an agent with policy \(\pi\) to maximize the expected sum of rewards.

**State.** The state \(s\) of the MDP consists of three separate parts \(\{G, M, B_{inter}\}\). \(G = \{0, 1, 2\}\) represents the race group, i.e. Indian (group 0), Asian (group 1) and African (group 2). \(M\) is equivalent of the adaptive margin \(m\). \(B_{inter}\) means the bias (skewness) of inter-class distance between \(g\)-th race group and Caucasians, which can be formulated as follows:

\[
B_{inter} = \left| d^g_{inter} - d^{Cau}_{inter} \right|
\]

where, \(d^g_{inter} = \frac{1}{N_g} \sum_{i=1}^{N_g} \max_{k=1: N_g, k \neq i} \cos(c_k, c_i)\)

(4)

where \(d^g_{inter}\) and \(d^{Cau}_{inter}\) is the inter-class distance of \(g\)-th race group and that of Caucasian, respectively. \(N_g\) is the number of identities belonged to \(g\)-th race group and \(c_i\) is the feature centre of \(i\)-th identity computed by the mean vector of embeddings. \(\cos(\cdot, \cdot)\) is the cosine distance function. \(d^{Cau}_{inter}\) can be computed by the same way as \(d^g_{inter}\). We suppose that different races have different demands for the margins, and the demands may change according to their \(B_{inter}\). When the bias (skewness) is larger, the race group may need larger margin to improve its generalization ability, and vice versa. Therefore, we take both \(G\) and \(B_{inter}\) into consideration when designing state \(s\). Each action will be chosen based on race group and the current skewness. Moreover, in order to make the space of states discrete, we map \(M\) and \(B_{inter}\) to discrete spaces \(\mathcal{M}\) and \(\mathcal{B}\), where \(\mathcal{M} = \{m_1, m_2, ..., m_{m_M}\}\) and \(\mathcal{B} = \{b_1, b_2, ..., b_{m_B}\}\).

**Action.** The action \(A = \{0, 1, 2\}\) is the adjustment of margin. We define 3 types of action as ‘staying the same’ (action 0), ‘shifting to larger value’ (action 1) and ‘shifting to smaller value’ (action 2), and shifting step is set to be a constant \(\epsilon\). The optimal action taken by the agent at time step \(t\) is formulated by \(a^t = \arg\max_a Q(s^t, a)\), where the
Q-value $Q(s^t, a)$ is the accumulated rewards of taking the action $a$. For example, at time step $t$, the agent chooses to take the action ‘1’ according to Q-value and state $s^t = \{0, m_2, b_1\}$, then the margin of Indian will update to $m_3 = m_2 + \epsilon$.

**Reward.** The reward, as a function $r(s^t, a^t)$, reflects how good the action taken by the agent is with regard to the state $s^t$. Since we suppose non-Caucasians should have the same generalization ability as Caucasians and the skewness between them should be minimized, we use the skewness of inter/intra-class distance between them to design the reward. The bias (skewness) of intra-class distance between Caucasians and $g$-th group can be formulated as:

$$B_{\text{intra}} = |d_{\text{intra}}^g - d_{\text{intra}}^{\text{Cau}}|$$

where, $d_{\text{intra}}^g = \frac{1}{N_g} \sum_{i=1}^{N_g} \sum_{x_j \in I_i} \cos(x_j, c_i)$ (5)

where $d_{\text{intra}}^g$ and $d_{\text{intra}}^{\text{Cau}}$ is the intra-class distance of $g$-th race group and that of Caucasian, respectively. $N_g$ is the number of identities belonged to $g$-th race group. $I_i$ is the set of all images in $i$-th identity, and $c_i$ is the centre of $i$-th identity computed by the mean vector of embeddings. And the bias (skewness) of inter-class distance can be computed by Eqn. 4. When the agent takes the action $a^t$ to adjust the margin of $g$-th race group, the reward of the action $a^t$ is computed by:

$$r(s^t, a^t) = R^{t+1} - R^t$$

where, $R = -(B_{\text{inter}} + B_{\text{intra}})$ (6)

**Objective function.** We choose to use deep Q-Learning [34, 51] to learn an optimal policy for agent. A two-layer fully connected network with a further hidden layer of 10 units is utilized to estimate the Q function. Each fully connected layer is followed by a ReLU activation function. The deep Q-learning network takes the state as input and produces the Q-value of all possible actions. We update the network by the minimizing the following loss function:

$$L_g = \mathbb{E}_{s^t, a^t} \left[ (y^t - Q(s^t, a^t))^2 \right]$$

where, $y^t = \mathbb{E}_{s^{t+1}} \left[ r^t + \gamma \max_{a^{t+1}} Q(s^{t+1}, a^{t+1} | s^t, a^t) \right]$ (7)

where $y^t = Q(s^t, a^t)$ is the temporal difference error. $y^t$ is the target value of $Q(s^t, a^t)$. $r^t$ is the reward of taking the action $a^t$, computed by Eq.6, and $\max_{a^{t+1}} Q(s^{t+1}, a^{t+1} | s^t, a^t)$ is the future reward estimated by the current deep Q-learning network with $s^{t+1}$.

To train DQN, we collect some offline samples in advance, i.e. \{(s^t, a^t, r^t, s^{t+1})\}, as input to feed network. These offline samples are generated by a sample CNN as shown in Algorithm 1. For each non-Caucasian group, we manually adjust the margin by actions $a^t$ and train the current sample network for one epoch with new margin. After one-epoch training, we compute the intra-class and inter-class distance of this race group and obtain the next state $s^{t+1}$ and reward $r^t$. We keep doing this until all states have been traversed. Then, we train DQN by these collected samples. After that, the adaptive margin policy can be generated according to the output of trained DQN, i.e. Q-value.

**Algorithm 1 Offline sampling.**

**Input:**

The unbalanced data with four race groups.

**Output:**

The samples, i.e. \{(s^t, a^t, r^t, s^{t+1})\}, which are used for training deep Q-learning network.

1. **for** $g$ **in** all groups **do**
2. **for** $a$ **in** all actions **do**
3. Compute $B_{\text{inter}}^t$ of $g$-th group by Eqn. 4, and obtain current state $s^t = \{G, M^t, B_{\text{inter}}^t\}$.
4. Take the action $a$ to adjust margin for $g$-th group.
5. Train the sample network for an epoch with the updated margin $M^{t+1}$ of $g$-th group; while margins of other groups remain unchanged.
6. Compute $B_{\text{inter}}^{t+1}$ of $g$-th group according to Eqn. 4, and obtain updated state $s^{t+1} = \{G, M^{t+1}, B_{\text{inter}}^{t+1}\}$.
7. Compute reward $r^t$ according to Eqn. 6.
8. Collect the sample \{(s^t, a^t, r^t, s^{t+1})\}.
9. **end for**
10. **if** $s^{t+1}$ is a new state that hasn’t appeared **then**
11. **Go** to step 2.
12. **end if**
13. **end for**

**Recognition network.** We utilize the policy to guide the training process of our recognition network. At each time step $t$, $s^t$ is computed by recognition network according to Eqn. 4 and sent to the agent, then the trained agent will take action $a^t$ to adjust the margins for each non-Caucasian group. Then, we update the margin of non-Caucasians in Eqn. 3, and use it to optimize the recognition network for one epoch. We keep doing this until the end of training.

5. Experiments

5.1. Experimental settings

**Datasets.** Because existing datasets are not race-aware except RFW [50], we use our BUPT-Globalface and BUPT-Balancedface datasets to train our models and use RFW [50] to fairly measure performance of different races. RFW [50] consists of four testing subsets, namely Caucasian, Asian,
Indian and African. Each subset contains about 10K images of 3K individuals for face verification. Moreover, in order to evaluate the generalization ability accurately, we additionally construct a validation set to compute intra-class and inter-class distance in RL instead of directly using training sets. The validation set contains 500 identities per race, and has no overlapping subjects with BUPT-Globalface, BUPT-Balancedface and RFW dataset [50].

**Evaluation Protocol.** Verification performance is measured by accuracy. We utilize average accuracy of four races as metric to evaluate total performance of deep models. And standard deviation (STD) and skewed error ratio (SER) are used as the fairness criterion. Standard deviation reflects the amount of dispersion of accuracies of different races. Error skewness is computed by the ratio of the highest error rate to the lowest error rate among different races. It can be formulated as \( SER = \frac{\max Error_g}{\min Error_g} \) where \( g \in \{ \text{Caucasian,Indian,Asian,African} \} \) means race group.

**Implementation details.** For preprocessing, we use five facial landmarks for similarity transformation, then crop and resize the faces to 112 × 112. Each pixel ([0, 255]) in RGB images is normalized by subtracting 127.5 and then being divided by 128. MxNet [11] is utilized to implement an adaptive margin loss and TensorFlow [4] is for Deep Q-learning. We use ResNet34 [19] for training recognition CNN and sampling CNN. For the case of training recognition CNN, the learning rate is started from 0.1 and decreased with a factor of 10 when errors plateau. The batch size is set as 256 for the small training dataset and 200 for the large one. We set momentum as 0.9 and weight decay as \( 5e^{-4} \). The architecture of DQN is mentioned in Section 4.2. The AdamOptimizer is used to optimize the whole network. The learning rate is set to be \( 1e^{-4} \) and the discount factor \( \gamma \) is set to be 0.99.

First, we collect training samples \( \{ (s^t, a^t, r^t, s^{t+1}) \} \) for deep Q-learning. In RL-RBN(soft), the margin of Caucasians is set to be 0 similar to N-Softmax [47], and the margin of non-Caucasians varies from 0 to 0.6 in steps of 0.2 in which an additive angular margin is used similar to Arcface [12]; In RL-RBN(cos) and RL-RBN(arc), the margin of Caucasians is set to be 0.15 and 0.3, and the margin of non-Caucasians varies from 0.15 to 0.45 in steps of 0.1 and varies from 0.3 to 0.6 in steps of 0.1, respectively. After that, we feed the samples \( \{ (s^t, a^t, r^t, s^{t+1}) \} \) to train deep Q-learning network and then generate an adaptive margin policy. Finally, the recognition CNN is trained with the guidance of this policy.

### 5.2. Cause of racial bias

Some papers [50, 54, 25] verified that non-Caucasians still perform poorly than Caucasians even with balanced training and faces of colored skin are inherently difficult to recognize for existing algorithms. In order to go deep into this phenomenon, we degrade images of RFW [50] by blur and noise, and observe the influence of these image degradations on the performance of Caucasians and Africans. To apply Gaussian blur, the sigma value of the Gaussian filter is 1.5 and kernel size varied from 1 to 10 in steps of one (Fig. 4). And we add noise to images by using Gaussian noise with zero mean and varying standard deviation from 5 to 50 in steps of 5. We train two ResNet-34 models with guidance of Softmax and Arcface loss [12] using BUPT-Balancedface dataset, and test them on the blurred and noisy RFW [50]. The accuracies of Caucasians and Africans are given in Fig. 5. We can see that the performance gap between Africans and Caucasians still exists even with balanced training. And both Caucasians and Africans are found to be sensitive to image blur and noise. More importantly, when blur and noise level increases, the performance gap between Caucasians and Africans widens. Therefore, we conclude that colored faces are more susceptible to noise and image quality than Caucasians. This may be one of the reasons why non-Caucasians are more difficult to recognize.

### 5.3. Experiments of our method

**Results on simulated dataset.** As we know, data bias in training set severely affects fairness of algorithms. In order to validate the effectiveness of our RL-RBN, we train our al-
Table 2. Verification accuracy (%) on RFW [50] trained with varying racial distribution. We boldface STD (lower is better) and skewed error ratio (SER) (1 is the best) since this is the important fairness criterion.

<table>
<thead>
<tr>
<th>Train</th>
<th>Test →</th>
<th>Method</th>
<th>Caucasian</th>
<th>Indian</th>
<th>Asian</th>
<th>African</th>
<th>Avg</th>
<th>Fairness</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 : 2 : 2 : 2</td>
<td>N-Softmax [47]</td>
<td>89.67</td>
<td>87.97</td>
<td>84.68</td>
<td>84.17</td>
<td>86.62</td>
<td>2.64</td>
<td>2.53</td>
</tr>
<tr>
<td></td>
<td>RL-RBN(soft)</td>
<td>91.35</td>
<td>90.77</td>
<td>89.87</td>
<td>90.13</td>
<td>90.53</td>
<td>0.66</td>
<td>1.17</td>
</tr>
<tr>
<td>5 : 4 : 1 : 2</td>
<td>N-Softmax [47]</td>
<td>89.88</td>
<td>88.52</td>
<td>85.13</td>
<td>83.42</td>
<td>86.74</td>
<td>2.08</td>
<td>1.64</td>
</tr>
<tr>
<td></td>
<td>RL-RBN(soft)</td>
<td>90.33</td>
<td>90.23</td>
<td>88.97</td>
<td>89.37</td>
<td>89.73</td>
<td>0.67</td>
<td>1.22</td>
</tr>
<tr>
<td>6 : 5 : 4 : 1</td>
<td>N-Softmax [47]</td>
<td>90.43</td>
<td>88.32</td>
<td>84.75</td>
<td>83.32</td>
<td>86.70</td>
<td>3.26</td>
<td>1.74</td>
</tr>
<tr>
<td></td>
<td>RL-RBN(soft)</td>
<td>90.17</td>
<td>90.02</td>
<td>87.67</td>
<td>88.27</td>
<td>89.03</td>
<td>1.25</td>
<td>1.25</td>
</tr>
<tr>
<td>7 : 1 : 1 : 1</td>
<td>N-Softmax [47]</td>
<td>90.67</td>
<td>87.77</td>
<td>84.37</td>
<td>82.97</td>
<td>86.44</td>
<td>3.46</td>
<td>1.83</td>
</tr>
<tr>
<td></td>
<td>RL-RBN(soft)</td>
<td>90.63</td>
<td>90.73</td>
<td>87.72</td>
<td>87.53</td>
<td>89.15</td>
<td>1.77</td>
<td>1.35</td>
</tr>
</tbody>
</table>

Table 3. Verification accuracy (%) of our policy on RFW [50] ([BUPT-Globalface, ResNet34, loss*]). M-RBN is the method using different fixed margins for different races inversely proportional to their number.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Caucasian</th>
<th>Indian</th>
<th>Asian</th>
<th>African</th>
<th>Avg</th>
<th>Fairness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Softmax</td>
<td>95.62</td>
<td>91.97</td>
<td>90.85</td>
<td>89.98</td>
<td>92.10</td>
<td>2.48</td>
</tr>
<tr>
<td>M-RBN(soft)</td>
<td>93.50</td>
<td>94.50</td>
<td>90.06</td>
<td>93.43</td>
<td>92.83</td>
<td>1.90</td>
</tr>
<tr>
<td>M-RBN(cos)</td>
<td>94.53</td>
<td>95.03</td>
<td>94.20</td>
<td>94.05</td>
<td>94.45</td>
<td>0.44</td>
</tr>
<tr>
<td>M-RBN(cos)</td>
<td>96.63</td>
<td>94.68</td>
<td>93.50</td>
<td>92.17</td>
<td>94.25</td>
<td>1.90</td>
</tr>
<tr>
<td>M-RBN(soft)</td>
<td>96.15</td>
<td>95.73</td>
<td>93.43</td>
<td>94.76</td>
<td>95.02</td>
<td>1.21</td>
</tr>
<tr>
<td>M-RBN(soft)</td>
<td>96.03</td>
<td>95.15</td>
<td>94.58</td>
<td>94.27</td>
<td>95.01</td>
<td>0.77</td>
</tr>
<tr>
<td>M-RBN(soft)</td>
<td>97.37</td>
<td>95.68</td>
<td>94.55</td>
<td>93.87</td>
<td>95.37</td>
<td>1.53</td>
</tr>
<tr>
<td>M-RBN(soft)</td>
<td>97.03</td>
<td>95.58</td>
<td>94.40</td>
<td>95.18</td>
<td>95.55</td>
<td>1.10</td>
</tr>
<tr>
<td>M-RBN(soft)</td>
<td>97.08</td>
<td>95.63</td>
<td>95.57</td>
<td>94.87</td>
<td>95.79</td>
<td>0.93</td>
</tr>
</tbody>
</table>

From the results, we can see several important observations. First, it also shows that racial bias indeed exists in existing algorithms. For example, when the racial distribution is 4:2:2:2, the accuracy of Norm-Softmax reaches 89.69% on Caucasian testing subset, but its accuracy dramatically decreases to 84.17% on African subset. Second, the results quantitatively verify our thought that the accuracy of each race is positively correlated with its number in training set. For example, increasing ratio of Caucasians (from 2/5 to 7/10) increases its accuracy from 89.67% to 90.67% in Norm-Softmax. Also, with the change of distribution, we observe a decrease in fairness between races. Third, after adopting adaptive margin loss guided with R-L, our RL-RBN(soft) significantly obtains more balanced performance than Norm-Softmax on different races. When racial distribution becomes more uneven, i.e. 7:1:1:1, our method can still perform better and decrease the SER from 1.83 to 1.35.

Results on BUPT-Globalface dataset. Training on BUPT-Globalface, we compare our RL-RBN with Softmax,

Cosface [48] and Arcface [12]. The scaling parameter is set to 60 and the margin parameters are set as 0.2 and 0.3 for Cosface [48] and Arcface [12], respectively. We show the results in Table 3 and Fig. 6. First, our RL-RBN(soft) obtains more perfect performance than Softmax. It achieves about 2.35% gains for average accuracy, and STD decreases from 2.48 to 0.44. Second, we find that large margin loss, i.e. Cosface [48] and Arcface [12], can alleviate racial bias to some extent through more separate inter-class. However, racial bias cannot be eliminated completely. Third, our RL-RBN(cos) and RL-RBN(soft) can find an optimal margin for each race group and obtain more balanced performance than Cosface and Arcface. It shows the superiority of our algorithm on learning balanced features from a biased dataset.

Results on BUPT-Balancedface dataset. We also com-


<table>
<thead>
<tr>
<th>Methods</th>
<th>Caucasian</th>
<th>Indian</th>
<th>Asian</th>
<th>African</th>
<th>Avg</th>
<th>Fairness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>STD</td>
</tr>
<tr>
<td>Triplet [42]</td>
<td>94.58</td>
<td>93.17</td>
<td>91.48</td>
<td>91.60</td>
<td>92.71</td>
<td><strong>1.47</strong></td>
</tr>
<tr>
<td>Softmax</td>
<td>94.18</td>
<td>92.82</td>
<td>91.23</td>
<td>91.42</td>
<td>91.41</td>
<td><strong>1.38</strong></td>
</tr>
<tr>
<td>RL-RBN(soft)</td>
<td>94.30</td>
<td>94.13</td>
<td>93.87</td>
<td>94.45</td>
<td>94.28</td>
<td><strong>0.20</strong></td>
</tr>
<tr>
<td>Cosface [48]</td>
<td>95.12</td>
<td>93.93</td>
<td>92.98</td>
<td>92.93</td>
<td>93.74</td>
<td><strong>1.03</strong></td>
</tr>
<tr>
<td>RL-RBN(cos)</td>
<td>95.47</td>
<td>95.15</td>
<td>94.52</td>
<td>95.27</td>
<td>95.10</td>
<td><strong>0.41</strong></td>
</tr>
<tr>
<td>Arcface [12]</td>
<td>96.18</td>
<td>94.67</td>
<td>93.72</td>
<td>97.98</td>
<td>94.04</td>
<td><strong>1.11</strong></td>
</tr>
<tr>
<td>RL-RBN(arc)</td>
<td>96.27</td>
<td>94.68</td>
<td>94.82</td>
<td>95.00</td>
<td>95.19</td>
<td><strong>0.93</strong></td>
</tr>
</tbody>
</table>

Table 4. Verification accuracy (%) of our policy on RFW [50] ([BUPT-Balancedface, ResNet34, loss*]).

Figure 8. Distribution of margins of non-Caucasians in RL-RBN(soft) and RL-RBN(arc) training on BUPT-Balancedface.


distributions of margins of non-Caucasians in RL-RBN(soft) and RL-RBN(arc) when trained on BUPT-Balancedface dataset. Guided by the agent, the margins of Asians and Africans are indeed larger than those of Indians, especially the Asians who are the most difficult to recognize when trained with Softmax or Arcface. Moreover, RL-RBN(arc) usually selects larger margins for non-Caucasians compared with RL-RBN(soft). This is because the performance of Caucasians is set to be the anchor and the performances of other races are improved to get close to the anchor in our method. In RL-RBN(soft), the margin of Caucasians is small and fixed at 0, overlarge margins of non-Caucasians will lead to out-of-balance performance again. This result proves the robustness and adaptability of our method.

**Compared with manual margin.** We also compare our method with manual-margin based RBN (M-RBN). The M-RBN simply sets different fixed margins for different races inversely proportional to the number of their samples. From Table 3, we can see that the performance of Asians is always a drag on fairness in M-RBN and our method is superior to M-RBN in fairness. This is because racial bias is a complex problem in which the number is not only fact affecting out-of-balance accuracy. Although the number of Asians is much larger than that of Indians and Africans in BUPT-Globalface dataset, this group still needs a larger margin because it is the most difficult race to recognize even with balanced training.

6. Conclusion

In this paper, we provide two ethnicity aware training datasets, i.e. BUPT-Globalface and BUPT-Balancedface. Then, a reinforcement learning based race-balance network is proposed to alleviate racial bias and learn more balanced features. It introduces the Markov decision process to adaptively find optimal margins for non-Caucasians. The comprehensive experiments prove the effectiveness of RL-RBN.

7. Acknowledgments

This work was partially supported by the National Natural Science Foundation of China under Grant Nos. 61871052 and 61573068.
**References**


[41] James Zou and Linda Schoibinger. Ai can be sexist and racists time to make it fair, 2018.