2023 IEEE 22nd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom) | 979-8-3503-8199-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/TrustCom60117.2023.00037

Big Data Assisted Object Detection with Privacy Protection

JianLin Zhang
College of Computer and Cyber Security
Fujian Normal University
FuZhou, China
zjl578441403@163.com

XiaoDing Wang*
College of Computer and Cyber Security
Fujian Normal University
FuZhou, China
wangdin1982@fjnu.edu.cn

Hui Lin
College of Computer and Cyber Security
Fujian Normal University
FuZhou, China
linhui@fjnu.edu.cn

***Abstract*—The issues of privacy and bias in datasets are rapidly becoming important challenges that the computer vision field needs to address. So far, there has been little attention paid to solutions for protecting the privacy of new datasets. In our work, we explored a object detection solution on the WIDER FACE dataset by anonymizing the dataset using face synthesis and enhancing the WIDER FACE dataset by balancing facial features along the dimensions of gender and skin color. Using both the original dataset and our enhanced dataset to train the target detection model, our target detection results show that our model can maintain detection performance while preserving privacy and partially balancing bias.**

***Index Terms*—Face Detection,Privacy Protection,Machine Vision**

I. Introduction

Neural network-based methods have made great breakthroughs in recent years. For example, chatgpt has developed rapidly in the field of text communication, the StyleGan [10]–[12] algorithm , and the stable diffusion algorithm have achieved great success in the field of AI image generation. Great results. These achievements benefit from better architectures, but on the other hand, they are also due to the ever-increasing dataset size [34].

With the addition of big data technology, target detection algorithms based on deep learning [36], [37], such as face recognition and automatic driving [33], have stronger robustness. Through big data, these deep learning methods can continuously improve their capabilities [35]. An expert in the field of AI, has stated that the future development path of deep learning should change from using big data to train models to using high-quality data created by methods such as data enhancement and synthetic data. However, how to ensure the privacy of big data has always been a problem that machine learning workers have been considering [29].

Researchers have identified two significant issues with both currently used datasets and newly proposed datasets. The first issue is the existence of serious data biases, as discovered by researchers such as [2], [25], who found that children and minorities are significantly underrepresented in common datasets. This bias in the dataset can be reflected in the model, resulting in biased output results. For instance, the Google image [20] labeling algorithm mistakenly labeled two Black people as gorillas, and facial analysis models have lower accuracy for minority women's faces than for White men's faces [2].

The second issue is the unauthorized use of images for training machine learning models, as described in [17]. While these data images were collected under a Creative Commons license, the license does not provide or specify any information regarding their use in training artificial intelligence models. This may potentially violate the license, as the trained models can be explored to reveal the entire training sample, such as addresses and bank accounts in GPT-2 [3], or possibly images [16]. Therefore, it is necessary to remove personal privacy information from the images. This not only protects personal information in the images from being stolen, but also strictly enforces national information security laws such as the Personal Information Protection Law and the Data Security Law.

One common variable in these two issues is the use of raw, unedited data, which often either violates privacy or contains biases. However, because collecting unbiased datasets with consistent consent of individuals is often very expensive or even unfeasible [25], in this paper, we investigate methods to mitigate these issues by modifying the training data. In summary, our contributions are as follows:

- We study measures to preserve data privacy while maintaining the performance of face object detection.
- We propose a novel approach to balance race, gender bias in training datasets while removing personally identifiable information.
- We measure the bias of our model with the image association test.

*Corresponding author

2324-9013/23/$31.00 ©2023 IEEE
DOI 10.1109/TrustCom60117.2023.00037

Authorized licensed use limited to: Fujian Normal University. Downloaded on June 12,2024 at 07:15:05 UTC from IEEE Xplore. Restrictions apply.

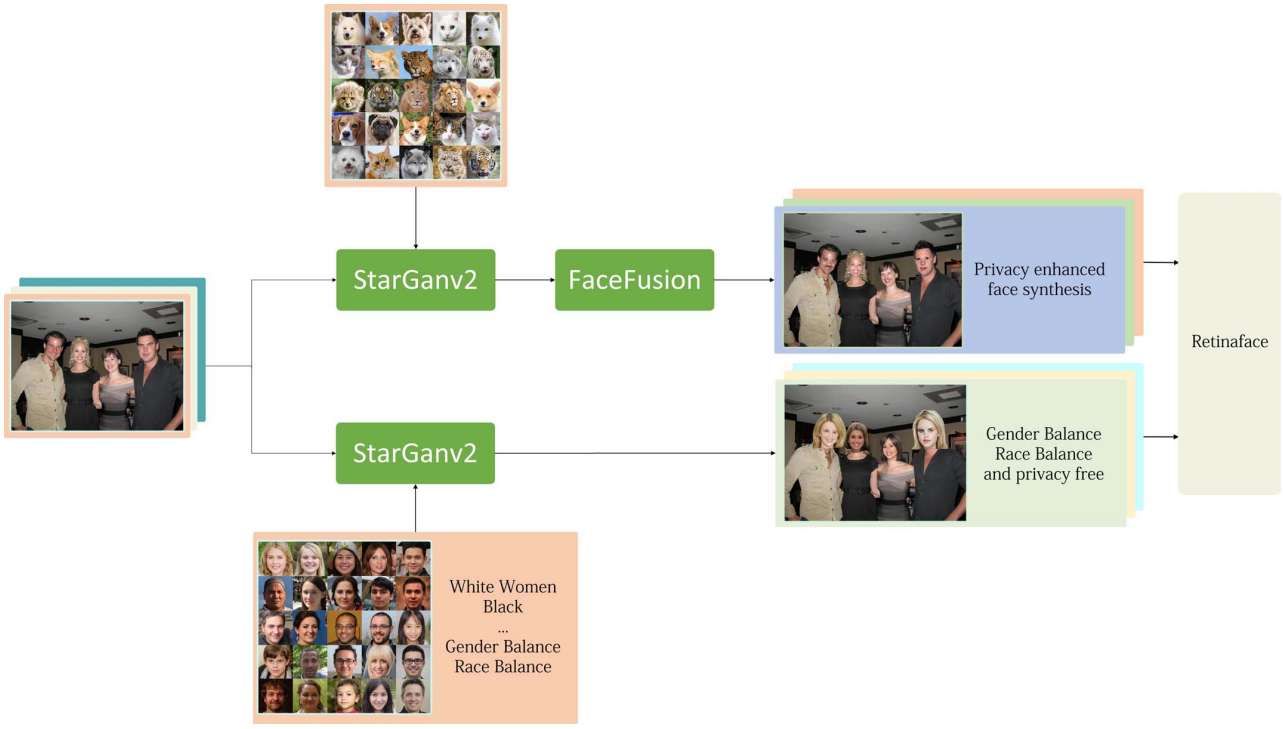


Fig. 1. Schematic illustration of the privacy-preserving approach we study. On the left is an example picture we extracted from the WIDER FACE dataset, the animal face dataset above is from StarGanv2, and the bottom is the face image we generated with StyleGan3. We use two face-changing methods to erase the privacy of some pictures in the WIDER FACE dataset, and then add the generated pictures to the dataset and use the expanded dataset to train the face detection algorithm.

II. Related Work

A. Measurement bias

According to experiments, many models have biases. For example, [2] showed that facial recognition tools have lower accuracy for women of color than for white men. Openly available models have been shown to have biases towards minority ethnic features [13], [20]. Furthermore, [21] found that unsupervised models trained on ImageNet contain intersectional biases related to race and gender. We investigated biases in our model using their test. To mitigate biases, [27] and others developed a method to eliminate the impact of bias in the collected dataset before conducting statistical analyses.

According to [23] and others, even with a balanced dataset, models can amplify implicit gender bias. They suggest using a generative adversarial network approach to remove personal information by obscuring private areas of images. [8] explored how to effectively transform input data to balance datasets while eliminating potential data biases. Our method uses non-sensitive facial images generated by StyleGAN3 [10] to replace the private facial information in the training dataset and balance the dataset bias.

B. Preserving Privacy

Even that image recognition algorithms inevitably expose some identifiable and sensitive elements, privacy protection is crucial in computer vision. Various methods have been developed to address privacy issues in computer vision models, such as face occlusion [22], reduced resolution quality [19], and others [18], [24]. [15] demonstrated how targeted blurring can be applied in privacy-sensitive areas to protect privacy while maintaining image utility. [8] evaluated the performance of privacy protection in modern object detection algorithms. In our approach, we use synthetic animal and human facial images to obscure sensitive information in the dataset and evaluate the performance using advanced face detection methods.

C. Synthesis of faces

The transfer of human faces with similar poses to target images is also a popular research topic in the field of computer vision. [28] pasted celebrity faces onto other images to generate new composite faces. [4], [5] significantly improved the effectiveness of face transfer by learning mappings between different visual domains, and their approach outperforms other algorithms in terms of visual quality and diversity.

D. Mitigating bias

In [23], the authors found that when using adversarial loss to enhance the model architecture to obscure parts of images that leak protected features (such as race and gender), the multi-label prediction performance on COCO only slightly decreased.

Authorized licensed use limited to: Fujian Normal University. Downloaded on June 12,2024 at 07:15:05 UTC from IEEE Xplore. Restrictions apply.

(a) . Original (b) . Animal synthesis (c) . Gender balance (d) . Race balance

Fig. 2. Example image. 4 versions of the same WIDER FACE image, each taken from one of our datasets. (a) is the original image, (b) is the image synthesized by animals, (c) (d) is the gender-balanced and race-balanced image synthesized with the AI image.

E. Face Detection

Despite the significant progress made in single-stage face detection, accurately and efficiently locating faces in the wild remains an open challenge. Retinaface [6] proposed a robust face detector that leverages the benefits of joint extra-supervision and self-supervised multi-task learning to perform pixel-level face localization on faces of various scales.

III. Methods

This article explores protecting data privacy and ensuring a balanced model bias without compromising face detection efficiency. We used StarGanv2 [5] to perform image facial synthesis transformations on the Wilder Face [26] dataset while maintaining detection performance. These modified images and the original Wilder Face dataset were used to train RetinaFace. The model's target detection performance was measured on both the modified and original datasets. Finally, we attempted to measure the representational bias of all fine-tuned Resnet50 backbones using image embedding correlation testing.

A. Synthesis of animal faces

The original WIDER FACE dataset contains 32203 training images, 393703 labels, and we extracted 2600 of them as the standard dataset. We developed a model for face image translation using StarGanv2.Our model is illustrated in Figure 3, where the mapping network or style encoder provides a specific domain style code S, which is injected into the generator via ADAIN [9]. We sample latent codes from a standard Gaussian distribution and input them into an MLP to generate the style code. The style encoder transforms the input style image into a style code through CNN. To generate face images, a generator G takes an image x and a style code $\tilde{s}$ as input, and learns to generate an output image G(x, $\tilde{s}$) through an adversarial loss:

$$\begin{aligned}\mathcal{L}_{adv} =& \mathbb{E}_{x,y}[\log D_y(X)]+ \\ & \mathbb{E}_{x,\tilde{y},z}[\log(1 - D_{\tilde{y}}(G(x,\tilde{s}))] \end{aligned} \tag{1}$$

In order to use the style code $\tilde{s}$ when generating images, we employ a style reconstruction loss:

$$\mathcal{L}_{sty} = \mathbb{E}_{x,\tilde{y},z}[\|\tilde{s} - E_{\tilde{y}}(G(x,\tilde{s}))\|] \tag{2}$$

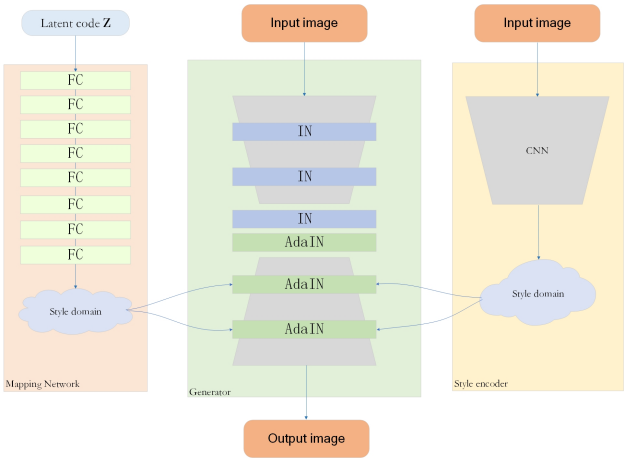


Fig. 3. Our model summary: The generator transforms an input image into an output image that reflects the domain-specific style code through the instance normalization (IN) [32] downsampling and and uses adaptive instance normalization (AdaIN) upsampling to output the image. The style code is injected into all AdaIN layers. The mapping network converts a latent code into style codes for multiple domains through MLP, randomly selecting one during training. The style encoder extracts the style code of a style image through CNN, and the generator performs style code synthesis to generate the image. The discriminator distinguishes real and fake images from the generator. Note that all modules, except the generator, have multiple output branches, with one selected during training for the corresponding domain.

And a style diversity loss to enhance the diversity of generated images:

$$\mathcal{L}_{ds} = \mathbb{E}_{x,\tilde{y},z_1,z_2}[\|(G(x,\tilde{s_1})) - (G(x,\tilde{s_2}))\|] \tag{3}$$

At the same time, in order to ensure that the original features such as face pose can be maintained in the generated image, we use the cycle consistency loss:

$$\mathcal{L}_{cyc} = \mathbb{E}_{x,y,\overline{y},z}[\|x - G(G(x,\tilde{s_2}),\hat{s})\|] \tag{4}$$

Among them, $\hat{s}$ is the style code of the original image.The overall objective function can be summarized as:

$$\min_{G,F,E} \max_{D} \mathcal{L}_{adv} + \lambda_{sty}\mathcal{L}_{sty} - \lambda_{ds}\mathcal{L}_{ds} + \lambda_{cyc}\mathcal{L}_{cyc} \tag{5}$$

Where $\lambda_{sty}, \lambda_{ds}, \lambda_{cyc}$ are the hyperparameters of each item, and they are all set to 1 in our training model.

Authorized licensed use limited to: Fujian Normal University. Downloaded on June 12,2024 at 07:15:05 UTC from IEEE Xplore. Restrictions apply.

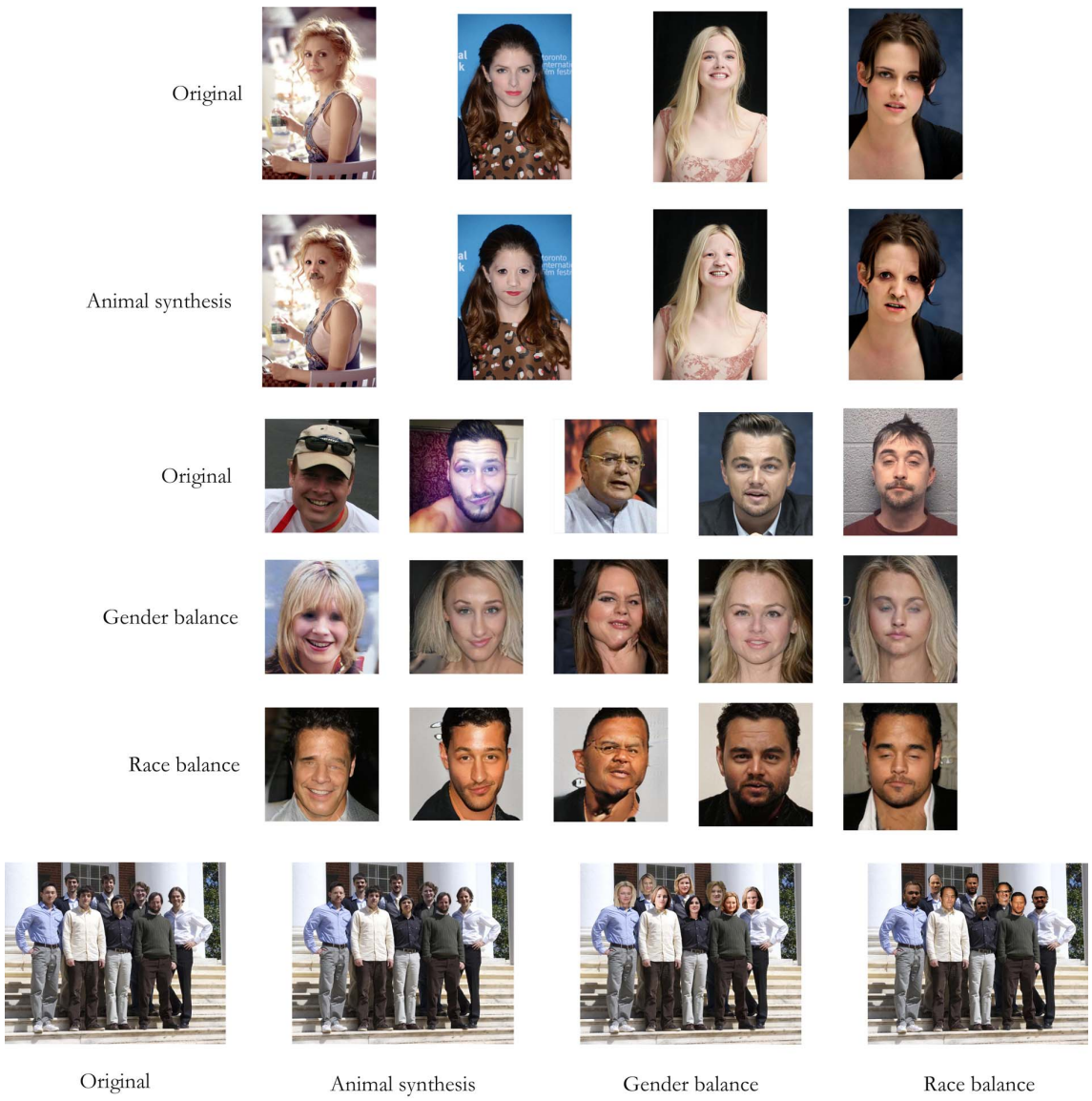


Fig. 4. Some image processing examples in our extended dataset.

We use the CelebA [14] and animal face [5] datasets to train our animal face synthesis model on StarGanv2, and then cut out the faces in the WIDER FACE dataset and use StarGanv2 to synthesize the animal face images. A synthetic face image is generated. And use [1]'s face fusion algorithm to fuse the synthetic face with the face in the data set into an image that does not contain the face privacy of the original image. Finally, the generated image is replaced with the corresponding image in the standard dataset to complete the face synthesis dataset.

B. Face Transformation

We modified the WIDER FACE dataset to synthesize face images without privacy and scramble the skin color of the face. In order to synthesize images, we use the pre-trained StyleGan3 network to generate AI faces and delete 1000 images with different skin colors and genders for backup, and then use StarGanv2 to synthesize the generated faces with the cut face images in WIDER FACE to generate Face images without personal privacy. We select photos of different ethnic and gender characteristics in the images generated by StyleGan3 to synthesize the faces in WIDER FACE to achieve the purpose of balancing race and gender.This potentially increases workload but reduces privacy breaches and biases.

From the generated image, we further select the face with better facial features and distort it so that its label is aligned with the face in the original image, and finally paste the generated face into the original image. Finally, the edited images are

Authorized licensed use limited to: Fujian Normal University. Downloaded on June 12,2024 at 07:15:05 UTC from IEEE Xplore. Restrictions apply.

replaced with the corresponding images in the standard dataset to generate gender-balanced and race-balanced datasets.

For the cross-balanced dataset, we add all images synthesized to the standard dataset and add labels.Figure 2 shows an example of the images included in our dataset, with more examples in Figure 4.

C. Face Detection

We use Retinaface, which is currently a relatively good face detection model, to test the effect of our model. We use our different data sets to train the detectors, get 5 trained ResNet50 networks, and use different networks to run the verification set of WIDER FACE, and the detection results are used as indicators of the respective detection capabilities of the models.

D. Bias measure: iEAT test

To measure the bias in image representations, we conducted tests using iEAT [21], which is adapted from the Implicit Association Test (IAT) in social psychology [7]. The test measures differential associations between certain target concepts (e.g. men, women) and a set of attributes (e.g. career, family) across all images. Firstly, we trained a RetinaFace model with a ResNet50 backbone to extract visual features on different datasets. We calculated the cosine distance between normalized representations of different visual features, and recorded the p-values and effect sizes d of the null hypothesis to test for bias in the model. The specific procedures are as follows:

The test statistic measures the differential association of target concepts X and Y with attributes A and B

$$s(X, Y, A, B) = \sum_{x \in X} s(x, A, B) - \sum_{y \in Y} s(y, A, B) \quad (6)$$

where s(w,A,B) is the differential association of w with the attribute, quantified by the cosine similarity of the vectors

$$s(w, A, B) = mean_{a \in A} \cos(w, a) - \quad mean_{b \in B} \cos(w, b) \quad (7)$$

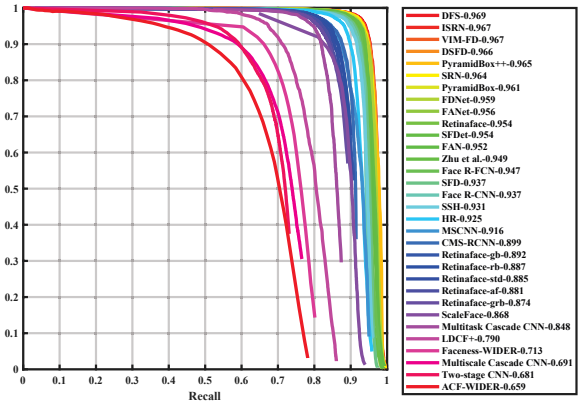
We test the significance of this association with a permutation test on all possible equal-sized partitions $(X_i, Y_i)_i$ of X ∪ Y to generate a null hypothesis , as if there were no biased associations. And use the p value to measure the impossibility of the null hypothesis unilaterally

$$p = Pr[s(X_i, Y_i, A, B) > s(X, Y, A, B)] \quad (8)$$

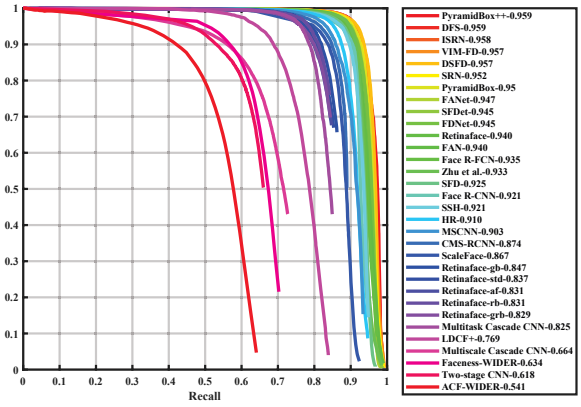
Whereas the effect size, a standardized measure of the split point between the relative associations of X and Y with A and B, can be described as

$$d = \frac{mean_{x \in X} s(s, A, B) - mean_{y \in Y} s(y, A, B)}{std_{w \in X \cup Y} s(w, A, B)} \quad (9)$$

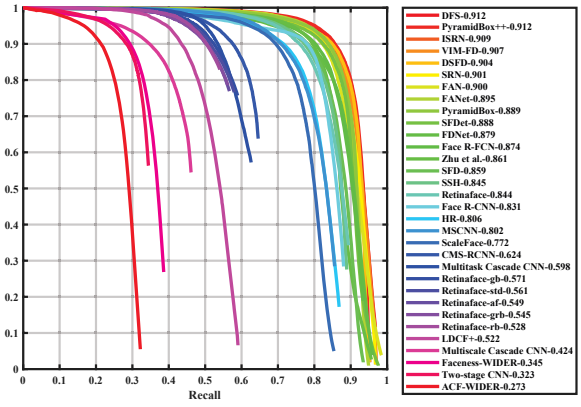
A larger effect size indicates a significant correlation of differences.



(a) easy



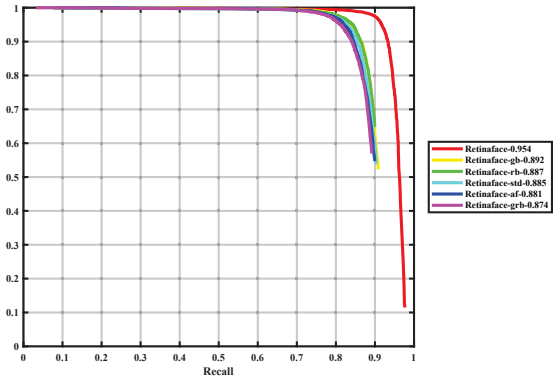
(b) medium



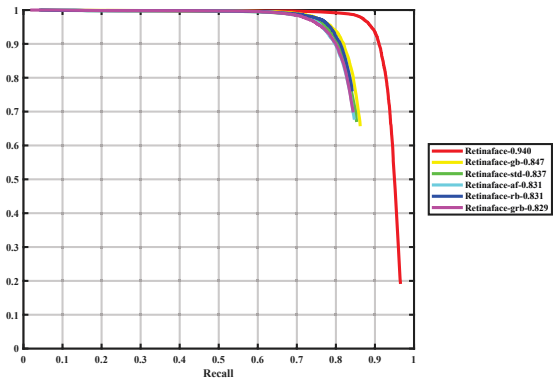
(c) hard

Fig. 5. The detection results of our model on WIDER FACE are compared with other detection models, where -std represents our standard model, -af represents the animal face synthesis model, -rb represents the race balance model, -gb represents the gender balance model, -grb represents our augmented cross-balanced model

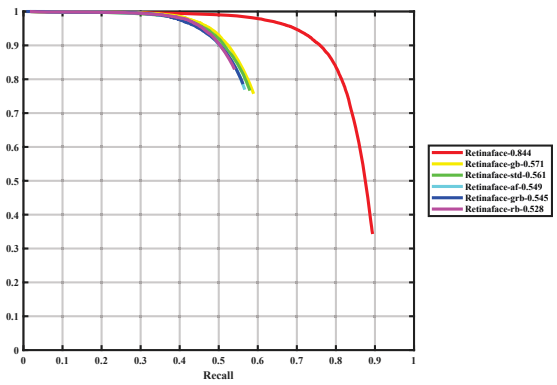
Authorized licensed use limited to: Fujian Normal University. Downloaded on June 12,2024 at 07:15:05 UTC from IEEE Xplore. Restrictions apply.



(a) easy



(b) medium



(c) hard

Fig. 6. Horizontal comparison of the detection results of our model on Retinaface

IV. Result

A. Face detection performance

As shown in Figure 5, although the performance of our model has decreased compared with the original Retinaface, both our standard model and the modified model still outperform some face detection models. From the detection results in Figure 5(c), we can see that our model has insufficient performance in the hard mode, but it should be known that the face is almost invisible in the hard mode, or contains only minimal privacy. Overall, all our models are able to maintain good detection performance.

The horizontal comparison results for all of our models are shown in Figure 6. As shown in Figures 6(b) and 6(c), our standard model Retinaface-std performs the best in the medium and hard modes, but only slightly outperforms our other models. In the easy mode, our modified model has detection performance comparable to the standard model, with no performance loss in evaluating synthetic faces, and even performs better than the standard model in detecting animal faces and achieving race balance. This may be due to the artifacts in the synthetic images. All models trained on the modified dataset perform equally well as our standard model in detecting objects in the WIDER FACE dataset. This suggests that with an appropriate face transformation method, we can train a face detector on datasets where not all images contain real faces, without sacrificing performance on real face images.

B. iEAT bias measurement

The IEAT test measures bias across many variables, but we focus on those we consider to be social biases and only discuss statistically significant results. The smaller the p-value, the stronger the significance, and the larger the d-value, the greater the bias. Comparing gender balance and race balance models to the standard model, bias was reduced by around 25% for ”Arab-Muslim” and by around 10% for ”gender-science” and ”weight”. All other differences were either of negligible size (such as ”religion”) or results where the difference in effect size was not statistically significant (such as in ”age”). Compared to the standard model, cross-balance increased bias in all significant categories (such as ”weight” and ”disability”). Similarly, the animal-synthesis model increased bias in ”disability” and ”Arab-Muslim”, while slightly decreasing bias in the ”gender-science” and ”weight” categories. The results of the IEAT association test show that our modified model slightly reduces bias compared to the standard model.

For completeness, we include results from random initialization as well as a Resnet50 backbone pretrained on Retinaface. While results from random initialization are not interpretable, the results from Retinaface pretrained backbone showed the least bias in the ”weight” and ”disability” categories, and the highest bias in the ”Arab-Muslim” category. Models trained on the standard model showed similar bias results, suggesting that our data preprocessing had a limited impact.

Authorized licensed use limited to: Fujian Normal University. Downloaded on June 12,2024 at 07:15:05 UTC from IEEE Xplore. Restrictions apply.

TABLE I
IEAT RESULT(P/D)

	RI	RP	Std	AS	GB	RB	CB
Insect-Flower	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.899	-0.549	-0.874	-0.981	-0.952	-1.146	-1.107
Gender-Science	0.289	0.168	0.072	0.039	0.053	0.014	0.164
	0.142	0.168	0.471	0.380	0.329	0.358	0.595
Gender-Career	0.106	0.079	0.636	0.578	0.846	0.952	0.734
	0.326	0.289	-0.132	-0.148	-0.124	-0.316	-0.226
Disability	0.567	0.025	0.048	0.037	0.053	0.037	0.084
	-0.483	0.935	1.042	1.043	1.021	1.026	1.089
Asian	0.147	0.288	0.365	0.124	0.016**	0.352	0.221
	0.536	0.462	0.364	0.426	1.062	1.071	0.244
Arab-Muslim	0.198	0.024	0.032	0.023	0.025	0.028	0.187
	0.453	0.326	0.732	0.745	0.637	0.629	0.884
Age	1.043	0.463	0.529	0.529	0.425	0.513	0.572
	-1.245	0.281	0.364	-0.378	-0.236	-0.085	0.373
Weight	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
	1.235	1.116	1.679	1.467	1.316	1.341	1.768
Weapon (Modern)	0.225	0.321	0.216	0.627	0.209	0.746	0.314
	0.351	0.515	0.184	-0.241	0.709	-0.415	0.193
Weapon	0.383	0.852	0.025	0.048	0.424	0.505	0.315
	0.102	1.234	0.609	0.395	0.902	-0.834	0.099
Skin-Tone	0.837	0.798	0.553	0.961	0.125	0.990	0.294
	1.452	0.833	0.617	0.308	0.012	0.133	0.189
Sexuality	0.205	0.782	0.790	0.540	0.413	0.543	0.212
	0.140	-0.687	-0.056	-0.057	-0.033	-0.960	-0.144
Religion	0.147	0.258	0.185	0.277	0.301	0.123	0.124
	0.379	0.310	0.415	0.444	0.323	0.306	0.252
Race	0.018	0.579	0.475	0.285	0.376	0.833	0.357
	1.054	-0.478	-0.429	-0.366	-0.212	-0.258	-0.366
Native	0.415	0.258	0.891	0.347	0.832	0.773	0.632
	0.200	-1.054	-0.436	-0.310	-0.425	-0.235	-0.425

Abbreviations:RI:random initialization,RP:Retinaface pre-training,Std:standard,AS:animal synthesis,GB:gender balance,RB:racial balance,CB:cross balance.
The highlight represents the size of the p-value: p<0.1 ; p<0.05 ; p<0.01

V. CONCLUSIONS

We propose a privacy-enhancing model based on object detection, which solves the problems of privacy leakage and bias in object detection by training object detection algorithms using synthesized face datasets while maintaining detection performance. Experimental results show that the Retinaface optimized for WIDER FACE can be trained with the same performance as the standard model without using real face datasets. Moreover, when using synthesized or edited faces to make the dataset more balanced in terms of gender and race, the detector learns the same things and achieves a more balanced detection performance. We also investigated whether the proposed two metrics could be used to eliminate model bias. Although the results were different from our expectations, the bias was still eliminated for some labels,and more metrics should be used to eliminate model bias.

ACKNOWLEDGMENT

This work is supported by National Natural Science Foundation of China under Grant No. 61702103 and U1905211, Natural Science Foundation of Fujian Province under Grant No. 2020J01167 and 2020J01169.

REFERENCES

[1] damo alibaba: Face fusion, https://www.modelscope.cn/models/damo/
[2] Buolamwini, J., Gebru, T.: Gender shades: Intersectional accuracy disparities in commercial gender classification. In: Conference on fairness, accountability and transparency. pp. 77–91. PMLR (2018)

Authorized licensed use limited to: Fujian Normal University. Downloaded on June 12,2024 at 07:15:05 UTC from IEEE Xplore. Restrictions apply.

[3] Carlini, N., Tramer, F., Wallace, E., Jagielski, M., Herbert-Voss, A., Lee, K., Roberts, A., Brown, T.B., Song, D., Erlingsson, U., et al.: Extracting training data from large language models. In: USENIX Security Symposium. vol. 6 (2021)

[4] Choi, Y., Choi, M., Kim, M., Ha, J.W., Kim, S., Choo, J.: Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 8789–8797 (2018)

[5] Choi, Y., Uh, Y., Yoo, J., Ha, J.W.: Stargan v2: Diverse image synthesis for multiple domains. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 8188–8197 (2020)

[6] Deng, J., Guo, J., Ververas, E., Kotsia, I., Zafeiriou, S.: Retinaface: Single-shot multi-level face localisation in the wild. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 5203–5212 (2020)

[7] Greenwald, A.G., McGhee, D.E., Schwartz, J.L.: Measuring individual differences in implicit cognition: the implicit association test. Journal of personality and social psychology **74**(6), 1464 (1998)

[8] He, P., Griffin, C., Kacprzyk, K., Joosen, A., Collyer, M., Shtedritski, A., Asano, Y.M.: Privacy-preserving object detection. arXiv preprint arXiv:2103.06587 (2021)

[9] Huang, X., Belongie, S.: Arbitrary style transfer in real-time with adaptive instance normalization. In: Proceedings of the IEEE international conference on computer vision. pp. 1501–1510 (2017)

[10] Karras, T., Aittala, M., Laine, S., Härkönen, E., Hellsten, J., Lehtinen, J., Aila, T.: Alias-free generative adversarial networks. Advances in Neural Information Processing Systems **34**, 852–863 (2021)

[11] Karras, T., Laine, S., Aila, T.: A style-based generator architecture for generative adversarial networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 4401–4410 (2019)

[12] Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., Aila, T.: Analyzing and improving the image quality of stylegan. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 8110–8119 (2020)

[13] Kayser-Bril, N.: Google apologizes after its vision ai produced racist results (2020), https://algorithmwatch.org/en/story/google-vision-racism/

[14] Liu, Z., Luo, P., Wang, X., Tang, X.: Deep learning face attributes in the wild. In: Proceedings of the IEEE international conference on computer vision. pp. 3730–3738 (2015)

[15] Orekondy, T., Fritz, M., Schiele, B.: Connecting pixels to privacy and utility: Automatic redaction of private information in images. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 8466–8475 (2018)

[16] Orekondy, T., Oh, S.J., Zhang, Y., Schiele, B., Fritz, M.: Gradient-leaks: Understanding and controlling deanonymization in federated learning. arXiv preprint arXiv:1805.05838 (2018)

[17] Prabhu, V.U., Birhane, A.: Large image datasets: A pyrrhic win for computer vision? arXiv preprint arXiv:2006.16923 (2020)

[18] Ren, Z., Lee, Y.J., Ryoo, M.S.: Learning to anonymize faces for privacy preserving action detection. In: Proceedings of the european conference on computer vision (ECCV). pp. 620–636 (2018)

[19] Ryoo, M., Rothrock, B., Fleming, C., Yang, H.J.: Privacy-preserving human activity recognition from extreme low resolution. In: Proceedings of the AAAI conference on artificial intelligence. vol. 31 (2017)

[20] Simonite, T.: When it comes to gorillas, google photos remains blind, https://www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/

[21] Steed, R., Caliskan, A.: Image representations learned with unsupervised pre-training contain human-like biases. In: Proceedings of the 2021 ACM conference on fairness, accountability, and transparency. pp. 701–713 (2021)

[22] Sun, Q., Ma, L., Oh, S.J., Van Gool, L., Schiele, B., Fritz, M.: Natural and effective obfuscation by head inpainting. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 5050–5059 (2018)

[23] Wang, T., Zhao, J., Yatskar, M., Chang, K.W., Ordonez, V.: Balanced datasets are not enough: Estimating and mitigating gender bias in deep image representations. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 5310–5319 (2019)

[24] Wu, Z., Wang, Z., Wang, Z., Jin, H.: Towards privacy-preserving visual recognition via adversarial training: A pilot study. In: Proceedings of the European conference on computer vision (ECCV). pp. 606–624 (2018)

[25] Yang, K., Qinami, K., Fei-Fei, L., Deng, J., Russakovsky, O.: Towards fairer datasets: Filtering and balancing the distribution of the people subtree in the imagenet hierarchy. In: Proceedings of the 2020 conference on fairness, accountability, and transparency. pp. 547–558 (2020)

[26] Yang, S., Luo, P., Loy, C.C., Tang, X.: Wider face: A face detection benchmark. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 5525–5533 (2016)

[27] Zhang, L., Wu, Y., Wu, X.: A causal framework for discovering and removing direct and indirect discrimination. arXiv preprint arXiv:1611.07509 (2016)

[28] Zhong, Y., Arandjelovic, R., Zisserman, A.: Faces in places: Compound query retrieval. In: BMVC-27th British Machine Vision Conference (2016)

[29] Fang, B. , et al. "Privacy preservation in big data:a survey." Big Data Research (2016)

[30] Johnathan Orsolini.: Men/women classification: A jpg dataset for male/female classification (2019), https://www.kaggle.com/playlist/men-women-classification/

[31] Kingma, Diederik , and J. Ba .: Adam: A Method for Stochastic Optimization. In: Computer Science (2014).

[32] Ulyanov D , Vedaldi A , Lempitsky V .Instance Normalization: The Missing Ingredient for Fast Stylization[J]. 2016.DOI:10.48550/arXiv.1607.08022.

[33] Yu Z, Hu J, Min G, et al. Mobility-aware proactive edge caching for connected vehicles using federated learning[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(8): 5341-5351.

[34] Mills J, Hu J, Min G. Communication-efficient federated learning for wireless edge intelligence in IoT[J]. IEEE Internet of Things Journal, 2019, 7(7): 5986-5994.

[35] Mills J, Hu J, Min G. Multi-task federated learning for personalised deep neural networks in edge computing[J]. IEEE Transactions on Parallel and Distributed Systems, 2021, 33(3): 630-641.

[36] Chen Z, Hu J, Min G, et al. Towards accurate prediction for high-dimensional and highly-variable cloud workloads with deep learning[J]. IEEE Transactions on Parallel and Distributed Systems, 2019, 31(4): 923-934.

[37] Wang J, Hu J, Min G, et al. Dependent task offloading for edge computing based on deep reinforcement learning[J]. IEEE Transactions on Computers, 2021, 71(10): 2449-2461.

Authorized licensed use limited to: Fujian Normal University. Downloaded on June 12,2024 at 07:15:05 UTC from IEEE Xplore. Restrictions apply.