BOSTON UNIVERSITY

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Bias Mimicking: A Simple Sampling Approach For Bias Mitigation

Maan Qraitem, Kate Saenko, Bryan A. Plummer

THU-AM-364

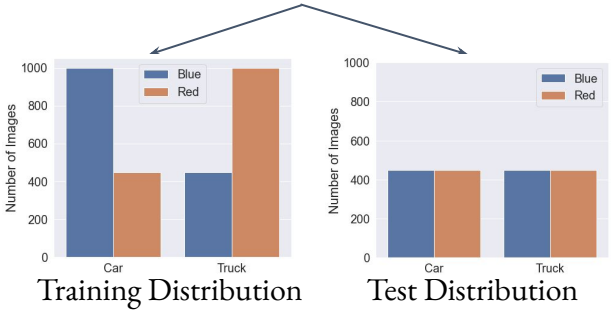
BOSTON UNIVERSITY

Summary

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Task Definition

Y: Target Labels ***B***: Bias Labels

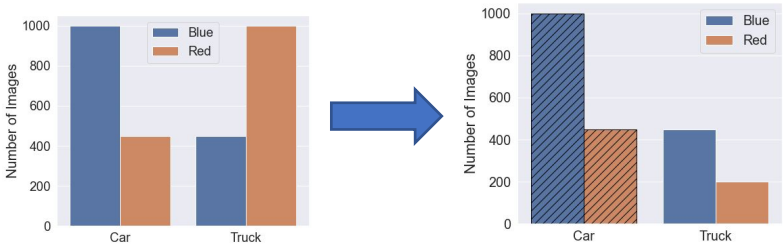


Training Distribution

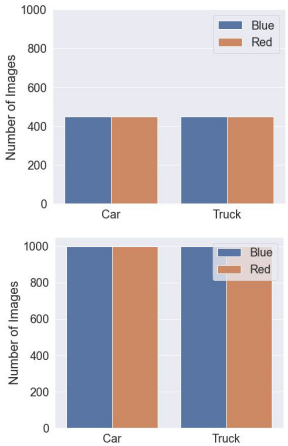
Test Distribution

- Distribution shift between train and test.
- How to prevent the model from learning the correlation between Y (Car) and B (Blue).

Bias Mimicking



- Guarantees statistical independence.
- We introduce a novel training procedure that:
  - Addresses prior sampling methods shortcomings.
  - Bridges the performance gap between sampling and non sampling methods.
  - Maintains sampling methods low cost and simplicity.



Simple Sampling Methods

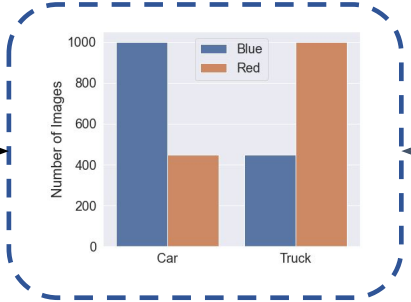
- $P_D(Y|B) = P_D(Y)$
- Simple to implement and cost effective.
- Suffer from shortcomings like overfitting over repeated samples.

BOSTON UNIVERSITY

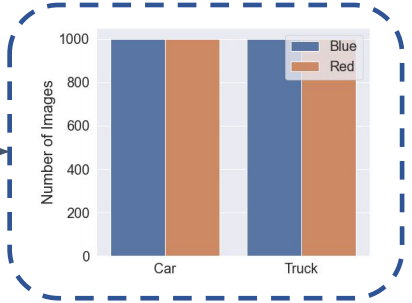
Motivation

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Issues with the data collection pipeline



Expectation



Neural Net Model

I think this a car

Observed in [1,2]

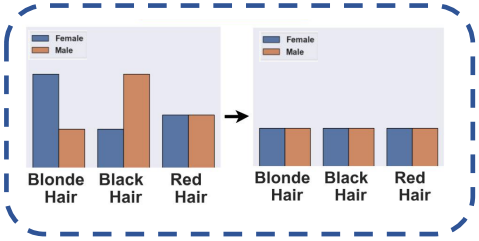
BOSTON UNIVERSITY

Sampling Methods

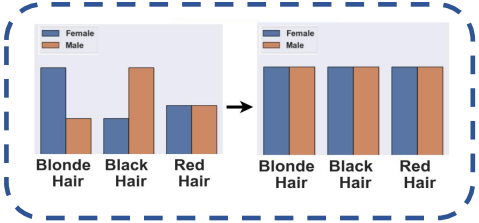
JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Sampling Methods Overview

- Simple Methods; few lines of code
- Ensure $P_D(Y|B) = P_D(Y)$
- Introduce no additional hyperparameters.
- They are missing from recent work benchmarks



Undersampling



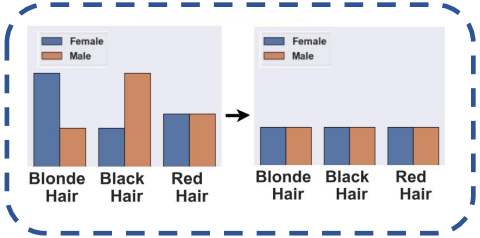
Oversampling

$$E_{x,c,s}[\frac{1}{p_D(x \in D_{c,s})} l(x, c; w)]$$

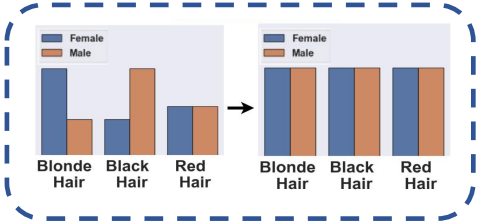
Upweighting

Sampling Methods

Sampling methods shortcomings



Limited distribution exposure per epoch which may compromise predictive performance



Model overfits over repeated samples [1]

$$E_{x,c,s}[\frac{1}{p_D(x \in D_{c,s})} l(x, c; w)]$$

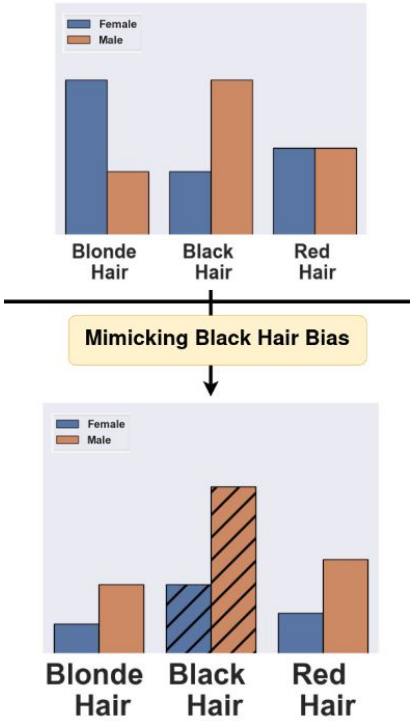
Instability issues with gradient descent [3]

Overview of the Bias Mimicking

- We want to guarantee $P_D(Y|B) = P_D(Y)$

- Bias Mimicking: Given class "c":
  - Ensure that $P_D(B|Y = c)$ is "mimicked" in each other class

- Turns out: this guarantees statistical independence!



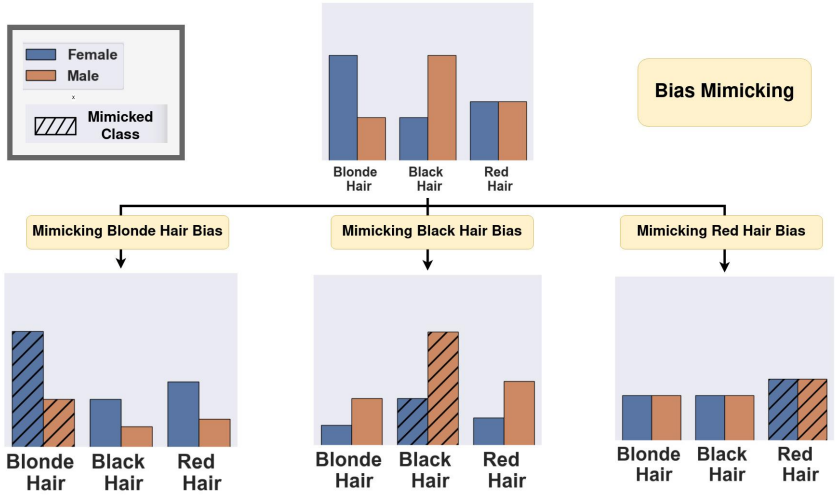
BOSTON UNIVERSITY

Method

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Training with BM

- We produce a distribution for each class.
- Each distribution preserves class c samples.
- Use all the distributions for training.
- Each distribution does not repeat samples.



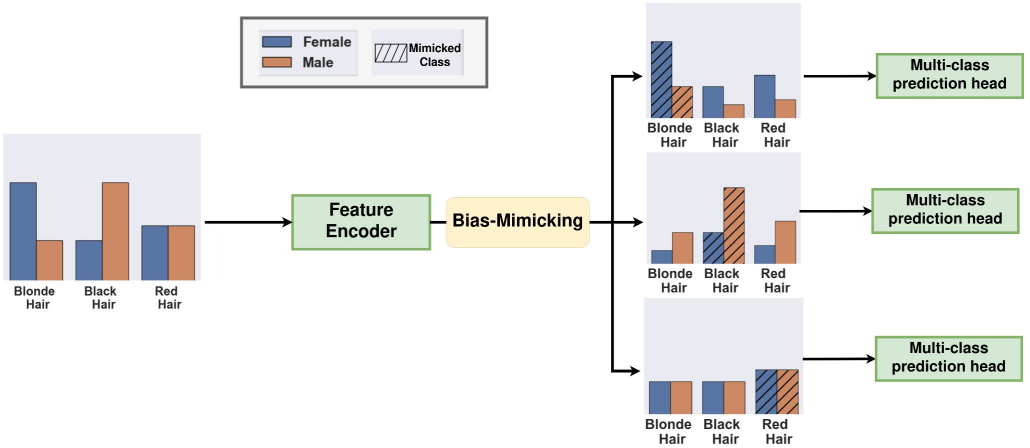
BOSTON UNIVERSITY

Method

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Training with BM

- Dedicate a multi class head for each distribution.
- **Issue:** Too many additional parameters.



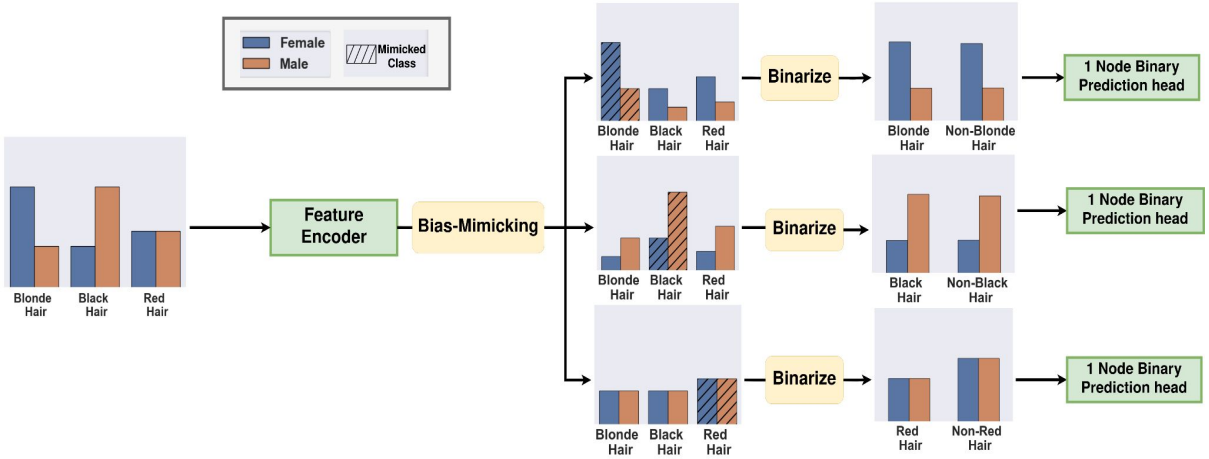
BOSTON UNIVERSITY

Method

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Training with BM

- Dedicate a binary head for each distribution.
- The binary prediction heads combined are equivalent to one multi class prediction head.



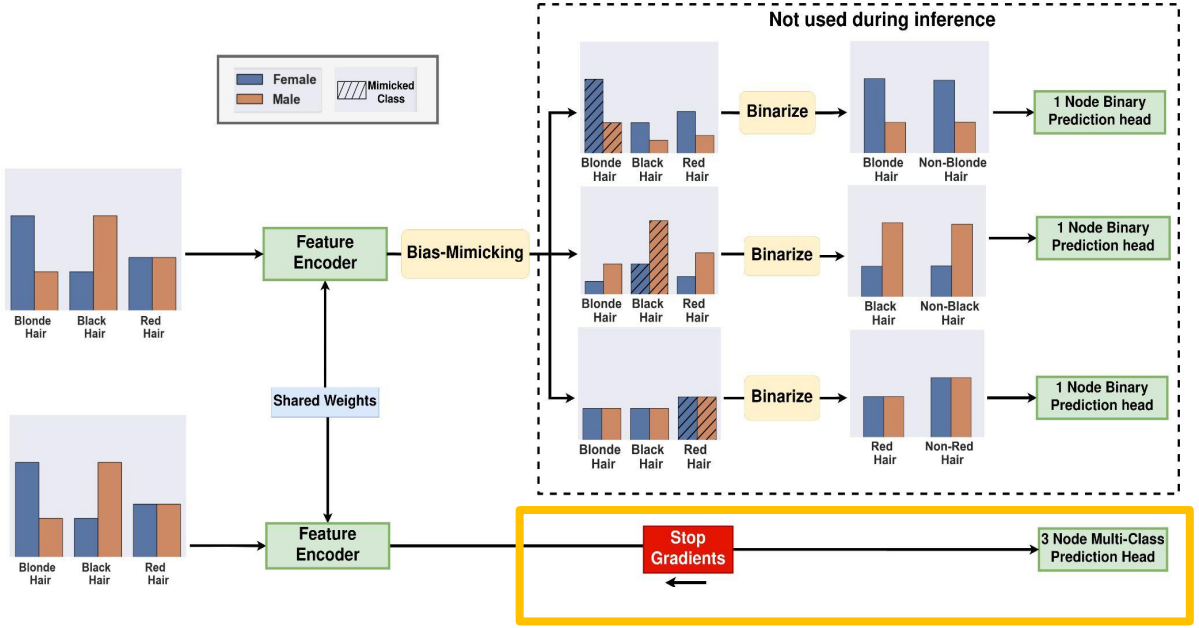
BOSTON UNIVERSITY

Method

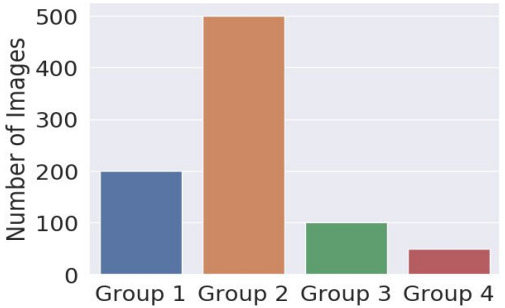
JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Inference with BM

- **Issue:** the scores may not be calibrated with respect to each other.
- **Solution:** train a multi-class prediction head on top of the debiased feature encoder.



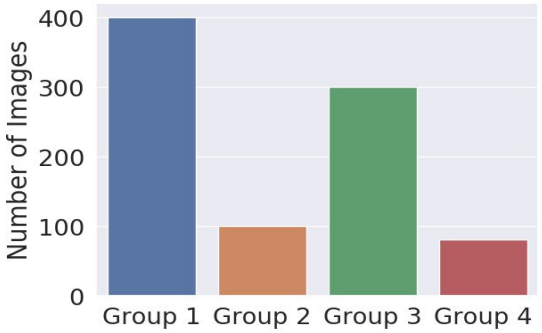
How to Bias Mimick?



Class to be preserved

Bias Mimicking

?



BOSTON UNIVERSITY

Method

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

How to Bias Mimick?

- We constrain the solution space such that the solution retains the most number of samples.
- We obtain the set of solutions using a linear program.

c: preserved class
c': mimicked class
s: bias
l: count

Set of biases

$$\max \sum_s l_s^{c'}$$

$$\text{s.t.} \quad l_s^{c'} \leq |D_{c',s}| \qquad s \in S$$

$$\frac{l_s^{c'}}{\sum_s l_s^{c'}} = P_D(B = s|Y = c) \qquad s \in S$$

BOSTON UNIVERSITY

Metrics

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

How to measure performance?

- Compute accuracy per group.
- **Unbiased Accuracy (UA)**: Take the mean of accuracies over subgroups. [2]
- **Bias Conflict (BC):** Take the mean accuracies of the under-represented subgroups only. [2]

BOSTON UNIVERSITY

Results

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

- Sampling methods show strong performance on some baselines.
- They lag behind on other benchmarks.
- Bias Mimicking shows consistent good performance unlike other sampling methods.

			Non-Sampling Methods				Sampling Methods			
		Vanilla	Adv	G-DRO	DI	BC+BB	OS	UW	US	BM
Utk-Face Age	**UA**	72.8	70.2	74.2	75.5	78.9	76.6	78.8	78.2	79.7
	BC	47.1	44.1	75.9	58.8	71.4	58.1	77.2	69.8	79.1
Utk-Face Race	**UA**	88.4	86.1	90.8	90.7	91.4	91.3	89.7	90.8	90.8
	BC	80.8	77.1	90.2	90.9	90.6	90.0	89.2	89.3	90.7
CelebA Blonde	**UA**	82.4	82.4	90.4	90.9	90.4	88.1	91.6	91.1	90.8
	BC	66.3	66.3	89.4	86.1	86.5	80.1	88.3	88.5	87.1
CIFAR-S	**UA**	88.7	81.8	89.1	92.1	90.9	87.8	86.5	88.2	91.6
	BC	82.8	72.0	88.0	91.9	89.5	82.5	80.0	83.7	91.1
Average	**UA**	83.0	80.1	86.1	87.3	87.9	85.9	86.6	87.0	**88.2**
	BC	69.2	64.8	85.8	81.9	84.5	77.6	83.6	82.8	**87.0**

BOSTON UNIVERSITY

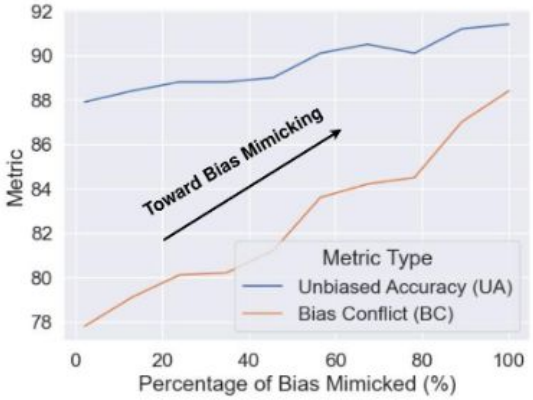
Sensitivity Analysis

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

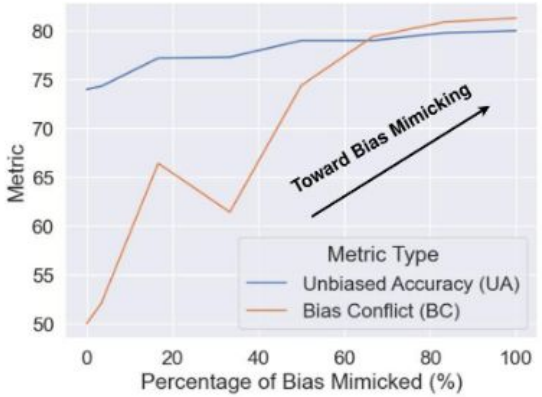
How sensitive is the model to the mimicking condition?

We vary the amount of bias mimicked between a percentage where:

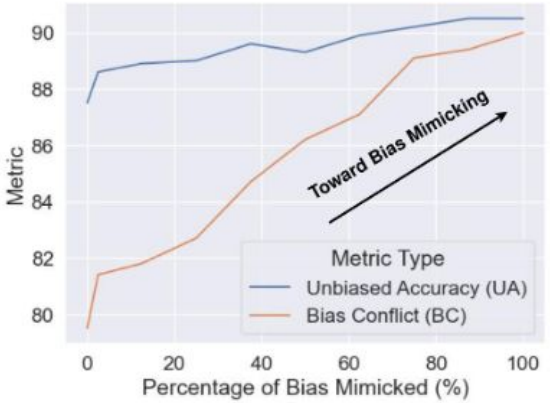
- 0%: distribution remains the same.
- 100%: Complete bias mimicking.



a) CelebA/Blonde



b) Utk-Face/Age



a) Utk-Face/Race

BOSTON UNIVERSITY

Conclusion

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

- **We showed** that simple sampling methods can be competitive on some benchmarks when compared to non sampling state-of-the-art approaches.
- **We introduced** a novel resampling method: Bias Mimicking that bridges the performance gap between sampling and nonsampling methods.
- **We conducted** an extensive empirical analysis of Bias Mimicking that details the method's sensitivity to the Mimicking condition. Refer to the paper for more details.

BOSTON UNIVERSITY

References

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

[1] Wang, Z., Qinami, K., Karakozis, I. C., Genova, K., Nair, P., Hata, K., & Russakovsky, O. (2020). Towards fairness in visual recognition: Effective strategies for bias mitigation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 8919-8928).

[2] Sagawa, S., Koh, P. W., Hashimoto, T. B., & Liang, P. (2019). Distributionally robust neural networks for group shifts: On the importance of regularization for worst-case generalization. arXiv preprint arXiv:1911.08731.

[3] An, J., Ying, L., & Zhu, Y. (2020). Why resampling outperforms reweighting for correcting sampling bias with stochastic gradients. arXiv preprint arXiv:2009.13447.

BOSTON UNIVERSITY

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Link to code on github