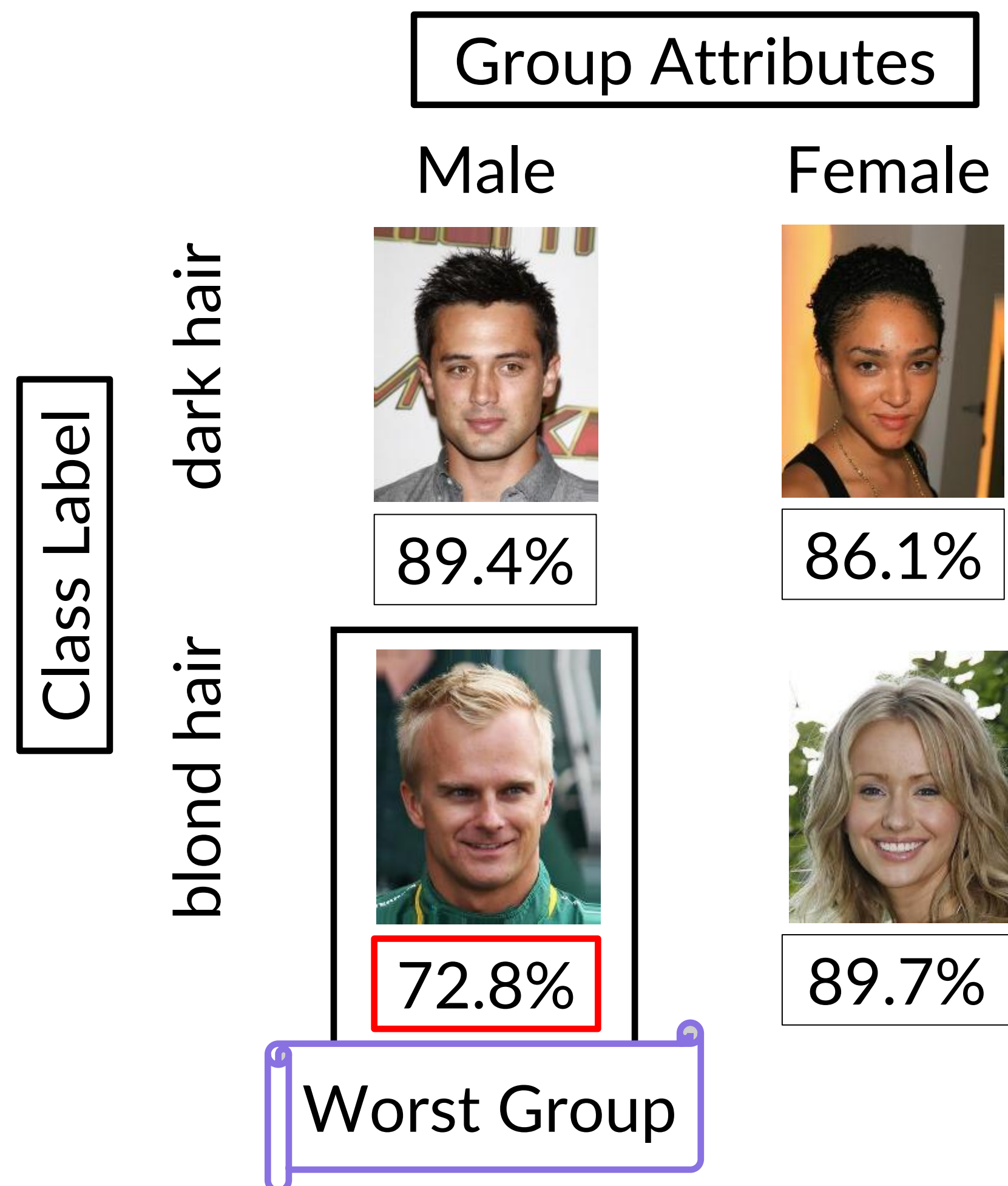
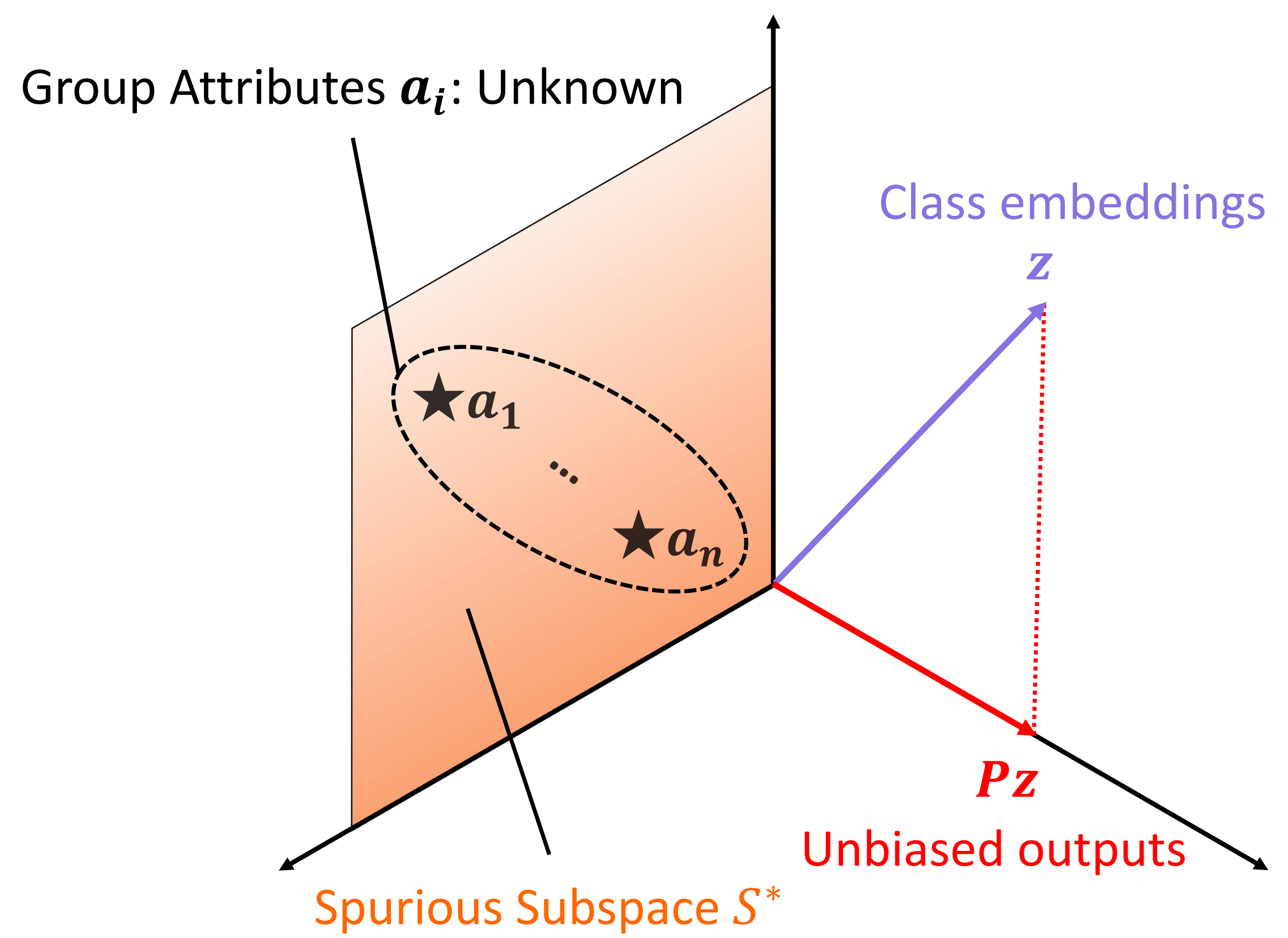


CLIP is not robust to group attributes



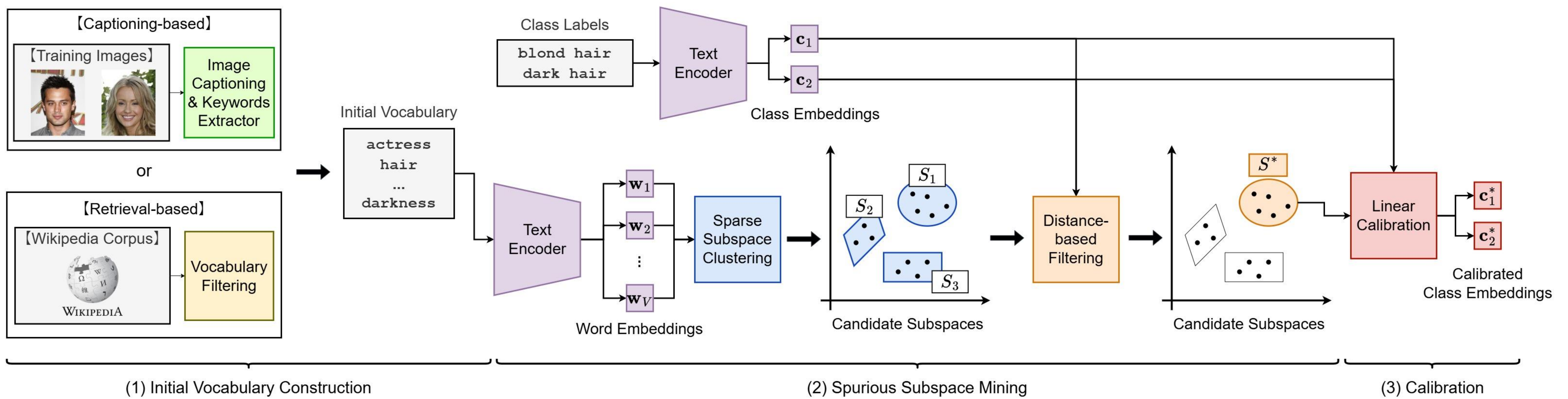
Our goal is to enhance the worst group accuracy (WG) without any knowledge of group attributes.

Debiasing by linear projection



Even if group attributes a_i are unknown, we can cancel out the negative impact of a_i by estimating the subspace S^* .

Spurious Subspace Mining (SSM)



- (1) Initial Vocabulary Construction: obtaining an initial vocabulary that is likely to contain group attributes
- (2) Spurious Subspace Mining: discovering multiple subspaces by sparse subspace clustering and specifying the spurious subspace
- (3) Calibration: calibrating class embeddings by linearly projecting them onto the orthogonal complement of the spurious subspace

Desirable properties of our method

Table1: Properties of Group Robust Classification Methods

Method	Knowledge-free	Linear	Calibration-based
ERM Linear Probe	✓	✓	-
ERM Adapter	✓	-	-
WiSE-FT	✓	✓	-
DFR	✓	✓	-
Orth-Cali	-	✓	✓
Contrastive Adapter	-	-	-
FairerCLIP	-	-	✓
ROBOSHOT	✓	✓	✓
Ours	✓	✓	✓

- Knowledge-Free: without any knowledge of group attributes
- Linear: training only a single linear projection
- Calibration-based: no updating model's parameters

Experimental Results

Table2: Comparison with Knowledge-free Methods

Method	CelebA			Waterbirds		
	WG↑	Avg↑	Gap↓	WG↑	Avg↑	Gap↓
Zero-shot CLIP	72.8	87.7	14.9	44.2	90.4	46.2
ERM Linear Probe	28.3	94.7	66.4	65.9	97.6	31.7
ERM Adapter	42.8	93.6	50.8	<u>77.6</u>	97.8	<u>20.2</u>
WiSE-FT	80.0	87.4	7.4	65.9	97.6	31.7
DFR (Subsample)	76.3	92.1	15.8	51.9	95.7	43.8
DFR (UPsample)	<u>83.7</u>	91.2	7.5	65.9	96.1	30.2
B2T	73.3	88.0	14.7	61.2	84.9	23.7
ROBOSHOT	82.6	85.5	2.9	45.2	79.9	34.7
Ours (Captioning-based)	82.2	84.2	<u>2.0</u>	79.4	88.5	9.1
Ours (Retrieval-based)	85.1	85.6	0.5	69.0	91.2	22.2

Our method outperforms all the knowledge-free baselines in WG and Gap accuracy on both datasets.