Contemplating Visual Emotions: Understanding and Overcoming Dataset Bias (Supplementary Material)

Rameswar Panda[1], Jianming Zhang[2], Haoxiang Li[3], Joon-Young Lee[2], Xin Lu[2], and Amit K. Roy-Chowdhury[1]

[1] Department of ECE, UC Riverside. {rpand002@,amitrc@ece.}@ucr.edu
[2] Adobe Research. {jianmzha,jolee,xinl}@adobe.com
[3] Aibee. hxli@aibee.com

Table 1. TABLE OF CONTENTS

[Summary.] In this Supplementary Material, we give out details omitted in the main text: (a) detailed information on our 3-level emotion hierarchy that we used for both creation of the **WEBEmo** dataset and curriculum guided learning, (b) information on different training and test datasets including motivation and the labeling protocol, (c) additional experiments including duplicate analysis, correlation analysis on all emotion categories, implementation details of the compared methods, visual persuasion experiment and sample qualitative prediction results among the top-5K predictions by our curriculum guided learning model. Please see below for more detailed analysis on each of the section.

Detailed Information on the Emotion Hierarchy

Table 2. Three Level Emotion Hierarchy Adopted in Current Work

Level - 1 Emotions	**Level - 2 Emotions**	**Level - 3 Emotions**
negative	anger	disgust envy exasperation Irritability rage
negative	fear	confusion horror nervousness
negative	sadness	disappointment neglect sadness shame suffering sympathy
positive	joy	cheerfulness contentment enthrallment optimism pride relief zest
positive	love	affection gratitude lust
positive	surprise	surprise

Discussion. As discussed in Section 4.1 of the main paper, we follow the Parrott's emotion wheel [5] and construct a three-level emotion hierarchy to retrieve stock images for constructing our large scale **WEBEmo** dataset. Tab. 2 show the three-level emotion hierarchy that we use in our current work. This emotional grouping is more interpretable and helps in learning a better recognition model by leveraging the hierarchal structure of emotions.

Detailed Information on our new Emotion Datasets

(a) Descriptive Statistics of WEBEmo Dataset

Table 3. Proposed Fine-grained Emotion Categories with associated Search Keywords in our **WEBEmo** dataset. The comprehensive list of search keywords help in reducing the positive set bias by diversifying the image search results.

Level - 3 Emotions	Search Keywords
disgust	contempt, disgust, disgusting, loathing, loath, loathsome
envy	envy, jealous, jealousy
exasperation	exasperation, frustrated, frustration, frustrating
irritability	aggravated, aggravation, agitated, annoy, annoyance, annoyed, grouchy, irritate, irritated, irritation
rage	anger, angry, bitterness, dislike, ferocious, ferocity, fury, hate, hatred, rage, resent, resentful, resentment, scorn, spite, wrath
confusion	confused, confusion, doubt, doubtful, hesitant, hesitation, perplexed, unsure
horror	fear, fearful, fright, horror, hysteria, panic, shock, shocked, terrified, terror
nervousness	anxiety, anxious, apprehension, apprehensive, distressful, dread, dreadful, nervous, nervousness, uneasiness,uneasy, worried, worry
cheerfulness	amused, amusing, cheer, cheerful, cheerfulness, delight, delightful, elation, enjoy, enjoyment, euphoria, fun, funny, glad, gladness, glee, happiness, happy, harmony, joy, satisfied
contentment	contentment, pleased, pleasure
enthrallment	rapture
optimism	confidence, confident, eager, hope, hopeful, optimism
pride	proud, success, successful, triumph
relief	calm, peaceful, relax, relaxed, relaxing, relief
zest	enthusiasm, excited, excitement, exhilarated, exhilarating, exhilaration, thrill, thrilling, zeal
affection	adoration, affection, care, compassion, fond, fondness, like, liking, love, loving, sentimental, tender, tenderness, worship
gratitude	appreciate, appreciation, grateful, gratitude, thank
lust	desire, lust, passion, sexual desire
disappointment	bored, boredom, disappointed, disappointment, disappointing
neglect	defeat, dejection, embarrassed, embarrassment, homesickness, humiliated, humiliation, insecure, insecurity, insult, insulted, insulting, loneliness, lonely, neglect, rejection
sadness	depressed, depression, despair, gloom, gloomy, glum, grief, hopeless, hopelessness, melancholy, miserable, misery, sad, sadness, sorrow, unhappy, woe
shame	guilt, guilty, regret, regretful, remorse, shame
suffering	agony, anguish, hurt, pain, suffer, suffering, torment, torture, trauma
sympathy	pathetic, pitiful, pity, sympathy
surprise	amazed, amazement, astonished, astonishment, surprise, surprised, surprising

(b) Descriptive Statistics of Emotion-6 Dataset

Motivation and Image Collection. Our main motivation on creating Emotion-6 dataset is to repeat the standard data collection/annotation protocol used by existing works and see how well it performs regarding the dataset biases. Specifically, we follow [7, 4] and create one emotion dataset of natural images by imitating the general notion of creating fully supervised datasets. Towards this, we first consider the most popular six pan-cultural basic emotion categories (*anger*, *fear*, *joy*, *love*, *sadness* and *surprise*) [5, 1] and collect about 150K images from both Google and Flickr using the secondary level emotion categories [5] as keywords, e.g., "disgust", "envy", "exasperation", "irritability" and "rage" are used as search keywords to collect images of "anger" category. Next, we ask five human subjects to label the true positive images within each emotion category and remove the rest non-emotional images from the collection, as in [7, 4]. In total, we acquire a total of 8350 images (*anger*: 1604, *fear*: 1280, *joy*: 1964, *love*: 724, *sadness*: 2221, *surprise*: 557), which is of the same order magnitude as the very recent emotion dataset reported in [4].

Fig. 1. Category-wise image distribution in Emotion-6 dataset.

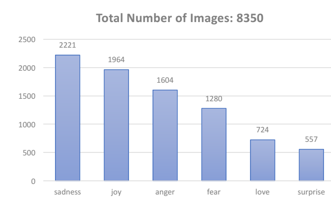


Table 4. Emotion Categories along with Search Keywords in **Emotion-6** Dataset.

Emotion Categories	Search Keywords
anger	contempt, disgust, hatefulness, loathing, nausea, revulsion, covetousness, envy, jealousy, malice, resentment, rivalry, annoyance, displeasure, exasperation, frustration, irritation, provocation, aggravation, agitation, grumpy, impatience, irritability, anger, bitter, dislike, fury, hostility, outrage, rage, spite
fear	agitation, anarchy, bewilderment, confusion, distraction, uncertainty, frightful, horror, hysteria, panic, scared, shock, terror, anxiety, apprehension, distress, nervousness, suspense, uneasiness, worry
joy	amusement, cheerful, delight, enjoyment, happiness, contentment, pleasure, satisfaction, attention, captivation, enraptured, enthrallment, fascination, eagerness, hope, optimism, positiveness, jubilance, pride, proud, gratefulness, relief, thankfulness, enthusiasm, excitement, exhilaration, thrill, zeal, zest
love	adoration, affection, attractiveness, compassion, appreciation, gratitude, obligation, desire, infatuation, lust, passion
sadness	disappointment, dismay, displeasure, dissatisfaction, distress, frustration, dejection, insecurity, isolation, loneliness, neglect, depression, despair, grief, misery, sadness, sorrow, unhappiness, embarrassment, guilt, humiliation, regret, shame, agony, anguish, hurt, suffering, compassion, consolation, empathy, pity, sympathy
surprise	amazement, astonishment, awe, curiosity, surprise

(See Fig. 2 for category-wise sample images from **Emotion-6** dataset.)

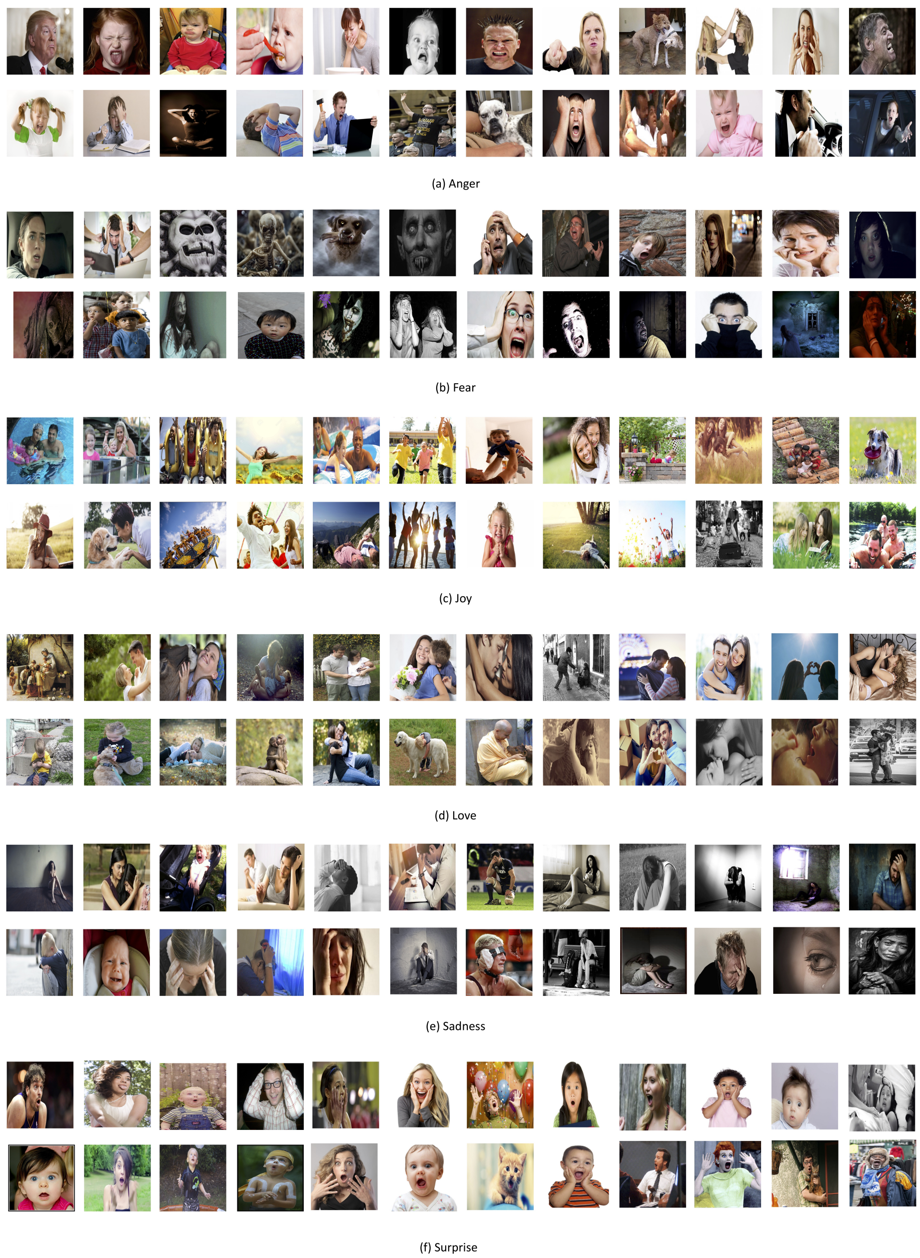


Fig. 2. Sample Images from **Emotion-6** Dataset. All the images are labeled by five humans following similar annotation protocol used in prior works. Despite best efforts, the dataset suffers from both positive and negative set biases. Best viewed in color.

(c) Descriptive Statistics of UnBiasedEmo Dataset

Motivation and Image Collection. Our main motivation on creating this dataset is to compare the generalization ability of various learning strategies in a less biased manner by testing on a cross-domain unbiased test set containing images of same object/scene with different emotions. Unlike existing emotion datasets, this dataset is acquired with a special focus on selecting diversified images with similar appearances/objects but with different emotions, such as the ones shown in Fig. 4. We first consider the same six basic emotion categories as Emotion-6 dataset and assembled a list of action keywords (shown in Tab. 5) related to each category, e.g., rage, fighting, irritating, annoyed, yelling, etc., for *anger* emotion. We select about 35 keywords covering a wide variety of concepts including human, animals and natural scenes. Given all meaningful combinations of action and concepts (722 in total), we query Google and retrieve images (usually between 200 to 300) returned for each query. We remove all unreadable images and obtain an initial pool of about 60,000 noisy images. Next, we follow a two-stage approach for cleaning the initial pool of images. First, we remove all non-emotional images from each action-concept pair, as in [4]. Second, we pool all the remaining images related to a concept and remove duplicates. Finally, we ask two human subjects to manually verify the images and keep those which are consistently labeled by both subjects. In total, we obtain 3045 images across six emotion categories. To the best of our knowledge, this is the first attempt to create such unbiased test set in visual emotion analysis.

Table 5. Emotion Categories with associated Action Keywords.

anger	**fear**	**joy**	**love**	**sadness**	**surprise**
angry	fear	happy	loving	sad	surprised
fighting	horrified	delighted	lovely	heartbroken	astonished
rage	scared	smiling	caring	sorrowful	amazed
punching	afraid	pleasing	passionate	pessimistic	shock
yelling	terrified	successful	tender	hurt	shocking
irritated	terrifying	peaceful	sentimental	depressed	awe
irritating	hiding		passion	depressing	
annoyed	panic		affection	weep	
mad	frightening		worship like	displeasure	
cursing	frightened		romantic	disap- pointed	
				unhappy	
				sorrow	
				homesick	
				crying	
				lonely	

Table 6. List of Concepts

conepts
dog, baby, guy, girl, teenager, people, couple, family, crowd, tiger, scenery, cat, amusement park, bird, house, event, soldier, teacher, horse, lion, team, gift, athletes, boxer, dancer, car, room, bridge, tower, flower, leaf, tree, train

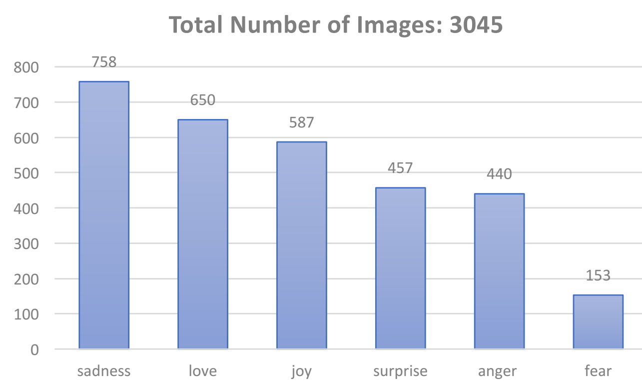


Fig. 3. Category-wise image distribution in **UnBiasedEmo** test dataset.

(See Fig. 4 for category-wise sample images from **UnBiasedEmo** test set.)

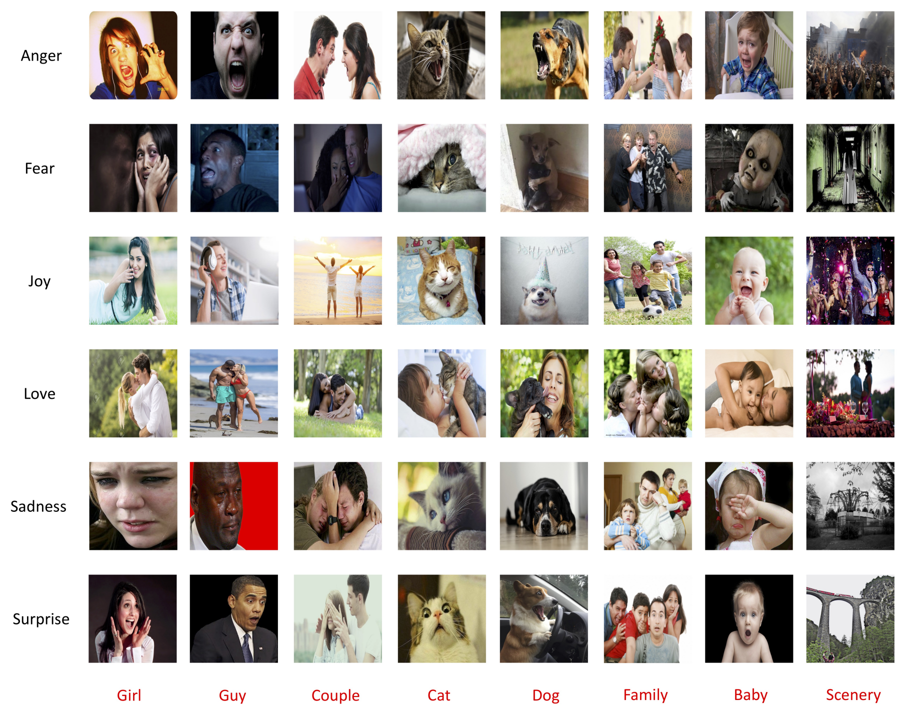


Fig. 4. Sample images from the **UnBiasedEmo** test dataset. The dataset contains images of same concept with different emotions. Figure shows six different emotions for eight concepts mentioned in red, e.g., girl, guy, couple etc. Specifically, images in each column represent one concept with different emotions. All the images in this dataset are manually labeled using two human subjects using a controlled annotation experiment. To the best of our knowledge, this is the first attempt to create such unbiased emotion dataset in visual emotion analysis. Best viewed in color.

Table 7. Proposed Action-Concept Pairs. Each action-concept pair is used to retrieve images (usually between 200 to 300) from Google. We obtain an initial pool of 60,000 images covering a wide range of fine-grained emotion concepts.

conepts	**action-concept pairs**
anger	angry dog, angry baby, angry guy, angry cat, angry couple, angry family, angry crowd, angry tiger, mad dog, mad baby, mad guy, mad teenager, mad girl, mad couple, mad teacher, annoyed baby, annoyed couple, annoyed girl, annoyed cat, annoyed teacher, irritating baby, irritating girl, irritated cat, irritated girl, irritated guy, irritated couple, irritated dog, irritated lion, yelling dog, yelling baby, yelling guy, yelling girl, yelling teacher, yelling teenager, yelling couple, yelling family, yelling crowd, yelling cat, yelling horse, yelling lion, rage baby, rage teenager, rage crowd, rage tiger, rage lion, fighting baby, fighting guy, fighting couple, fighting events, fighting family, fighting horse, fighting lion, fighting tiger, punching baby, punching guy, punching face, punching girl, teacher cursing, boxer punching, angry athletes, angry soldier, yelling soldier
fear	fear baby, fear scenery, scared cat, scared family, scared girl, scary fish, scary clown, scary house, scared dog, scared baby, scared guy, scared teenager, scared couple, scared teacher, scared house, scary bridge, scary tower, scary tree, scary room, scary train, scary scenery, scary event, scary gift, horrified guy, panic face, frightening face, frightening baby, frightening girl, frightening couple, frightening scenery, frightening events, frightening room, frightening tree, frightening bridges, frightening gif, terrified horse, terrified baby, terrified girl, terrified guy, terrified couple, terrifying baby, terrifying guy, terrifying couple, terrifying events, terrifying room, terrifying tree, terrifying bridge, terrifying tower, terrifying house, hiding girl, panic scenery, guy afraid, girl afraid, couple afraid
joy	happy dog, happy baby, happy guy, happy scenery, happy cat, happy family, happy beach, happy sunset, happy crowd, amusement park, happy bird, lovely mountain, lovely river, happy tiger, delighted dog, delighted baby, delighted guy, delighted girl, delighted cat, delighted family, delighted teenager, delighted tree, delighted face, smiling dog, smiling cat, smiling baby, smiling guy, smiling girl, smiling family, smiling teenager, smiling team, smiling event, smiling scenery, smiling teacher, smiling soldier, smiling dancer, pleasing scenery, pleasing tree, pleasing bird, pleasing cat, peaceful scenery, peaceful tree, peaceful bridge, peaceful house, peaceful flower, peaceful family, peaceful baby, peaceful girl, successful baby, successful dog, successful guy, successful guy, successful teenager, successful scenery, successful family, successful team, successful event, successful teacher, successful boxer, successful dancer, successful people

Table 8. Proposed Action-Concept Pairs. Each action-concept pair is used to retrieve images (usually between 200 to 300) from Google. We obtain an initial pool of 60,000 images covering a wide range of fine-grained emotion concepts.

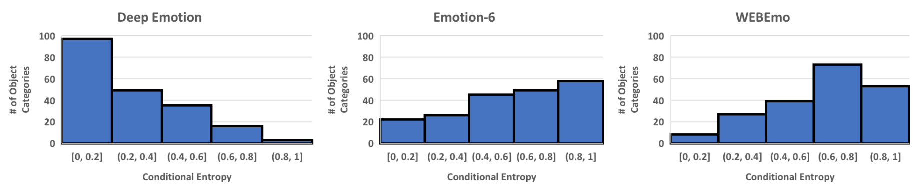
conepts	**action-concept pairs**
love	loving dog, loving baby, loving scenery, lovely couple, bird in love, loving guy, loving girl, loving teenager, loving gift, loving flower, romantic couple, romantic family, romantic scenes, romantic cat, romantic events, romantic night, romantic lover, passionate couple, caring baby, caring guy, caring girlfriend, caring couple, caring dog, caring cat, caring teacher, caring soldier, tender scenes, tender couple, sentimental couple, baby affection, dog affection, cat affection, tiger affection, lion affection, girl affection, teenage affection, couple affection, people affection, family affection, athletes affection, affectionate people, affectionate scenery, affectionate events, affectionate room, affectionate tree, affectionate bridge, worship couple, like couple, affectionate birds, affectionate boyfriend, affectionate mother, affectionate father
sadness	sad dog, sad baby, sad guy, sad scenery, sad cat, sad couple, sad family, sad beach, sad sunset, sad crowd, sad amusement park, sad bird, sad mountain, sad river, sad girl, sad teenager, sad teacher, sad people, sad house, sad soldier, sad team, sad athletes, sad dancer, sad room, sad leaf, sad tree, heartbroken girl, brokenhearted guy, heartbroken couple, heartbroken dog, heartbroken teenager, heartbroken athletes, sorrowful guy, sorrowful girl, sorrowful teenager, sorrowful people, sorrowful dog, pessimistic girl, pessimistic people, pessimistic gif, baby hurt, soldier hurt, girl hurt, couple hurt, depressing scenery, depressed baby, depressed guy, depressed girl, depressed couple, depressed people, depressed soldier, depressed teenager, depressed teacher, depressed room, depressed house, displeasure guy, disappointed baby, disappointed people, disappointed girl, disappointed teenager, disappointed couple, disappointed crowd, unhappy baby, unhappy girl, unhappy guy, unhappy couple, unhappy teacher, unhappy team, unhappy teenager, unhappy soldier, homesick gir, lonely guy, lonely girl, lonely people, baby crying, girl crying, guy crying, couple crying, people crying, teenager crying, teacher crying, soldier crying, dancer crying, athletes crying, crying scenes, crying dog, crying cat, baby weep, girl weep, teenager weep, couple weeping, soldier weeping, athletes weeping, dog weeping
surprise	surprised dog, surprise baby, surprise guy, surprised cat, surprise couple, surprised family, surprised crowd, surprised girl, nice surprise, surprise gift, birthday surprise, astonished cat, astonished baby, astonished girl, astonished guy, astonished teenager, astonished couple, astonished people, amazed baby, amazed guy, amazed girl, amazed couple, amazed people, amazed family, baby shocked, girl shocked, guy shocked, shocked couple, shocked family, awe girl, awe baby, awe face, awe tree, awe bridge, amazing scenery

Additional Experimental Results

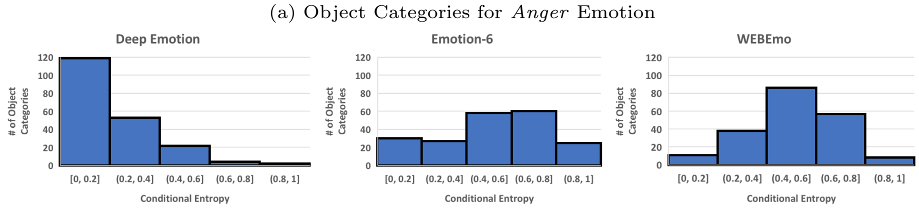
(a) Duplicates between our WEBEmo and Public Benchmarks.

We use a perceptual hashing method [**?**] to identify the duplicate images and find that only about 0.1% images in our **WEBEmo** dataset have duplicates in the Deep Emotion dataset [7]. There exists no duplicate images between our **WEBEmo** dataset and the Deep Sentiment dataset [6].

(b) Correlation Analysis with Object/Scene Categories.



(a) Object Categories for *Anger* Emotion

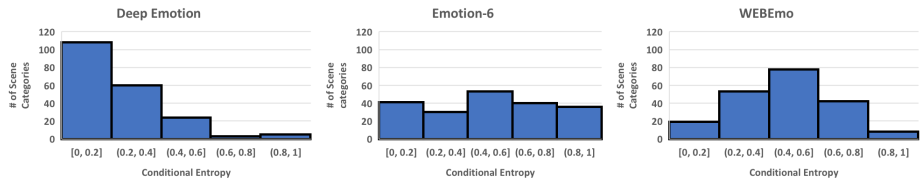


(b) Object Categories for *Fear* Emotion.

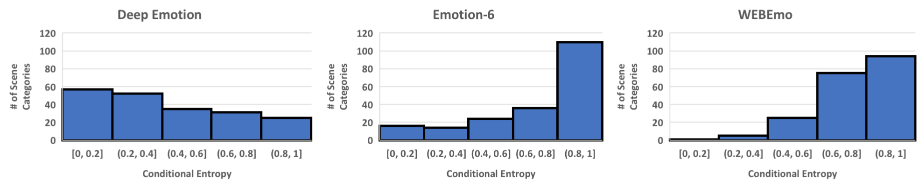
Fig. 5. Distribution of object categories w.r.t conditional entropy. (a) objects in *anger* emotion, (b) objects in *fear* emotion. While analyzing correlations between objects and emotions, we observe that more number of object categories (zero conditional entropy) in both Deep Emotion and Emotion-6 datasets are only present in either positive or negative set of an emotion category leading to the presence of bias. On the other hand, only few object categories in our **WEBEmo** dataset are within the low entropy range leading to a much less biased dataset (see right most plots).

[**Analyzing Object/Scene Categories.**] Apart from the distributions, we further examine the zero entropy categories and find most of them will lead to a dataset bias. For example, objects like balloon, candy store and parachute are only present in negative set of *sadness* in both Deep Emotion and Emotion-6 datasets. Categories like balloon are strongly related to happiness, but still there should be a few negative balloon images such as sad balloon in the negative set[4]. Completely missing the negative balloon images will lead to dataset bias. On the other hand, for the same *sadness* emotion, we have only one object category (i.e.,

[4] For example, see: `https://tinyurl.com/yazvkjmv`



(a) Scene Categories for *Fear* Emotion



(b) Scene Categories for *Sadness* Emotion.

Fig. 6. Distribution of scene categories w.r.t conditional entropy. (a) scenes in *fear* emotion, (b) scenes in *sadness* emotion. While analyzing correlations between scenes and emotions, we see that both datasets (Deep Emotion and Emotion-6) are biased towards to specific scene categories. In contrast, only few scene categories in our **WEBEmo** dataset are within the low entropy range showing that most of the scenes are well distributed across positive and negative emotion sets in our dataset., e.g., only one scene category has zero entropy for the *sadness* emotion in our dataset (see right most plots).

fly) with zero entropy in our **WEBEmo** dataset. Our **WEBEmo** dataset contains some negative balloon images leading to non-zero entropy of balloon category and overall a less biased dataset. We also perform a toy experiment related to this (see Figure 7 for an example) and observe that models trained on both Deep Emotion and Emotion-6 datasets fail to recognize the basic emotion category of negative from these images while model trained using our **WEBEmo** dataset correctly recognizes the sadness emotion from the below two images. This once again shows that emotions are correlated with object/scene categories and carefully analyzing the correlations can help to identify the presence of bias in emotion datasets and also help to understand the model predictions.

Fig. 7. Two sample images of sad balloon with negative emotion. Models trained on both Deep Emotion and Emotion-6 fail to predict the basic emotion while model trained using our dataset correctly predicts the sadness emotion from these two images. Both images are taken from Google Images with the search keyword sad balloon park. Best viewed in color.

(c) Details on Exploration Study

As described in Section 5.1 (Exploration Study) of the main paper, we analyze cross-dataset generalization performance by comparing with three different methods: (1) Direct Learning, (2) Self-Directed Learning and (3) Joint Learning. Here, we first present the common experimental settings that are applicable to all the methods and then describe the implementation details of each methods.
[**Common Settings.**] We follow the following experimental settings in all our compared methods.

- We choose ResNet-50 as our default choice of network and initialize the networks from the same ImageNet checkpoint while comparing with all methods.
- We use the same training and testing split across all three methods to report our results.
- During training, all input images are resized to 256×256 pixels and then randomly cropped to 224×224.
- We use batch normalization and train the networks using stochastic gradient descent with same parameters for a comparison with all three baseline learning strategies.

[**Direct Learning.**] Direct learning refers to the learning strategy that directly learn a deep network using the noisy web images of 25 fine-grained emotion categories. Apart from 25 categories, we implement two additional direct learning strategies with two and six emotion categories respectively. Our extensive results in the main paper show that the performance of direct learning baseline is much worse compared to our proposed curriculum guided learning across different tasks. This is not surprising since emotions are highly complex and ambiguous that directly learning models to categorize such fine-grained details fails to learn discriminative features.
[**Self-Directed Learning.**] Following [6, 2], we implement this baseline by starting the training with a small clean set and then progressively adapt the model by refining the noisy web data. More specifically, we first manually label 500 images as the initial clean set of images and then start the training with such small set. After training the network using such clean labeled image set, we test the remaining images and include the images in the training set whose probability of being in an emotion category is more than threshold. We then update the network using the new set of images to obtain an improved model for recognizing emotions. We iterate this procedure until all the images are progressively included in the training set. Our experiments show that self-directed learning baseline shows better generalization compared to the direct learning. However, it still suffers from the requirement of initial labeled data which can be difficult to obtain in many real-world settings.
[**Joint Learning.**] We implement this baseline in a multi-task setting i.e., simultaneously learning to classify images with two, six and twenty-five emotion categories. Our experimental analysis show that the joint learning baseline is more competitive than the other two learning strategies since it learns a shared representation from multiple tasks. However, the proposed curriculum guided learning still outperforms it in terms generalization across other datasets since

Table 9. Results on persuasion dataset. Features learned using our curriculum guided webly supervised learning significantly outperforms the ImageNet baseline.

Methods	Accuracy (%)
ImageNet	70.33
Direct Learning	73.66
Self-Directed Learning	74.80
Joint Learning	76.33
Curriculum Learning	**78.33**

it is able to learn more discriminative fine-grained emotion feature by ordering training from easy to difficult in a sequential manner.
[Impact of Emotion Categories.] As discussed in Section 5.1 (Impact of Emotion Categories), we compare with two different curriculum guided training strategy by varying the number of emotion categories. Here, we additionally compare with a single stage learning strategy using only 2 emotion categories and observe that it produces inferior results, with an accuracy of 75.21% on the self test set and a mean accuracy of 68.04% on other two datasets, compared to 81.41% and 72.76% respectively by the three stage curriculum guided learning. Note that this performance is lower than the performance of the two-stage curriculum learning strategy (2-6), which once again corroborated the fact that the generalization ability of learned models increase with increased number of fine-grained emotion categories. We also observe similar drop in performance of direct learning baseline while training with two emotion categories compared to the training with 6 and 25 emotion categories.

(d) Visual Persuasion

[Goal.] The goal of this experiment is to analyze the effectiveness of our learned emotion feature in predicting communicative intents from persuasive images [3]. In particular, our aim here is to recognize the overall favorability of famous politicians towards the target in terms of either positive or negative from their persuasive portraits. Predicting global favorability towards a target is more challenging and different from sentiment analysis since images with negative sentiment can also show positive favorability towards a target and vice versa.
[Datasets and Settings.] We use visual persuasion dataset consisting of 1124 images of 8 U.S. politicians [3] and annotate each image as either positive or negative by looking into the favorability towards a target person as an integrated measure. We use the learned models as pure feature extractors and adopt standard training/testing split [3] to report binary label accuracies.
[Results.] We have the following key findings from Table 9: (1) Features learned using our curriculum guided learning outperforms all other baselines by a significant margin. We believe this is because of the two introduced components working in concert: first, our hierarchical data collection procedure to cover diverse concepts and second, our curriculum guided learning strategy to learn more discriminative emotion feature. (2) Although being trained on generic web images, these results indicate the potential of our learned features in predicting communicative intents from persuasive images (e.g., politician photos). (3) ImageNet features show inferior results compared to features learned using our **WEBEmo** dataset (70.33% vs 78.33%).

(e) Impact of Dataset Size

We randomly sample a subset of 25,000 images (size similar to Deep Emotion dataset) from our **WEBEmo** dataset and follow the curriculum guided learning to train a model. We observe that the model trained using this subset produces an accuracy of 69.04% on the self test set and a mean accuracy of 64.49% on the other datasets, compared to 81.41% and 72.76% respectively by the full dataset which is about 10 times larger than this subset. Both self test and mean others accuracy increases as the size of dataset increases. Interestingly, model trained using the manually labeled Deep Emotion dataset only achieves a mean accuracy of 63.53% compared to 64.49% by the reduced subset while testing on the other datasets. This once again shows the effectiveness of our approach in learning a generalization recognition model.

(f) Sample Results

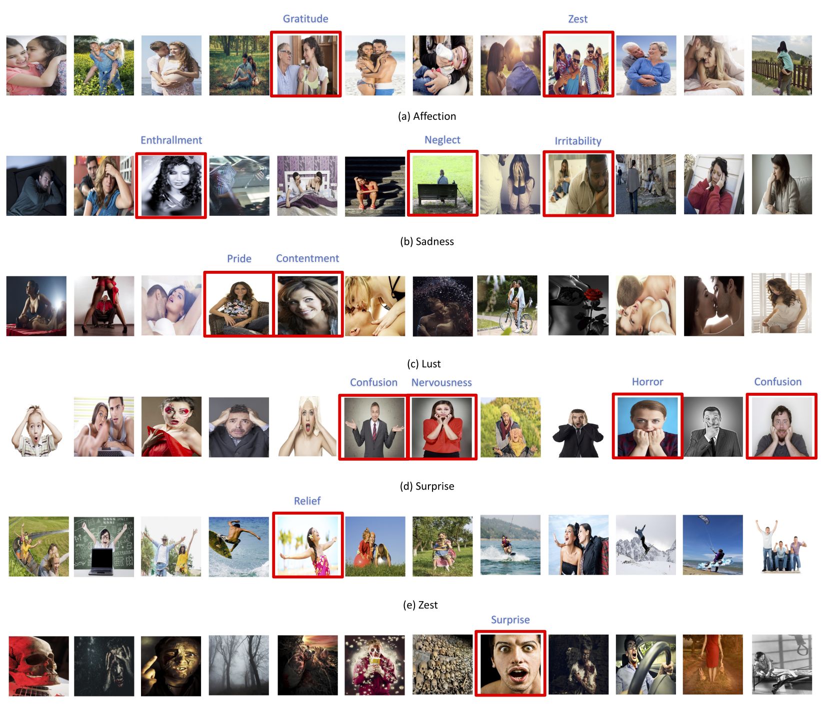


Fig. 8. Sample results among the top-5K predictions for each category by the curriculum guided learning based classifier. The images are listed in descending order of confidence. False alarms are shown with red borders and ground-truth labels at the top. Best viewed in color.

References

1. Ekman, P.: An argument for basic emotions. Cognition & emotion (1992)
2. Gan, C., Yao, T., Yang, K., Yang, Y., Mei, T.: You lead, we exceed: Labor-free video concept learning by jointly exploiting web videos and images. In: CVPR (2016)
3. Joo, J., Li, W., Steen, F.F., Zhu, S.C.: Visual persuasion: Inferring communicative intents of images. In: CVPR (2014)
4. Kim, H.R., Kim, S.J., Lee, I.K.: Building emotional machines: Recognizing image emotions through deep neural networks. arXiv preprint arXiv:1705.07543 (2017)
5. Parrott, W.G.: Emotions in social psychology: Essential readings. Psychology Press (2001)
6. You, Q., Luo, J., Jin, H., Yang, J.: Robust image sentiment analysis using progressively trained and domain transferred deep networks. In: AAAI (2015)
7. You, Q., Luo, J., Jin, H., Yang, J.: Building a large scale dataset for image emotion recognition: The fine print and the benchmark. In: AAAI (2016)