Paper selected: Undoing the damage of Dataset Bias

Chaeeun Ryu 2018312824

1. What was the bias problem tackled and in what task does it appear?

The paper was mainly dealing with all kinds of diverse bias including selection bias, capture bias, and negative set bias in popular image datasets. The paper has distinguished bias into two kinds: object-specific bias and contextual bias. The bias problem of object recognition the authors raised was that data augmentation or increasing the number of data samples with other datasets often lead to a decrease in the performance of the model on a test set because the dataset is biased, when ideally, more data should lead to better generalization ability, referred to Torralba et al.[1]. The paper captured two main biases in a max-margin model, which are contextual bias and object-specific bias in object classification task(identifying the presence of an object in the image) and object detection task(localization of an object in the image), respectively.

2. Why is the tackled bias important?

The tackled bias is important because undoing the bias leads to better object recognition ability, particularly for object detection and image classification. The paper demonstrated that there is dataset-specific bias for each dataset, and the bias vectors it captured from the model outperformed the baseline model(SVM) by carrying out object detection tasks and classification tasks although the bias vectors were not trained to perform these tasks. This is verified from demonstrating that the images from the i-th dataset are better classified by the corresponding bias vectors in all training cases and experiments than by the bias vectors of other datasets. Also, the authors note that it is very difficult, if not impossible to get an unbiased training dataset, as suggested also by Ponce et al.[2], so dealing with bias is inevitable and important.

3. Write a couple of paragraphs explaining what you have understood about the proposed methods to detect or deal with the bias.

The paper proposed a discriminative framework which resembles a max-margin learning model(Support Vector Machine). The authors jointly combined two models: a visual world model that can generalize well on average for all datasets, and a biased model that works particularly well for each dataset, but has low generalization ability. The goal of the model is to properly control the generalization ability by weighing the importance of biases for each specific dataset.The dataset-specific weight vector(w_i)is calculated in a linearly additive way in a formula $w_i = w_{vw} + \Delta_i$ by combining visual world weight vectors(w_{vw}), which is common for all datasets and the bias vector of each dataset(Δ_i), which is different for every dataset.

As it resembles SVM referred to Evgeniou et al.[3], it aims to solve an objective function

$$min_{w_{vw},\Delta_i,\xi,\rho} \frac{1}{2}\left\|w_{vw}\right\|^2 + \frac{\lambda}{2}\sum_{i=1}^{n}\left\|\Delta_i\right\|^2 + C_1\sum_{i=1}^{n}\sum_{j=1}^{s_i}\xi_j^i + C_2\sum_{i=1}^{n}\sum_{j=1}^{s_i}\rho_j^i$$

of maximizing its margin while minimizing errors. The model calculates two weight vectors which are visual weight vectors(w_{vw}) and bias weight vectors(Δ_i) and calculates two losses ξ and ρ, each indicating the loss(ξ) incurred across all datasets when using w_{vw} and loss(ρ) occurred when incorrectly classified an example using bias vector Δ_i, respectively .Hyper parameter λ controls the weight between generalization ability and dataset-specific performance, thus increasing λ leads to increasing the influence of a common set of weights for all datasets, and the vice versa enlarges the influence of bias weight that is independent for each dataset. Also, hyperparameters C1 and C2 control the relative importance between two constraints of optimizing loss of visual world and individual datasets. C2 is set to be some fraction of C1 to model the trade-off relationship between the loss on each independent dataset and the visual world.

The model solves its optimization problem by using stochastic subgradient descent, updating the visual world weight and bias vector at the same time, respectively with a different weight updating formula. After defining the algorithm, the authors applied it to a classification task and object detection task.

In the experiments, the model outperformed the baseline (SVM) by 2.8% mAP (mean Average Precision) in classification task and 0.7% mAP for the object detection task. Overall, the model outperformed the baseline for both seen and unseen datasets in various experiments. The authors note that this is so because the model treats the data from each dataset as biased samples of the real world, thereby not having full reliance on what it learned from each dataset but also taking the knowledge of the visual world into account. Surprisingly from the experiments, the learned bias had strong correlation to the specific dataset it belonged to, and could carry out image recognition tasks for the unseen samples of the test set without being trained to do so. In conclusion, explicitly modeling these biases is useful for bias mitigation and leads to better generalization ability.

4. Besides the mentioned bias, state and explain 5 other concepts that you have learned while reading the selected paper or reference papers. (10)

Firstly, we learned about transfer learning from Pan et al. [6], which leverages knowledge, earned from the previous training of the model, for training newer models to related tasks. The three kinds of transfer learning are inductive transfer learning, unsupervised transfer learning, and transductive transfer learning. Transfer learning is particularly beneficial when there is insufficient data in the new training process because we can use the knowledge gathered from the previous training as an additional input for this new task. Secondly, domain adaptation, which is a subclass of transfer learning, uses labeled data from one or more source domains and applies it to the target domain for solving new tasks. Domain adaptation mentioned in Saenko et al. [7] is different from transfer learning in that the feature space of source domain and target domain is always the same, while for transfer learning, it could be different. The third concept we learned is slack variables, which are used when converting inequality constraints to equality constraints in order to solve optimization problems. In the framework of this paper, two slack variables were used to allow two constraints to be violated. Since the model used in this experiment as a baseline was SVM, this was the fourth concept we learned in order to understand this paper. A Support Vector Machine, described in Kruczkowski et al.'s paper[9], is a supervised machine learning algorithm used in classification and regression problems. Each data point is plotted in an n-dimensional space, where n is the number of features, and then the algorithm finds the hyperplane that would separate the plotted data into each class. The goal is to find the greatest margin from the hyperplane to the closest data point. Finally, the paper uses HOG templates to visualize the influence of biases. We learned that Histogram of Oriented Gradients(HOG) is a feature descriptor that is used in object detection, image segmentation, and other visual processing tasks. It counts and visualizes occurrences of gradient orientation in an image, as described by Ahn et al.[8]. A gradient, in this context, is a directional change in the color of an image.

5. What is your overall opinion about this paper? Are there things you agree or disagree with? Is there something you want to add about tackling the bias problem addressed by the paper? (5)

It was interesting how the bias vectors are learned to carry out image recognition tasks without being trained to do so. The images and graphs included were also interesting, and they helped us understand the paper. However, one improvement we wanted to suggest to the paper was that as the authors first brought up the problem about how augmenting the data does not lead to better generalization ability of the model, it would have been better if they talked about solving the problem they suggested by addressing how much they were able to solve their initial problem through their discriminative framework. We would have also liked to see a more social perspective of the bias they dealt with, for example, how these biases affect certain groups of people and from what social aspect they incurred.

References

1. Torralba, A., Efros, A.A.: Unbiased look at dataset bias. In: CVPR, pp. 1521–1528 (2011)
2. Ponce, J., Berg, T.L., Everingham, M., Forsyth, D., Hebert, M., Lazebnik, S., Marszalek, M., Schmid, C., Russell, B.C., Torralba, A., Williams, C.K.I., Zhang, J., Zisserman, A.: Dataset Issues in Object Recognition. In: Ponce, J., Hebert, M., Schmid, C., Zisserman, A. (eds.) Toward Category-Level Object Recognition. LNCS, vol. 4170, pp. 29–48. Springer, Heidelberg (2006)
3. Evgeniou, T., Pontil, M.: Regularized multi–task learning. In: 10th ACM SIGKDD International Conf. Knowledge Discovery and Data Mining, pp. 109–117 (2004)
4. dbredep, undoing_bias_torch, 2019, Github repository, https://github.com/adikhosla/undoing-bias
5. Aditya Khosla, Undoing the Damage of Dataset Bias, 2013, Github repository, https://github.com/dbredep/undoing_bias_pytorch
6. Pan, S.J., Yang, Q.: A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering 22 (2010)
7. Saenko, K., Kulis, B., Fritz, M., Darrell, T.: Adapting Visual Category Models to New Domains. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part IV. LNCS, vol. 6314, pp. 213–226. Springer, Heidelberg (2010)
8. Ahn, S., Park, J., & Chong, J.: Blurring Image Quality Assessment Method Based on Histogram of Gradient. *Proceedings of the 19th Brazilian Symposium on Multimedia and the Web*, 181–184. Presented at the Salvador, Brazil. (2013) doi:10.1145/2526188.2526226
9. Kruczkowski, M., & Szynkiewicz, E. N.: Support Vector Machine for Malware Analysis and Classification. *Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT) - Volume 02*, 415–420. (2014) doi:10.1109/WI-IAT.2014.127