View Reviews

Paper ID
1232

Paper Title
Exact Feature Distribution Matching for Arbitrary Style Transfer and Domain Generalization

Reviewer #1

Questions

2. Summary. In 5-7 sentences, describe the key ideas, experimental or theoretical results, and their significance.

They propose a simple, yet powerful method for arbitrary style transfer and domain generalization tasks. To solve the limitation of existing approaches, which limit the alignment of feature distribution in either first-order or second-order levels, they propose to perform exact feature distribution matching. This technique can be simply implemented by sorting feature channels of input content and style examples.

Despite the simplicity of their approach, their proposed method shows effectiveness in style transfer, domain generalization in image classification, and person re-ID. They also include analysis on the running time of their approach.

3. Strengths. Consider the significance of key ideas, experimental or theoretical validation, writing quality, data contribution. Explain clearly why these aspects of the paper are valuable. Short bullet lists do NOT suffice.

1. Their idea of achieving exact histogram matching is novel and reasonable. Although the idea of EHM and sort-matching algorithm is not theirs, but the application to style transfer and domain generalization task is new.

2. Their proposed method is simple, hence, easy to plug-in various models, but shows strong performance on many tasks. Especially, I think the effectiveness in image recognition tasks are clear and solid.

3. Their experimental validation is solid. The experiments are widely done including different tasks and architectures with additional analysis.

4. This paper is well-organized, easy to follow.

4. Weaknesses. Consider the significance of key ideas, experimental or theoretical validation, writing quality, data contribution. Clearly explain why these are weak aspects of the paper, e.g. why a specific prior work has already demonstrated the key contributions, or why the experiments are insufficient to validate the claims, etc. Short bullet lists do NOT suffice.

1. In domain generalization experiments, their main baseline is MixStyle since the formulation of the method is similar and directly comparable. However, they may lack some baselines such as [1][2][3]. These methods rely on shuffling style features when normalizing features. Especially, [1] seems to be directly comparable to this submission since the technique of [1] can be plugged in ResNet. I do not think the existence of [1][2][3] degrades the novelty of this submission, but the comparison to these methods will be interesting to see.

2. It may be also worth noting the training time of the proposed approach, i.e., how much the sort-matching increases training time compared to other baselines.

[1] CrossNorm and SelfNorm for Generalization under Distribution Shifts, ICCV2021
[2] Style Normalization and Restitution for Generalizable Person Re-identification, CVPR 2020
[3] Adversarially Adaptive Normalization for Single Domain Generalization, CVPR2021

5. Paper rating (pre-rebuttal).

Strong Accept

7. Justification of rating. What are the most important factors in your rating?

Although I have a very small concern as shown in weakness, I do not have a reason to reject this paper. This paper presents one interesting direction in domain generalization and style-transfer.

8. Are there any serious ethical/privacy/transparency/fairness concerns? If yes, please also discuss below in Question 9.

No

9. Limitations and Societal Impact. Have the authors adequately addressed the limitations and potential negative societal impact of their work? Discuss any serious ethical/privacy/transparency/fairness concerns here. Also discuss if there are important limitations that are not apparent from the paper.

They need to include limitations and societal impact section.

10. Is the contribution of a new dataset a main claim for this paper? Have the authors indicated so in the submission form?

No dataset contribution claim

Reviewer #2

Questions

2. Summary. In 5-7 sentences, describe the key ideas, experimental or theoretical results, and their significance.

This paper proposes using exact histogram matching technique to match a feature distribution of an example to that of the other for the tasks of example to example style transfer and domain generalization. It is argued that image feature distribution is too complex to be modeled by Gaussian, so exact histogram matching is more effective than moment matching such as those done by previous works. Experiments demonstrate good qualitative synthesis result in style transfer task and competitive result in domain generalization task.

3. Strengths. Consider the significance of key ideas, experimental or theoretical validation, writing quality, data contribution. Explain clearly why these aspects of the paper are valuable. Short bullet lists do NOT suffice.

It has been observed that matching feature distribution statistics results in similar "style" between images. The central idea of the paper is to do exact feature distribution matching using exact histogram matching techniques. This is very interesting to explore, does better matching of feature distribution result in better matching in style between images?

On style transfer task, the proposed method produces good result, some with less artifacts & distortion than other methods. On domain generalization task, it outperforms state of art mixstyle. I think, as an augmentation technique, histogram matching produce more diverged examples than matching mean/std only.

Presentation's of paper is good, writing is comprehensive and figures are nicely done.

4. Weaknesses. Consider the significance of key ideas, experimental or theoretical validation, writing quality, data contribution. Clearly explain why these are weak aspects of the paper, e.g. why a specific prior work has already demonstrated the key contributions, or why the experiments are insufficient to validate the claims, etc. Short bullet lists do NOT suffice.

I don't see serious flaws in proposed method.

The main idea digs deeper what has been proposed (matching feature distribution for style transfer). While it is novel and interesting, it is more of an incremental improvement.

5. Paper rating (pre-rebuttal).

Weak Accept

7. Justification of rating. What are the most important factors in your rating?
The proposal is interesting and well executed in this work. Experimentation is somewhat extensive and demonstrates state of the art result.

8. Are there any serious ethical/privacy/transparency/fairness concerns? If yes, please also discuss below in Question 9.
No

9. Limitations and Societal Impact. Have the authors adequately addressed the limitations and potential negative societal impact of their work? Discuss any serious ethical/privacy/transparency/fairness concerns here. Also discuss if there are important limitations that are not apparent from the paper.
N/A

10. Is the contribution of a new dataset a main claim for this paper? Have the authors indicated so in the submission form?
No dataset contribution claim

Reviewer #3

Questions

2. Summary. In 5-7 sentences, describe the key ideas, experimental or theoretical results, and their significance.
This paper propose a extact feature distribution matching method for style transfer (AST) and domain generalization (DG) tasks. The main contribution is that it doesn't assume the feature distribution to follow Gaussian distribution. So it would improve the downstream tasks. The paper demonstrate the improvement in a set of experiments. It would be useful to further advance the SOTA of AST and DG results.

3. Strengths. Consider the significance of key ideas, experimental or theoretical validation, writing quality, data contribution. Explain clearly why these aspects of the paper are valuable. Short bullet lists do NOT suffice.
The idea of exact feature distribution matching to improve AST and DG seems reasonable. There are some experimental validation but can be improved.

The paper is well written. The codes are also provided to reproduce the results.

4. Weaknesses. Consider the significance of key ideas, experimental or theoretical validation, writing quality, data contribution. Clearly explain why these are weak aspects of the paper, e.g. why a specific prior work has already demonstrated the key contributions, or why the experiments are insufficient to validate the claims, etc. Short bullet lists do NOT suffice.
There are several aspects to be improved. First, in the AST experiments, a formal user study should be conducted to evaluate whether and how much the proposed method is better than the other methods. Only show some sample results won't be sufficient for readers to gain quantitative insight about the paper's contribution. Second, it's understandable that effect matching would generate better results than approximate matching like AdaIN. However, the higher compuational complexity would be the major concern. The paper only shows an example test that the two methods have roughly the same computation time and said the feature size is usually finite. The argument is relatively weak. There should be an analysis of the featuer size range of commonly used methods, sufficiently sample the range and compare the computation time of the two methods.

5. Paper rating (pre-rebuttal).
Borderline

7. Justification of rating. What are the most important factors in your rating?
The experiment should be improved to include a user study for better assess AST quality. The discussion of computational complexity should also be improved.

8. Are there any serious ethical/privacy/transparency/fairness concerns? If yes, please also discuss below in Question 9.

No

9. Limitations and Societal Impact. Have the authors adequately addressed the limitations and potential negative societal impact of their work? Discuss any serious ethical/privacy/transparency/fairness concerns here. Also discuss if there are important limitations that are not apparent from the paper.

There are discussion but can be further improved.

10. Is the contribution of a new dataset a main claim for this paper? Have the authors indicated so in the submission form?

No dataset contribution claim

14. Final recommendation based on ALL the reviews, rebuttal, and discussion (post-rebuttal).

Weak Accept

15. Final justification (post-rebuttal).

The author feedback has provide concrete evidence to address my questions. So I will give the final rating as weak accept.