Unsupervised shadow removal using generative adversarial networks

Vladyslav Andronik

Shadow removal problem

- Shadow removal is a challenging problem which demands detecting shadow first and filling it using background information.
- Detected shadows can help with better scene understanding[1,2]. However, they degrade the performance of downstream computer vision tasks as object detection[3], tracking[4] etc.

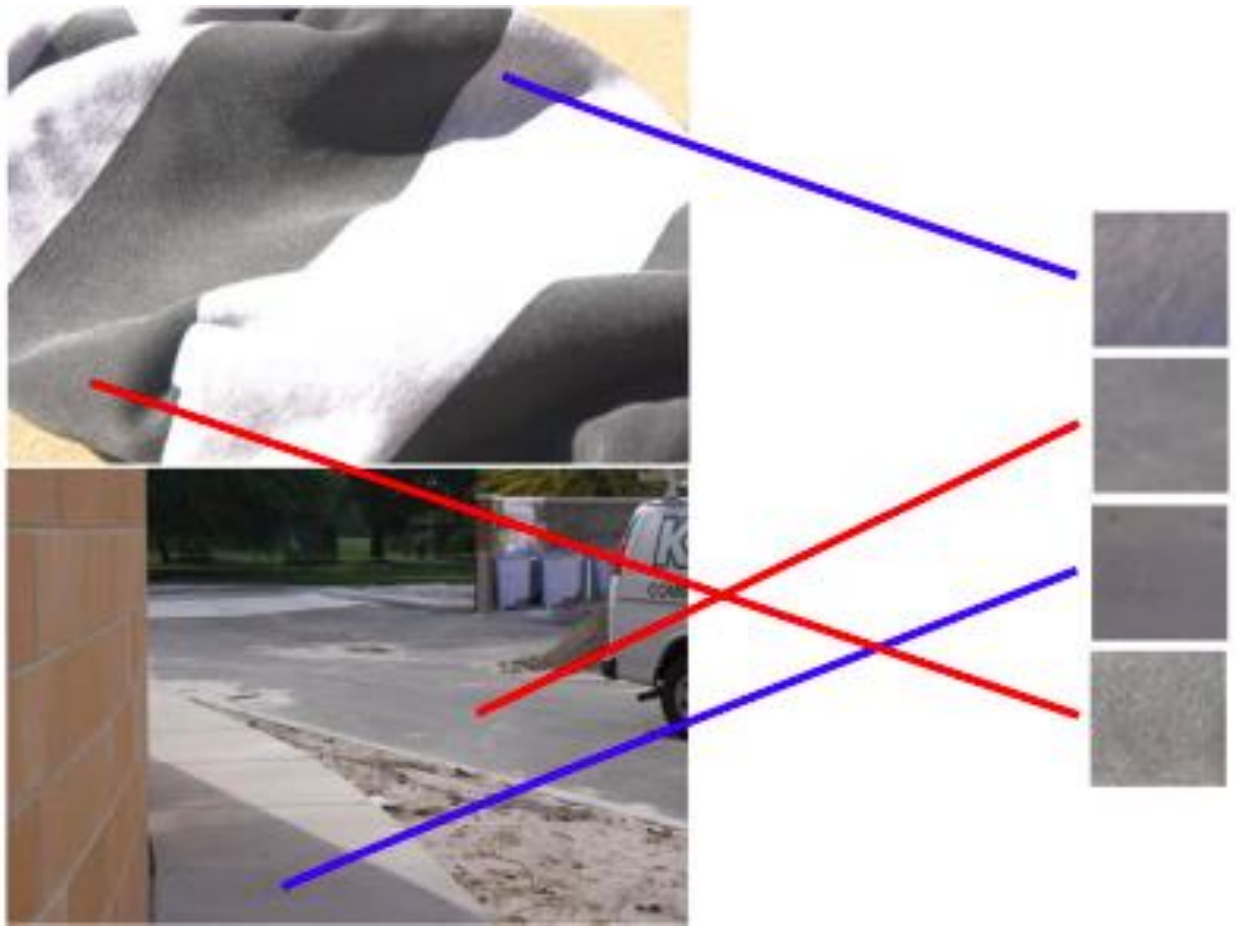


Figure 1: It is hard to tell from local perspective where the shadow region is located[5].

Current state of research

- Nowadays, the sphere is dominated by deep learning approaches. However, most of them relies on tediously annotated datasets where one needs to fix the camera and add/remove the objects to obtain the pair of shadow and shadow free images.
- We also argue that this approach is constraining the complexity of scenes and might expose undesired discrepancy in illuminance and color (see Figure 2).
- For that reason we conducted research on using unsupervised learning for shadow removal.

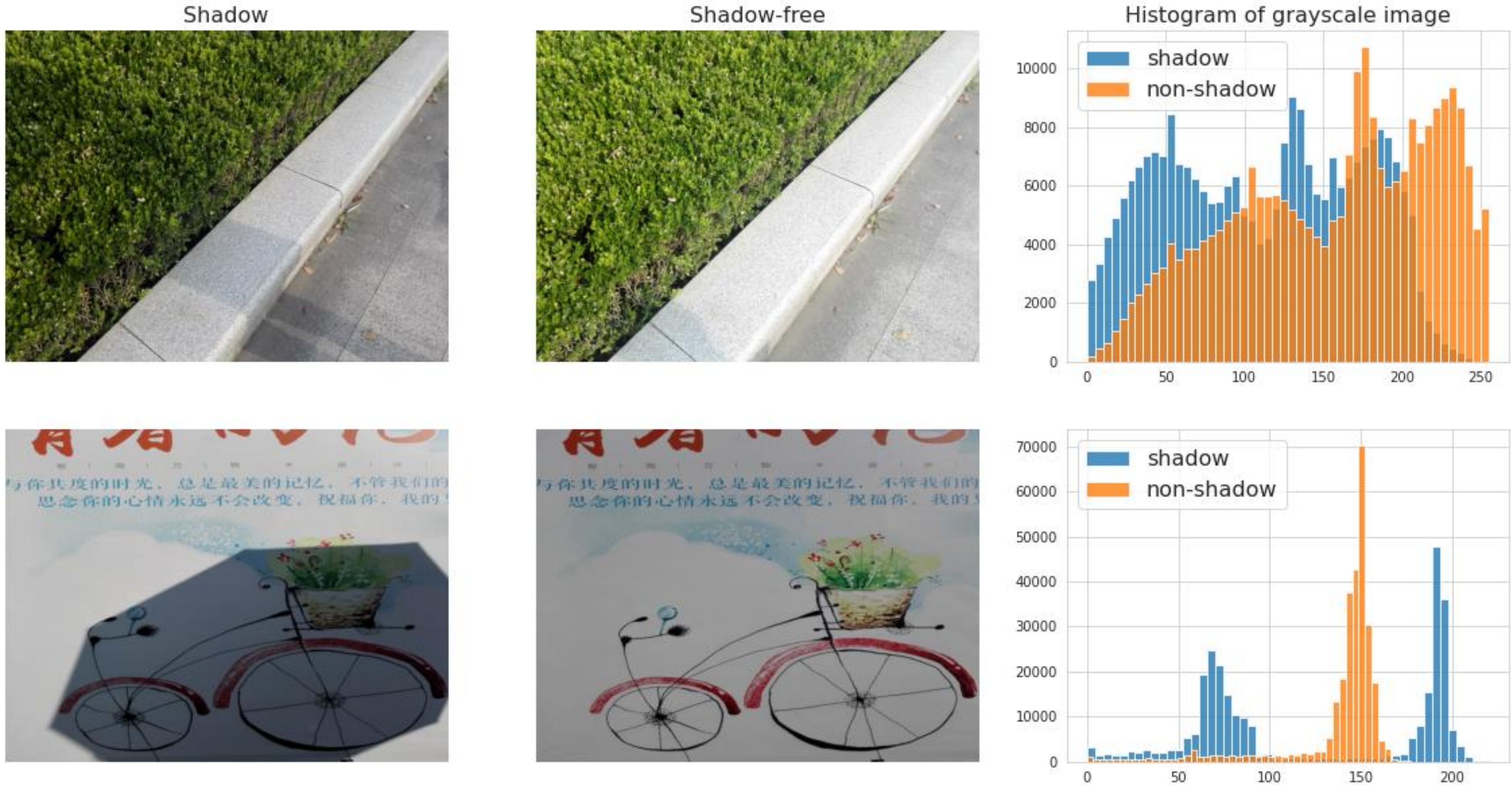


Figure 2:
Change in color between shadow and shadow free images comparing non-shadow regions.

Method

- We trained four generative adversarial networks in a single end-to-end framework with cycle consistency losses. Multi context aggregation and attention modules are presented for more efficient shadow removal.
- Dilated convolutions are used to extend the receptive field at the second part of the bottleneck layer (see Figure 3, next slide) which helps with more efficient shadow removal.
- Attention module is integrated in both generator and discriminator to help distinguishing between domains. We also use attention maps to generate shadow images in the inverse transform generator.
- Shadow generator network is conditioned with attention map and shadow binary mask received from shadow removal network to raise the generalization capacity.

Method

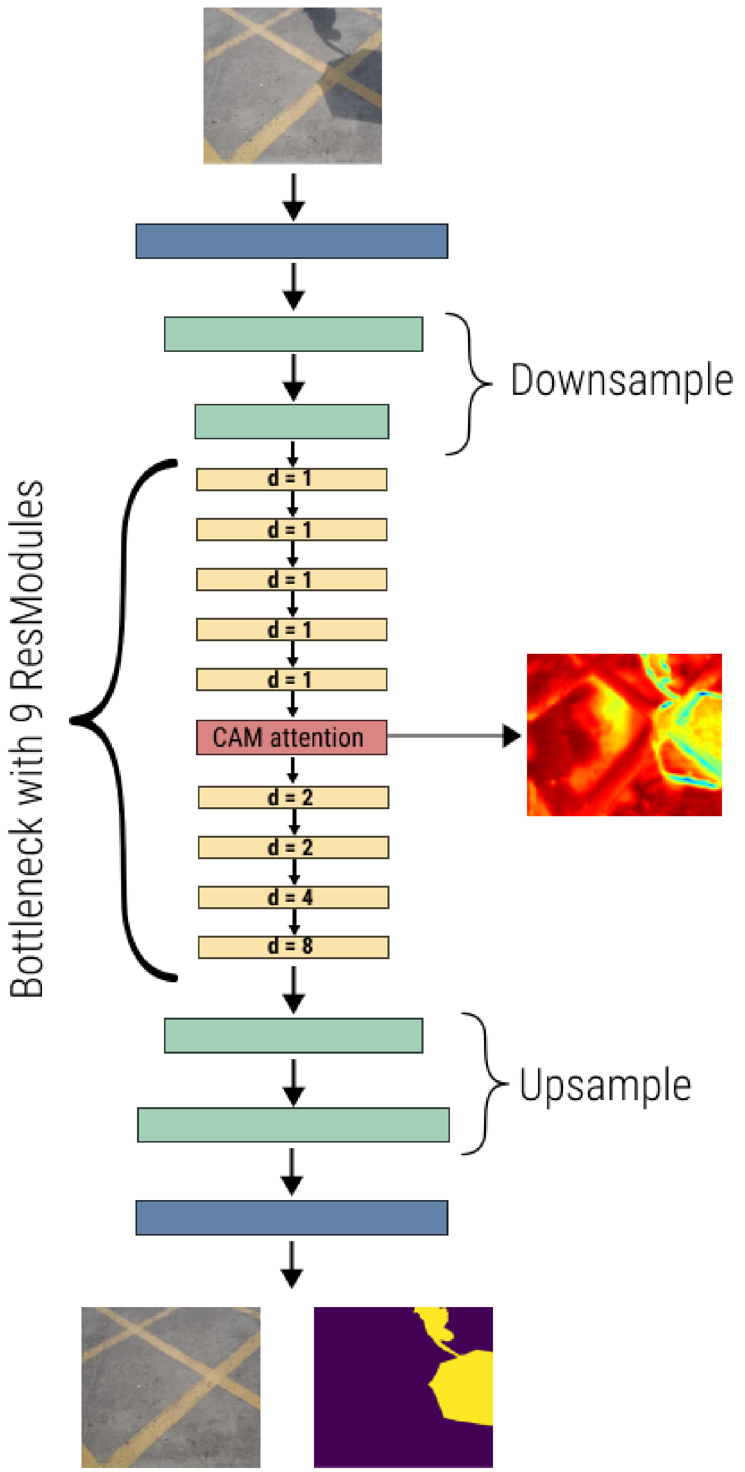


Figure 3: Shadow removal generator

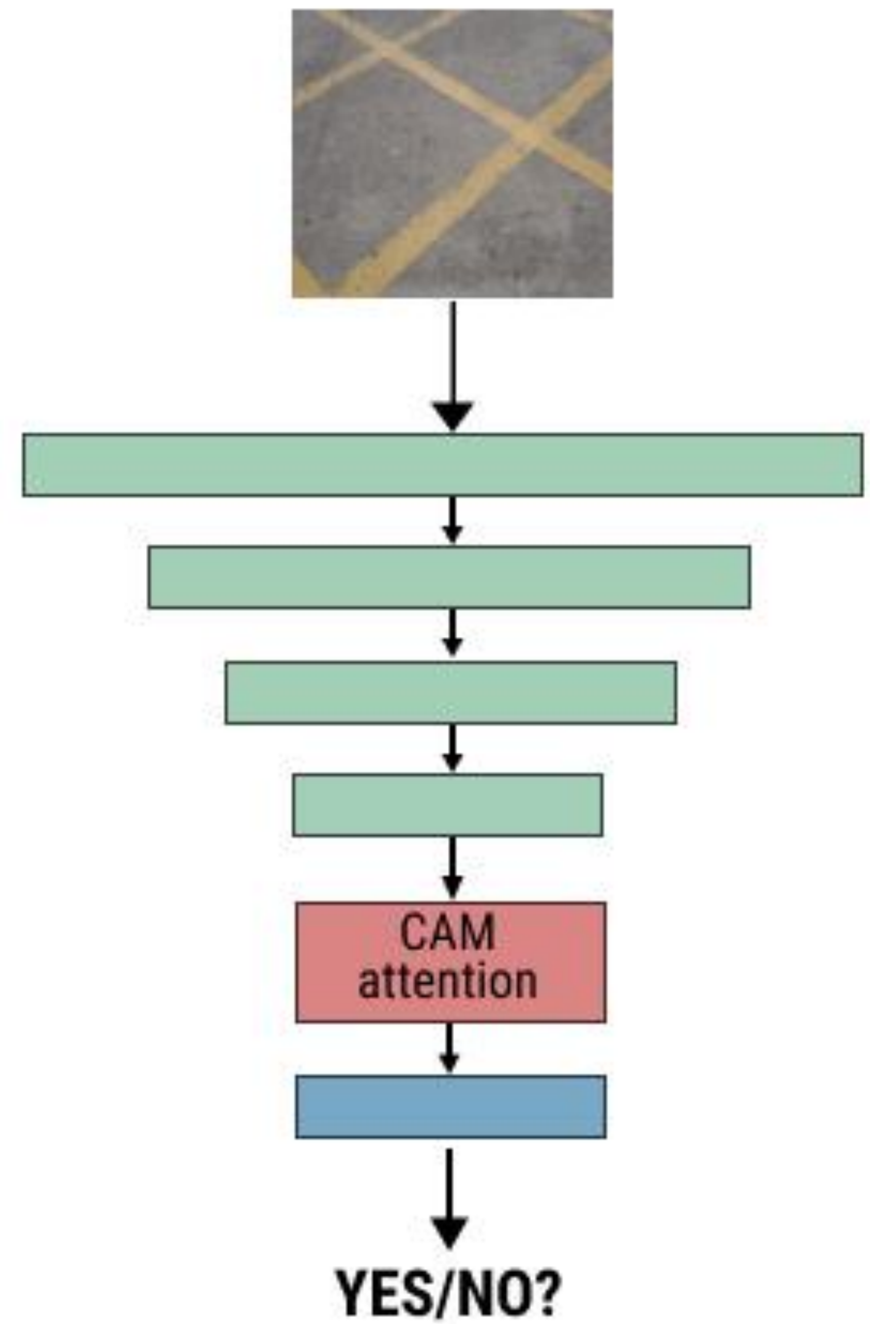


Figure 4: Shadow free Discriminator

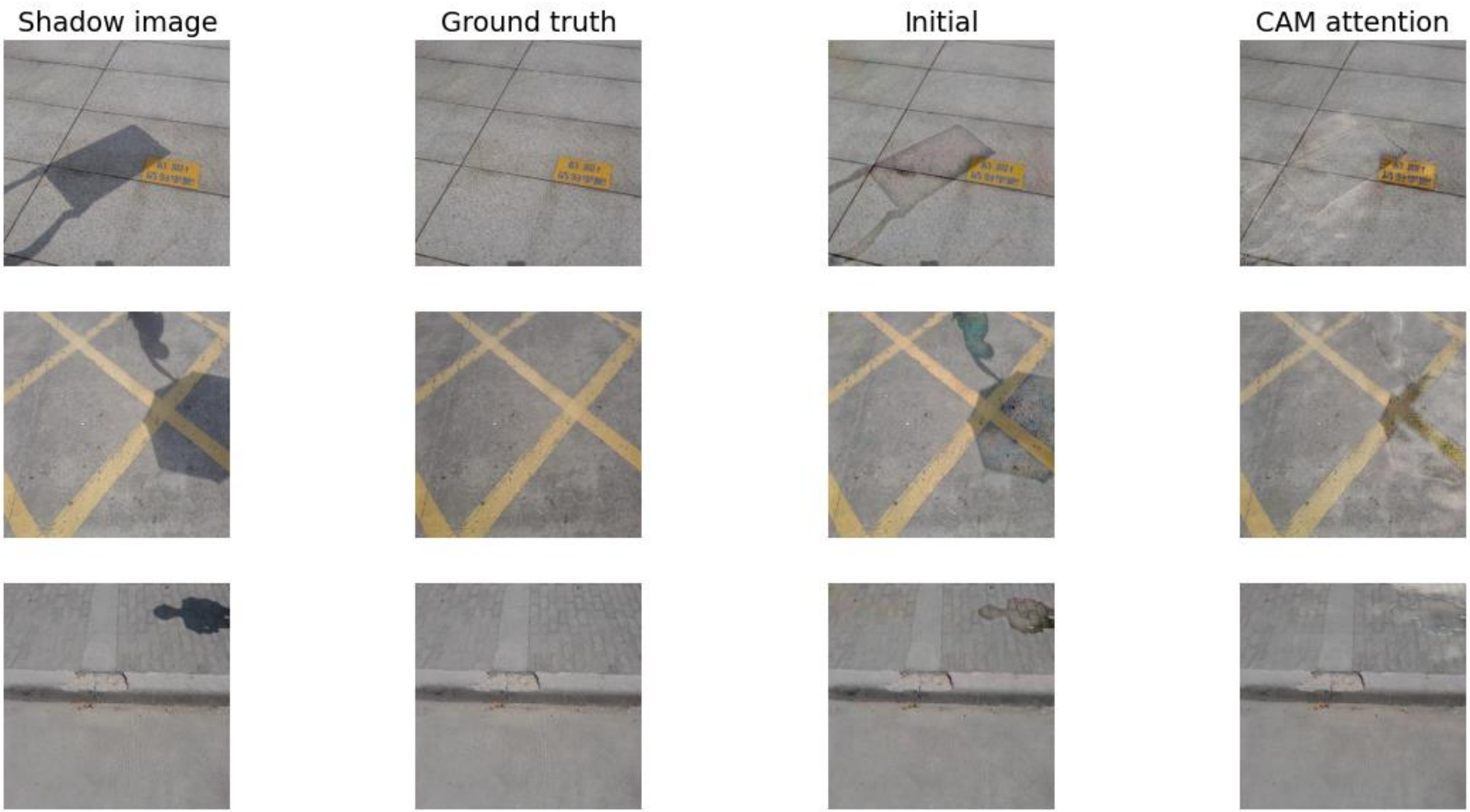
Results

- We conducted experiments on proper use of dilated convolutions coupled with attention modules. Dilated convolutions did not give gain in quality however we proceeded using them with attention.
- Attention modules allowed to reach significantly better results in shadow removal(presented by first two columns in Table 1) while degrading the detection performance.
- Our generative attentional networks show capacity outperforming the existing approach in shadow removal task.

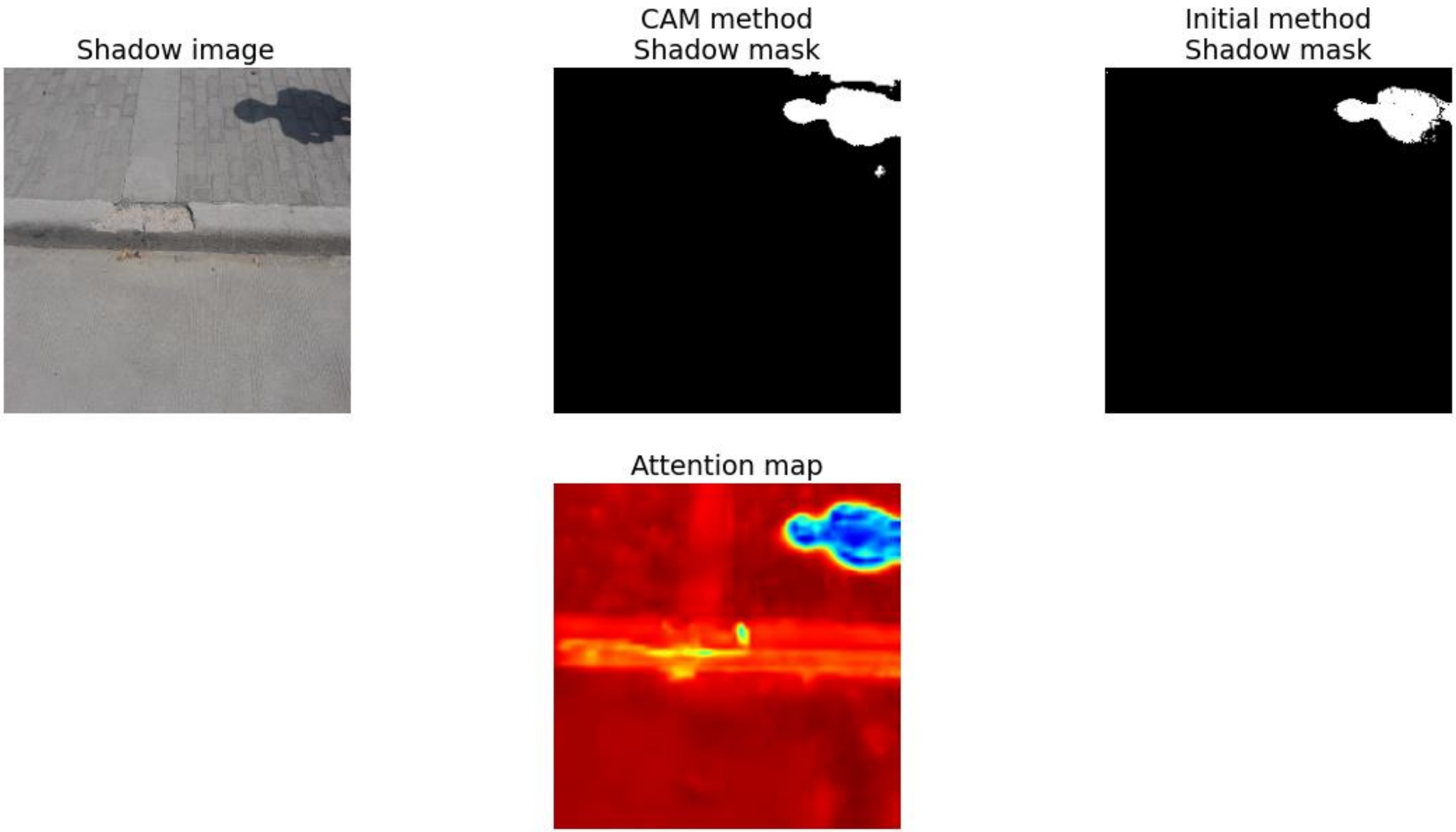
Methods	Global RMSE	Shadow region RMSE	IOU
MaskShadowGAN[14] with our training strategy	3.0099	23.6703	**80.0894**
+dilation in $G_{s\to f}$ and $G_{f\to s}$	3.1253	26.9753	75.5382
+CAM attention in all G and D	2.3902	**15.0150**	73.8277
+CAM attention in $G_{s\to f}$ and D_f giving A_s as input to D_s (*)	2.3139	15.7047	71.1142
+A_s with no M_s to (*)	2.3261	15.5398	70.3951
+CAM weights from $G_{s\to f}$ to $G_{f\to s}$	**2.2436**	15.6087	70.7965

Table 1:
Quantitative comparison of different experiments. All models are evaluated on validation set. RMSE indicates error in shadow removal while IOU — shadow detection quality. Values in bold show best score.

Results



Results



Shadow image
CAM method
Shadow mask
Initial method
Shadow mask
Attention map

Results analysis

- One could see a strong evidence that there is a trade-off between shadow removal and detection performances.
- We discovered that solutions with attention module presents instability issues whilst drastically improving shadow removal. However, they suffer from over detection and corruption in generated samples.
- On the other hand, solutions with no attention are better in shadow detection, but struggle to efficiently use background information, thus being worse in removal task.

Future work

- After analyzing the results, we assume that over detection problem arises from signal vanishing in the bottleneck layer. For that reason, we will look into using the multiple skip-connections that might be crucial in end-to-end learning.
- To cope with stability and consistency issues, we will research different normalization techniques that proved to be successful in generative methods.
- Last but not least, style generators have showed capacity for general purpose usage. That is why we would try them to encourage the networks to save the global context in generated samples.
- We will extend our approach to datasets with more complex scenes to better understand its application to real-world problems.

Conclusion

- This work presented a solution to unsupervised shadow removal problem with the use of generative adversarial networks with attention modules and multi context aggregation.
- Our network produces better results compared to the existing approach in the field.
- Analysis showed that attention maps obtained from auxiliary classifier encourage the networks to concentrate on more distinctive regions between domains.
- However, GANs demand more accurate and consistent architecture to solve the problem in a more efficient way.
- We have also showed how attention modules can improve the quality of shadow removal while introducing the problems with the shadow over-detection. For that reason we will research further to address this problem.

Thank you!

References

(1) D. L. Waltz, “Generating semantic descriptions from drawings of scenes with shadows,” Cambridge, MA, USA, Tech. Rep., 1972.

(2) H. Barrow and J. Tenenbaum, “Recovering intrinsic scene characteristics from images,” in Comp. Vision Systems, 1978.

(3) I. Mikic, P. C. Cosman, G. T. Kogut and M. M. Trivedi, "Moving shadow and object detection in traffic scenes," Proceedings 15th International Conference on Pattern Recognition. ICPR-2000, Barcelona, Spain, 2000, pp. 321-324 vol.1, doi: 10.1109/ICPR.2000.905341.

(4) R. Cucchiara, C. Grana, M. Piccardi, A. Prati and S. Sirotti, "Improving shadow suppression in moving object detection with HSV color information," ITSC 2001. 2001 IEEE Intelligent Transportation Systems. Proceedings (Cat. No.01TH8585)

(5) R. Guo, Q. Dai and D. Hoiem, "Paired Regions for Shadow Detection and Removal," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 12, pp. 2956-2967, Dec. 2013, doi: 10.1109/TPAMI.2012.214.