Using Generative Adversarial Networks (GANs) to Improve Autonomous Driving Under Adverse Weather Conditions

Trisha Thakur **Jonathan Tang** **Uttkarsh Berwal**

Department of Computer Science
University of Toronto
Toronto, ON M5S
{trisha.thakur, jonathanj.tang, uttkarsh.berwal}@mail.utoronto.ca

Abstract

A current barrier to the widespread adoption of autonomous driving systems in everyday use is their resilience under adverse weather conditions, as most state-of-the-art systems experience severe performance degradation in such circumstances. Towards a remedy, we investigate the use of Generative Adversarial Networks (GANs) in improving the quality of adverse weather images to enhance the performance in a fundamental subproblem of autonomous driving, object detection. We compare the effectiveness of conditional GAN (Pix2pix) and CycleGAN architectures in pre-processing adverse weather images and evaluate performance improvement on a representative object detection model DETR. We discover minute performance enhancements under Pix2pix pre-processing and significant improvements under the CycleGAN pre-processing.

1 Introduction

Weather conditions have significant impacts on road safety and daily traffic conditions. With the advent of autonomous driving systems, safety plays a critical role in their acceptance into the society. Such systems are required to be robust to variations in environmental conditions such as adverse weather conditions and lighting variations for widespread adoption. We focus on one of the sub-problems concerning autonomous driving, being image-based object detection, which faces challenges when regions of input images are obscured by weather conditions such as rain, fog, snow, and sand. Most object detection models of autonomous vehicles are trained on datasets without significant adverse weather conditions, making them unreliable when adverse situations arise. In the past decade, a variety of Generative Adversarial Networks (GANs) variants have arisen that can learn image-to-image translations, transforming images from one style to another including the addition and removal of adverse weather conditions to an image. A number of papers have recognized the potential of GANs to improve performance of autonomous driving systems [5, 10, 8].

In this project, we investigate whether GANs can be used to improve image-based object detection performance in autonomous driving applications, by preprocessing images to remove the effects of adverse weather conditions. We will compare two existing GAN architectures, CycleGANs and Conditional GANs, by examining the performance of existing object detection architectures on images from adverse weather datasets that are processed by the two GAN architectures. Our main contribution is providing experimental evidence for the use of the two GAN architectures in improving object detection with deweathered images in the autonomous driving domain.

Preprint. Under review.

2 Related Work

Various powerful Generative Adversarial Networks (GANs) architectures have been developed for performing **image-to-image translation** tasks [3, 11, 2], wherein the models are trained to learn a mapping between two different image styles and can be used to translate one type of image to another. The **Pix2Pix** model is a conditional GAN [3] that has proven to be useful in converting sketches to photos, day scenes to night scenes, etc. While the **CycleGAN** is another architecture [11] that builds on the Pix2Pix model with a cycle consistency loss that has also been successful in learning the mapping between aerial photos of streets and the corresponding GoogleMap images.

In terms of adverse weather condition images and their normal weather counterparts, GANs have demonstrated the capability of removing or improving images with adverse weather conditions in various specific scenarios, creating realistic replicas of the original. A conditional GAN architecture was developed for image deraining which brought about an improvement in the performance of the selected object detection model [10]. In the area of autonomous driving object detection, there are several state of art models that are constantly being improved upon. For the purpose of this study, we made use of a more general purpose object detection, Facebook DE:TR end-to-end object detection model [1]. This model was found to be performing better than Faster R-CNN which has been used for autonomous driving object detection applications [9].

3 Method

The basic architecture of GANs rely on two networks 'Generator' and 'Discriminator'. The job of the generator is to generate new $fake$ images, whereas the discriminator's job is to differentiate between $fake$ and $real$ images. In our paper, we are investigating two types of GANs -

Conditional GAN - Pix2Pix: When using GAN we can not control which image to generate, therefore, we use Conditional GAN which helps us in conditional generation of the images. We simply have to pass an extra parameter to the generator and discrminator to do this. The first part of our investigation relies on using conditional GAN, which is borrowed from the famous Pix2Pix model [3] for image-to-image translation to replicate bad weather conditions. This model uses a conditional GAN objective with a reconstruction loss.

$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{x,y}[\log D(x, y)] + \mathbb{E}_{x,z}[\log(1 - D(x, G(x, y))] -$ Conditional GAN objective [3]

$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y,z}[||y - G(x, z)||_1] -$ Reconstruction Term [3]

This gives us the final objective as $- G* = arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda\mathcal{L}_{L1}(G)$ [3]

Here, x denotes observed images, y represents output images, z is the noise vector, and D is the discriminator.

CycleGAN : Like any other GAN, this model also has a generator and a discriminator, but in addition to that it utilizes a cycle consistency loss function which is introduced as regularization term. This model maps real to fake images and then, from the fake images we try to regenerate the real ones, therefore we have an inverse mapping. Our objective is to learn the mapping between two domains X and Y, given by the function $G : X \to Y$. The objective function for the discriminator D_Y, and for mapping $X \to Y$, is given by [11] $- \mathcal{L}_{GAN}(G, D_Y, X, Y) = \mathbb{E}_x[\log(1 - D_Y G(X))] + \mathbb{E}_y[\log D_Y G(Y))]$

Similarly, for the mapping [11] $F : Y \to X$, and the discriminator D_X, it can be written as $\mathcal{L}_{GAN}(G, D_X, Y, X)$

$\mathcal{L}_{cyc}(G, F) = \mathbb{E}_x[||F(G(x)) - x||_1] + \mathbb{E}_y[||G(F(y)) - y||_1] -$ consistency loss [11]

The full objective is [11] $-$
$G* = arg \min_{G,F} \max_{D_X,D_Y} \mathcal{L}_{GAN}(G, D_Y, X, Y) + \mathcal{L}_{GAN}(G, D_X, Y, X) + \lambda\mathcal{L}_{cyc}(G, F)$

Here, X and Y denotes observed images, whereas D_X and D_Y are the discriminators.

After training our models using paired dataset for conditional GAN (Pix2Pix) and unpaired data for CycleGAN, we used them to generate fake images without the bad weather. Next step is to pass these

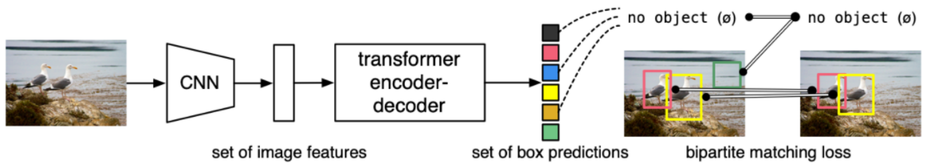

Figure 1: Network architecture DE:TR

images, through an object detection model to observe any improvement in the object detection. For this step, we are used the Facebook Research model DE:TR [1] which is a combination of a CNN and transformer.

4 Experiments and Discussion

4.1 Images Generated By Pix2Pix

Figure 2: Images generated by the Pix2Pix model are on the left of each pair and the images on the right are images for fog, rain, snow and sand-storm in the order of appearance starting from top left

In the first step, we tried generating fake images after removing the bad weather for different weather conditions, images shown in Fig. 2.The model was trained on raining and deraining images only because of the limitation of available data. Then it was tested in four weather conditions namely - fog, rain, snow, and sandstorm. From the images above, we can notice that there is a visible improvement in the case of rain and snow, a slight improvement in fog, and very minuscule improvement in sand-storm.

4.2 Images Generated By CycleGAN

In the second step, we repeated the process using cycle GAN, images shown in Fig. 3. In the case of fog, rain, snow CycleGAN removes the bad weather obscurity significantly and performs slightly better than conditional GAN. In the case of sand-storm, the improvement is not that visible but there is a change in coloration of the image.

4.3 Model Comparison with DE:TR

After passing the images through the DE:TR object detection model (shown in Fig. 4) we can observe that accuracy improved significantly when we pre-process the images using GANs to remove bad weather. We observe that more objects are detected with greater accuracies when we used CycleGAN compared to conditional GAN.

Figure 3: Images generated by the CycleGAN model are on the left of each pair and the images on the right are images for fog, rain, snow and sand-storm in the order of appearance starting from top left

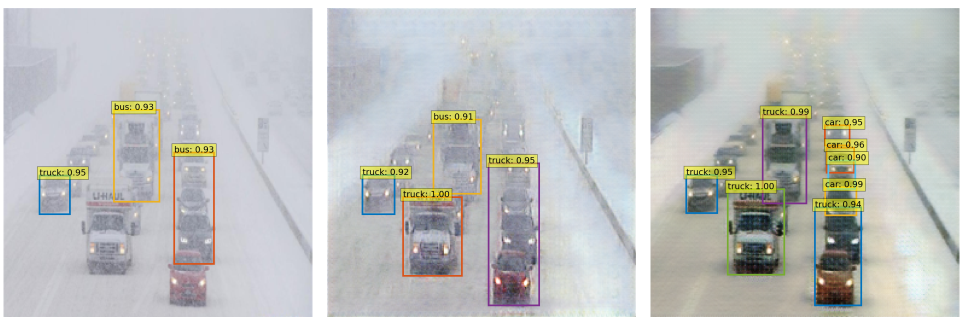

Figure 4: Images passed through the DE:TR model. First from left is the original weathered snow-storm image, second is a deweathered image generated by Pix2Pix and the last is the deweathered image generated by CycleGAN

Therefore, our hypothesis was right. In the future, we can look at various weather conditions and improve the GAN architecture even further.

To reproduce our results the code to this whole project can be found on here (https://github.com/berwalut/413finalproject). For our code base we used a combination of existing code from different sources. [4] [7] [6]

5 Conclusion

In this investigation, we found that both the Pix2Pix and CycleGAN models improved the accuracy of the object detection model DETR, with better performance of rain and snow weather conditions. CycleGAN had slightly better detection accruacies as compared to the Pix2Pix model. Since the models were only trained on rain and derain training set, it was expected that performance on non-rain weather conditions would not be significant. In future, collecting or generating a dataset for other weather conditions would be critical towards better detection.

Contributions

Trisha Thakur - Main focus was understanding and implementing conditional GAN architecture and the object detection model

Jonathan Tang - Main focus was understanding and implementing CycleGAN and object detection model

Uttkarsh Berwal - Main focus was understanding and implementing conditional and CycleGAN.

References

[1] Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, and Sergey Zagoruyko. End-to-end object detection with transformers. In Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm, editors, *Computer Vision - ECCV 2020 - 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part I*, volume 12346 of *Lecture Notes in Computer Science*, pages 213–229. Springer, 2020. doi: 10.1007/978-3-030-58452-8_13. URL `https://doi.org/10.1007/978-3-030-58452-8_13`.

[2] Yunjey Choi, Min-Je Choi, Munyoung Kim, Jung-Woo Ha, Sunghun Kim, and Jaegul Choo. Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018*, pages 8789–8797. IEEE Computer Society, 2018. doi: 10.1109/CVPR.2018.00916. URL `http://openaccess.thecvf.com/content_cvpr_2018/html/Choi_StarGAN_Unified_Generative_CVPR_2018_paper.html`.

[3] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 5967–5976. IEEE Computer Society, 2017. doi: 10.1109/CVPR.2017.632. URL `https://doi.org/10.1109/CVPR.2017.632`.

[4] junyanz. pytorch-cyclegan-and-pix2pi. `https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix`, 2017.

[5] Mourad A. Kenk and Mahmoud Hassaballah. DAWN: vehicle detection in adverse weather nature dataset. *CoRR*, abs/2008.05402, 2020. URL `https://arxiv.org/abs/2008.05402`.

[6] Facebook Research. detr. `https://github.com/facebookresearch/detr`, 2004.

[7] TheLethargicOwl. Single-image-de-raining-keras. `https://github.com/TheLethargicOwl/Single-Image-De-Raining-Keras`, 2019.

[8] Michal Uricár, Pavel Krízek, David Hurych, Ibrahim Sobh, Senthil Kumar Yogamani, and Patrick Denny. Yes, we GAN: applying adversarial techniques for autonomous driving. *CoRR*, abs/1902.03442, 2019. URL `http://arxiv.org/abs/1902.03442`.

[9] Gang Wang, Jingming Guo, Yupeng Chen, Ying Li, and Qian Xu. A pso and bfo-based learning strategy applied to faster r-cnn for object detection in autonomous driving. *IEEE Access*, 7: 18840–18859, 2019. doi: 10.1109/ACCESS.2019.2897283.

[10] He Zhang, Vishwanath Sindagi, and Vishal M. Patel. Image de-raining using a conditional generative adversarial network. *IEEE Trans. Circuits Syst. Video Technol.*, 30(11):3943–3956, 2020. doi: 10.1109/TCSVT.2019.2920407. URL `https://doi.org/10.1109/TCSVT.2019.2920407`.

[11] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017*, pages 2242–2251. IEEE Computer Society, 2017. doi: 10.1109/ICCV.2017.244. URL `https://doi.org/10.1109/ICCV.2017.244`.