Topic: cGAN

Adversarial Network for edge detection

Zhiliang Zeng, Ying Kin Yu, Kin Hong Wong
Department of Computer science and Engineering
The Chinese University of Hong Kong

2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR)

Introduction (1/2)

- **Domain/Fields:** A type of neural network called the conditional generative adversarial network (cGAN) to address the edge detection problem. cGAN is an innovative framework to do the image synthesis task.
- **Issues :** Producing a high-quality edge map directly without further processing, to avoid further steps of non maximum-suppression
- **Problems** Most of the existing neural network-based systems use the convolutional network or its variant. They usually produce thick edges and the application of non-maximum- suppression to suppress the edge is necessary

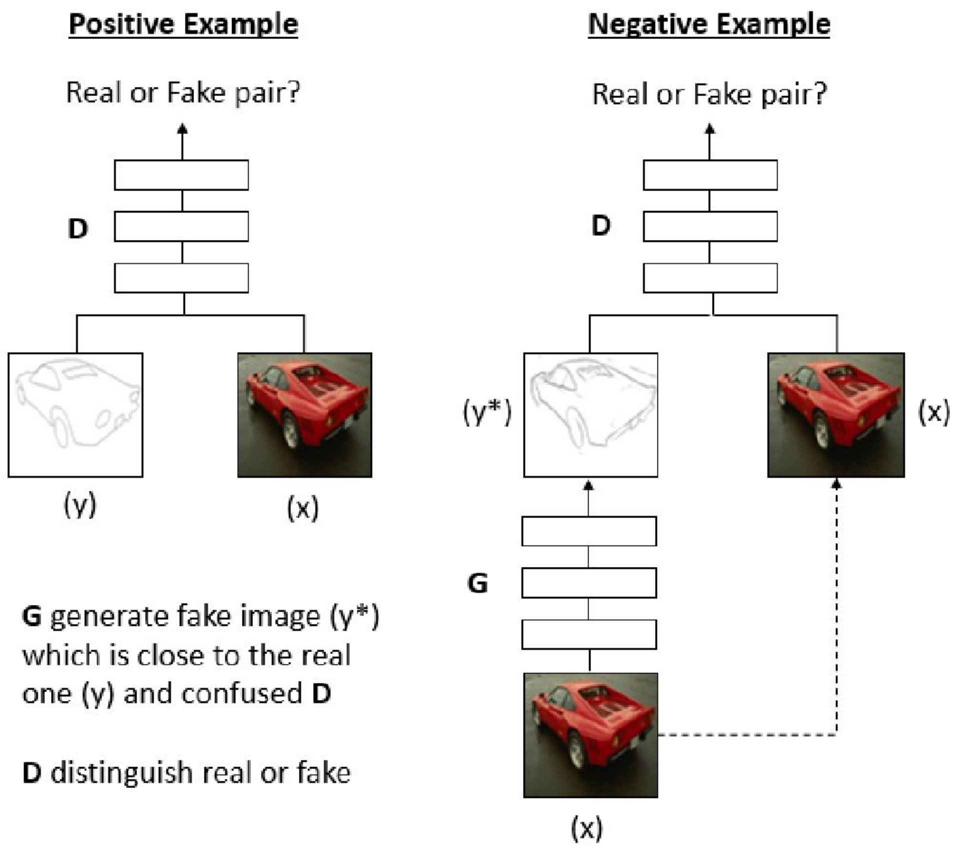
- **The Proposed Method**
  - This paper proposes an innovative edge detection algorithm based on generative adversarial network.
  - This paper trains vanilla cGAN with dataset BSD500A CNN model is also proposed for the system to learn by using the generated input images.
  - Since the original BSD500 dataset is a small one, it use augmentation to enlarge the size of the dataset method is implemented

- **The Contribution**
  - The proposed method suggests one of approaches about regarding how we can use adversarial network for edge detection.
  - Furthermore,cGAN can produce an image which is close to the real one

- GAN consists of two networks.
- One is the generator (G) and the other is the discriminator (D).
- After training, the generator (G) can produce images that are very close to the real images in the dataset.
- Define the images from the dataset as real images and the images produced by the generator are fake images.
- Objective of D is to distinguish whether the input image is real or not while the goal of G is to produce a fake image that looks like the real one.

The proposed method (2)

GAN is hard to train, therefore alternative method called the conditional GAN (cGAN) to alleviate the problem.



The proposed method (3)

- Directly predicting whether y or y is real or fake, our discriminator is used to predict whether a pair of (x; y*), or (x; y) is real pair or faker.[17]
- We use the generator to map input image from x to y.
- As seen in the image there is no noise z that is fed into the generator
- In vanilla GAN, G is the mapping noise z to one label image y in dataset.

Vanilla GAN

$$\mathcal{L}_{GAN} = \mathbb{E}_{x\sim p_{data}(x)}[logD(x)] \\ + \mathbb{E}_{z\sim p_z(z)}[1 - log(D(G(z)))]$$

GAN is hard to be implemented because the first equation cannot provide sufficient weight gradient information for the generator to learn well.

cGAN

$$\mathcal{L}_{cGAN} = \mathbb{E}_{(x,y)\sim p_{data}(x,y)}[logD(x,y)] \\ + \mathbb{E}_{x\sim p_{data}(x)}[log(1 - D(x, G(x)))]$$

Conditional GAN maps real image x in the dataset to y

Generator targets to predict the edge map, i.e. there are only two classes in the map

Positive side = edges

Negative class = background

$$\mathcal{L}1 = \begin{cases} \alpha\|G(x_i) - y_i\|, & \text{if} y_i > \eta \\ 0, & 0 < y_i \leq \eta \\ \beta\|G(x_i) - y_i\|, & \text{otherwise} \end{cases}$$

If the corresponding pixel value yi in the label is greater than

η, it is set to 0.5 in the program.

α is a hyper parameter that enlarges the L1 error of the pixel value xi in the positive class of the predicted edge map.

If yi equals to zero, it is a background pixel

The proposed method (6)

- β is applied as a hyper parameter to reduce the negative log-likelihood error.[23],[28]

$$\alpha = 1 + \frac{Y^-}{Y^+ + Y^-}$$

$$\beta = \frac{Y^+}{Y^+ + Y^-}$$

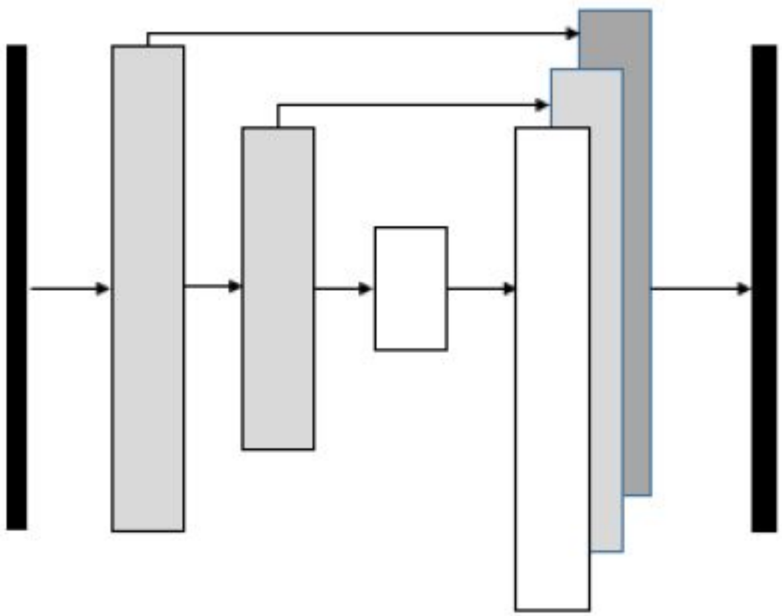
Objective loss function lambda which is the weight is set as 100.0

$$\mathcal{L}_{obj} = arg\,\min_G \max_D [\mathcal{L}_{cGAN} + \lambda \mathcal{L}1]$$

Network Architecture

- It follows UNET[25] framework to design the generator.
- Decoder VGG16
- Only one neuron in the output layer

- Network Architecture



Modified UNET

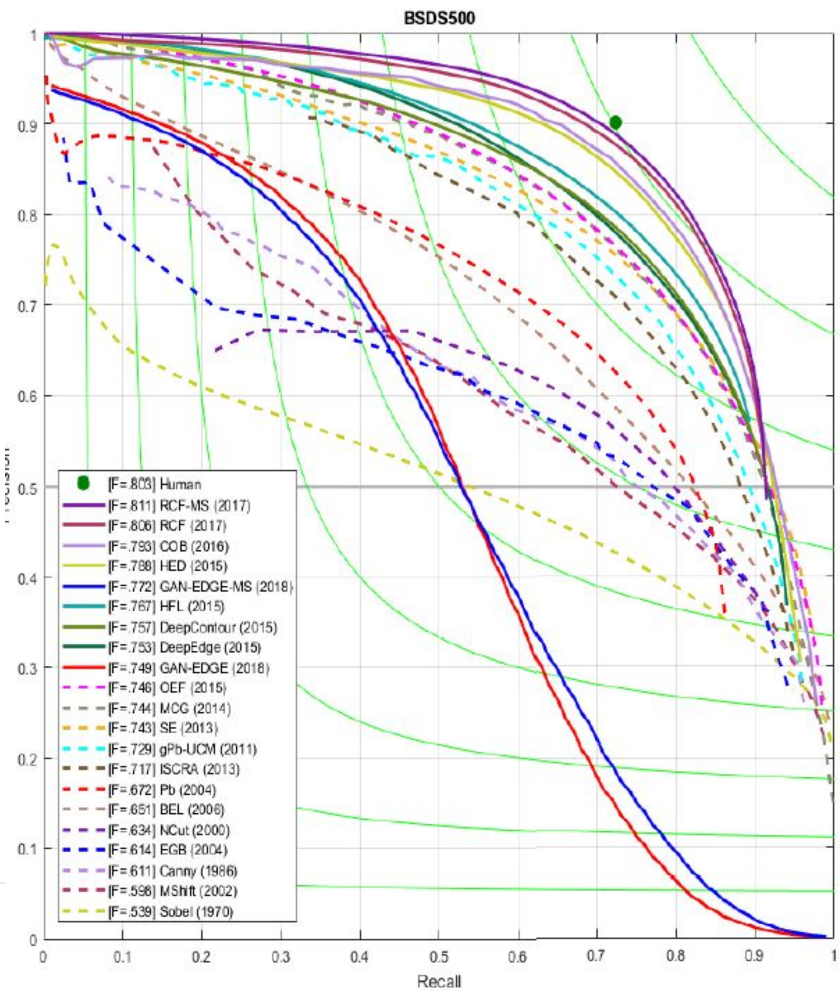
Proposed Method (9)

Experiment:

- BSD500 dataset was used to evaluate our generator[2]

- 500 examples in the dataset, in which 200 for training, 200 for testing, and 100 for validation
- Augmentation was used rotating and scaling the images to increase the amount of dataset to a total of 28800

Proposed Method (10)



Both of our single-scale and multiscale approaches achieved good performances that were comparable to the other latest methods.

Conclusion

- Achieved ODS and OIS scores on natural images that are comparable to the state-of-the-art methods.
- The model is computationally efficient.
- cGAN can produce an image which is close to the real one.
- Without using any pre-trained network parameters, the proposed method is still able to produce high quality edge images

References

- [2] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. Contour detection and hierarchical image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 33(5):898–916, May 2011. 1, 3, 4
- [14] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, editors, Advances in Neural Information Processing
- [17] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. CVPR, 2017. 1, 2, 3
- [23] Y. Liu, X. H. Ming-Ming Cheng, K. Wang, and X. Bai. Richer convolutional features for edge detection. In IEEE CVPR, 2017. 1,2, 3, 4
- [25] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. CoRR, abs/1505.04597, 2015. 1, 3
- [27] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-saascale image recognition. CoRR, abs/1409.1556, 2014. 1, 3