

CycleGAN

Dicky Ikbal Pratama (H071191060); Andi muhammad yusril (H071191018)

Departemen Matematika

Fakultas MIPA

Universitas Hasanuddin

CONTENTS

I	Introduction	1
II	Cycle GAN approach	1
II-A	Adversarial Loss	2
II-B	The Cycle consistency loss	2
II-C	Full Objective	2
III	Training details	2
IV	Evaluation	3
V	CycleGAN with Tensorflow	4
VI	Conclusions	4
References		5

CycleGAN

Abstract—Terjemahan gambar-ke-gambar adalah kelas masalah penglihatan dan grafik di mana tujuannya adalah untuk mempelajari pemetaan antara gambar input dan gambar output menggunakan satu set pelatihan pasangan gambar yang disejajarkan. Namun, untuk banyak tugas, data pelatihan berpasangan tidak akan tersedia. Kami menyajikan pendekatan untuk belajar menerjemahkan gambar dari domain sumber X ke domain target Y tanpa adanya contoh berpasangan. Tujuan kami adalah mempelajari pemetaan $G:X \rightarrow Y$ sedemikian rupa sehingga distribusi gambar dari $G(X)$ tidak dapat dibedakan dari distribusi Y menggunakan kerugian adversarial. Karena pemetaan ini sangat kurang dibatasi, kami menggabungkannya dengan pemetaan terbalik $F:Y \rightarrow X$ dan memperkenalkan kehilangan konsistensi siklus untuk mendorong $F(G(X)) \approx X$ (dan sebaliknya). Hasil kualitatif disajikan pada beberapa tugas di mana data pelatihan berpasangan tidak ada, termasuk transfer gaya koleksi, transfigurasi objek, transfer musim, peningkatan foto, dll. Perbandingan kuantitatif terhadap beberapa metode sebelumnya menunjukkan keunggulan pendekatan kami.

I. INTRODUCTION

CycleGAN adalah model terjemahan gambar-ke-gambar, sama seperti Pix2Pix. Tantangan utama yang dihadapi dalam model Pix2Pix adalah bahwa data yang dibutuhkan untuk pelatihan harus dipasangkan yaitu gambar domain sumber dan target harus berada di lokasi yang sama, dan jumlah gambar kedua domain juga harus sama.

The Cycle Generative Adversarial Network, atau CycleGAN, adalah sebuah pendekatan untuk melatih jaringan saraf konvolusi dalam untuk tugas-tugas terjemahan gambar-ke-gambar. Jaringan mempelajari pemetaan antara gambar input dan output menggunakan kumpulan data yang tidak berpasangan. Misalnya: Membangkitkan citra RGB dari SAR, citra multispektral dari RGB, memetakan rute dari citra satelit, dll.

Model ini merupakan perpanjangan dari arsitektur Pix2Pix yang melibatkan pelatihan simultan dari dua model generator dan dua model diskriminator. Selain fitur Pix2Pix, kita dapat menggunakan dataset yang tidak berpasangan dan juga kita dapat mengonversi gambar dalam arah sebaliknya (target ke sumber gambar) menggunakan model yang sama.

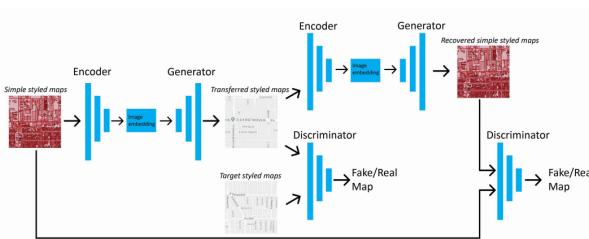


Fig. 1. Overview of CycleGAN architecture: Translating from satellite image to map routes domain

Pix2pix membuat asumsi bahwa data berpasangan tersedia untuk masalah terjemahan gambar yang sedang dipecahkan. Berbeda dengan itu, Cycle GAN, yang dibahas di bagian selanjutnya dari artikel ini, dibuat untuk mendukung bekerja dengan data yang tidak berpasangan.

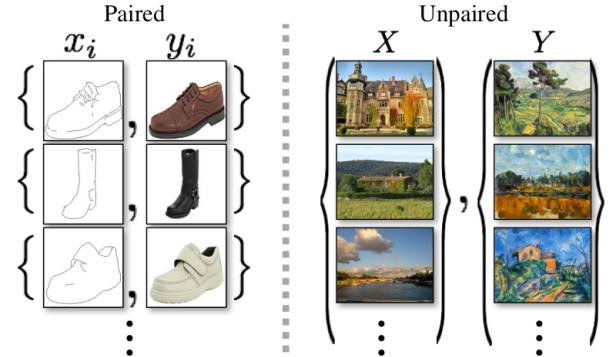


Fig. 2. Overview of CycleGAN architecture: Translating from satellite image to map routes domain

Tujuan dari model ini adalah untuk menghasilkan hasil yang sebanding dengan Pix2pix, tetapi tetap dapat belajar tanpa memerlukan data berpasangan. Inilah yang dilakukan CycleGAN. Ini dibangun di atas pekerjaan dari Pix2pix dengan memperkenalkan pembelajaran siklik dalam proses optimasi.

II. CYCLE GAN APPROACH

Di Pix2pix, model G dilatih untuk menerjemahkan gambar dari domain X ke domain Y . CycleGAN melakukan hal yang sama, tetapi selain itu juga melatih model F yang menerjemahkan gambar ke arah yang berlawanan - dari domain Y ke domain X .

Penerjemahan balik adalah sebuah konsep di mana setelah menerjemahkan dari A ke B , ada proses penerjemahan lain dari B kembali ke A untuk memeriksa seberapa dekat konten asli dibandingkan dengan yang melalui proses penerjemahan.

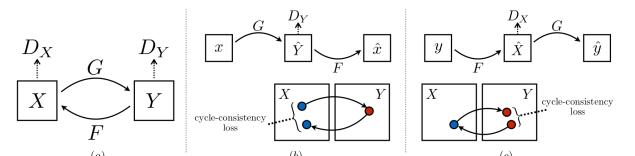


Fig. 3. CycleGAN

CycleGAN terdiri dari :

- 2 Generator - F dan G
- 2 Diskriminator - D_x dan D_y
- 2 Additional loss

- Forward cycle-consistency loss: $x \rightarrow G(x) \rightarrow F(G(x))x$
- Backward cycle-consistency loss: $y \rightarrow F(y) \rightarrow G(F(y))y$

Siklus GAN mempelajari fungsi pemetaan antara dua domain, X dan Y. Sampel pelatihan yang tersedia $x_{i=1}^N, x_i \in X$ and $y_{i=1}^M, y_i \in Y$. Generator G ditugaskan dengan pekerjaan mengubah sampel dari domain X ke Y, sedangkan generator F melakukan sebaliknya, mengubah sampel dari Y ke X. Diskriminasi Dx dan Dy melakukan klasifikasi biner di mana mereka mencoba untuk menentukan apakah sampel termasuk dalam domain X dan Y, masing-masing.

A. Adversarial Loss

Kerugian permusuhan diterapkan pada keduanya G dan F. Untuk pembangkit G dan diskriminasi Dy, dapat diformulasikan sebagai:

$$\begin{aligned} \mathcal{L}_{GAN}(G, D_Y, X, Y) = & \mathbb{E}_{y \sim p_{\text{data}}(y)} [\log D_Y(y)] \\ & + \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log(1 - D_Y(G(x)))] \end{aligned} \quad (1)$$

Fig. 4.

G mencoba untuk meminimalkan kerugian ini sementara diskriminasi Dy berusaha memaksimalkannya. Ini dapat dinyatakan sebagai

$$\min_G \max_{D_Y} \mathcal{L}_{GAN}(G, D_Y, X, Y)$$

Fig. 5.

Hal ini serupa untuk F dan Dx dan dapat diformulasikan sebagai:

$$\min_F \max_{D_X} \mathcal{L}_{GAN}(F, D_X, Y, X)$$

Fig. 6.

B. The Cycle consistency loss

Kehilangan konsistensi Cycle mencoba menangkap beberapa perbedaan sampel yang direkonstruksi dari sampel asli. Misalnya, seberapa jauh perbedaan gambar asli dibandingkan dengan gambar yang telah direkonstruksi oleh generator.

- cycle consistency loss:
 - Forward cycle consistency: $x \rightarrow G(x) \rightarrow F(G(x))x$
 - Backward cycle consistency: $y \rightarrow F(y) \rightarrow G(F(y))y$
- Kerugian tersebut kemudian dirumuskan sebagai:

$$\begin{aligned} \mathcal{L}_{\text{cyc}}(G, F) = & \mathbb{E}_{x \sim p_{\text{data}}(x)} [\|F(G(x)) - x\|_1] \\ & + \mathbb{E}_{y \sim p_{\text{data}}(y)} [\|G(F(y)) - y\|_1]. \end{aligned}$$

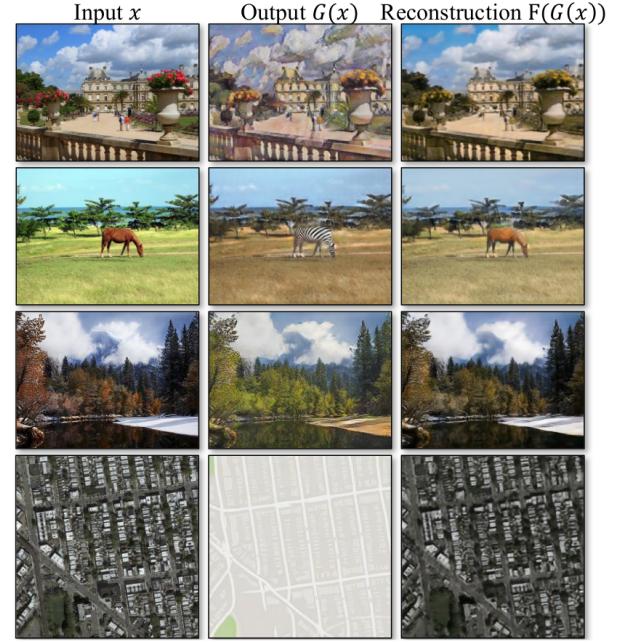


Fig. 7. The input images x, output images G(x) and the reconstructed images F(G(x)) from various experiments. From top to bottom: photoCezanne, horseszebras, winter→summer Yosemite, aerial photosGoogle maps.

C. Full Objective

Fungsi kerugian akhir dinyatakan sebagai:

$$\begin{aligned} \mathcal{L}(G, F, D_X, D_Y) = & \mathcal{L}_{GAN}(G, D_Y, X, Y) \\ & + \mathcal{L}_{GAN}(F, D_X, Y, X) \\ & + \lambda \mathcal{L}_{\text{cyc}}(G, F), \end{aligned}$$

Dan tujuan lengkapnya adalah:

$$G^*, F^* = \arg \min_{G, F} \max_{D_X, D_Y} \mathcal{L}(G, F, D_X, D_Y).$$

III. TRAINING DETAILS

Untuk semua percobaan, kami menetapkan $\lambda = 10$ dalam Persamaan 3. Kami menggunakan pemecah Adam dengan ukuran batch 1. Semua jaringan dilatih dari awal dengan tingkat pembelajaran 0,0002. Kami mempertahankan tingkat pembelajaran yang sama untuk 100 epoch pertama dan secara linier meluruhkan laju ke nol selama 100, epoch berikutnya. Discriminators menggunakan 70×70 Patch GAN.

IV. EVALUATION

CycleGAN tampaknya bekerja dengan baik dalam eksperimen yang diberikan dalam makalah. Dua gambar berikut menunjukkan bagaimana kinerja CycleGAN dibandingkan dengan metode sebelumnya.

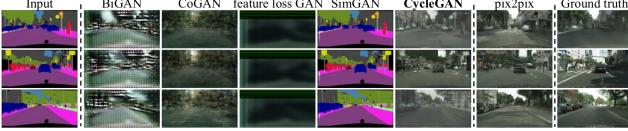


Fig. 8. Different methods for mapping labelsphotos trained on Cityscapes images. From left to right: input, BiGAN/ALI , CoGAN, feature loss + GAN, SimGAN , CycleGAN (ours), pix2pix trained on paired data, and ground truth.

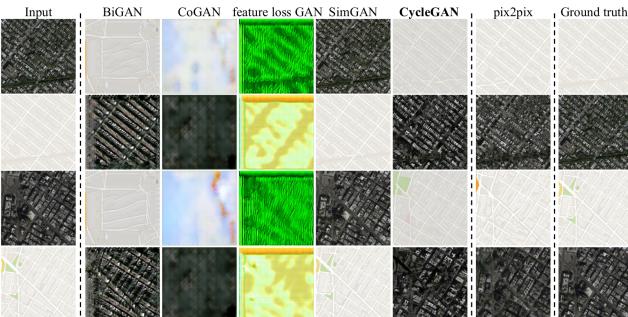


Figure 6: Different methods for mapping aerial photos \leftrightarrow maps on Google Maps. From left to right: input, BiGAN/ALI [7, 9] CoGAN [32], feature loss + GAN, SimGAN [46], CycleGAN (ours), pix2pix [22] trained on paired data, and ground truth.

Fig. 9. Different methods for mapping aerial photosmaps on Google Maps. From left to right: input, BiGAN/ALI ,CoGAN , feature loss + GAN, SimGAN, CycleGAN (ours), pix2pix trained on paired data, and ground truth.

Makalah ini melaporkan hasil beberapa percobaan:

- Masalah terjemahan gambar-ke-gambar yang tidak berpasangan
- Masalah terjemahan gambar-ke-gambar yang dipasangkan
- Dengan fokus pada pentingnya siklus dan kerugian yang merugikan
- Dengan fokus pada generalisasi pada kumpulan data yang berbeda.

Ada juga eksperimen yang dilakukan melalui Amazon Mechanical Turk (AMT) yang menunjukkan bahwa gambar yang dihasilkan oleh Cycle GAN berhasil membuat peserta berpikir bahwa itu nyata selama sekitar 25%.

Loss	Map \rightarrow Photo		Photo \rightarrow Map	
	% Turkers labeled real	% Turkers labeled real	% Turkers labeled real	% Turkers labeled real
CoGAN [32]	0.6% \pm 0.5%		0.9% \pm 0.5%	
BiGAN/ALI [9, 7]	2.1% \pm 1.0%		1.9% \pm 0.9%	
SimGAN [46]	0.7% \pm 0.5%		2.6% \pm 1.1%	
Feature loss + GAN	1.2% \pm 0.6%		0.3% \pm 0.2%	
CycleGAN (ours)	26.8% \pm 2.8%		23.2% \pm 3.4%	

Fig. 10. Table 1: AMT “real vs fake” test on mapsaerial photos at 256 \times 256 resolution.

Pada masalah terjemahan gambar-ke-gambar yang dipasangkan, Cycle GAN sebanding dengan Pix2pix:

Loss	Per-pixel acc.	Per-class acc.	Class IOU
CoGAN [32]	0.40	0.10	0.06
BiGAN/ALI [9, 7]	0.19	0.06	0.02
SimGAN [46]	0.20	0.10	0.04
Feature loss + GAN	0.06	0.04	0.01
CycleGAN (ours)	0.52	0.17	0.11
pix2pix [22]	0.71	0.25	0.18

Fig. 11. Table 2: FCN-scores for different methods, evaluated on Cityscapes labels \rightarrow photo.

Loss	Per-pixel acc.	Per-class acc.	Class IOU
CoGAN [32]	0.45	0.11	0.08
BiGAN/ALI [9, 7]	0.41	0.13	0.07
SimGAN [46]	0.47	0.11	0.07
Feature loss + GAN	0.50	0.10	0.06
CycleGAN (ours)	0.58	0.22	0.16
pix2pix [22]	0.85	0.40	0.32

Fig. 12. Table 3: Classification performance of photo \rightarrow labels for different methods on cityscapes

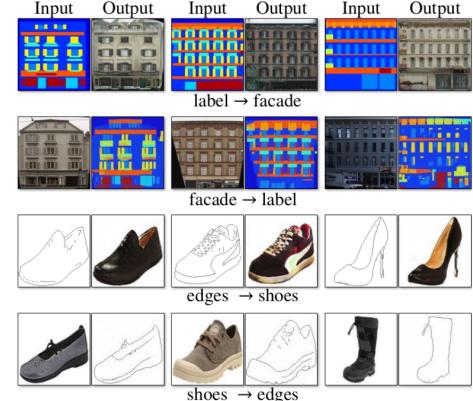
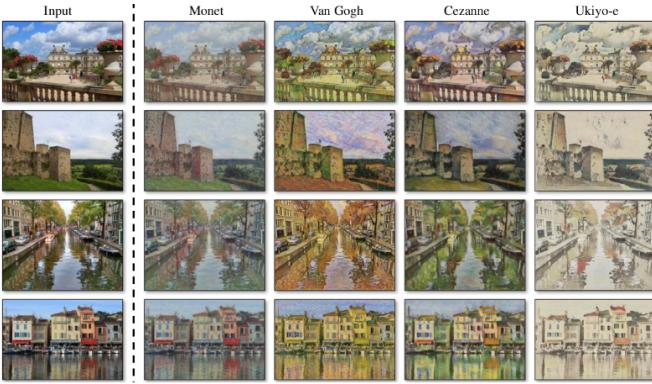


Fig. 13. Example results of CycleGAN on paired datasets used in “pix2pix” [22] such as architectural labelsphotos and edgesshoes.

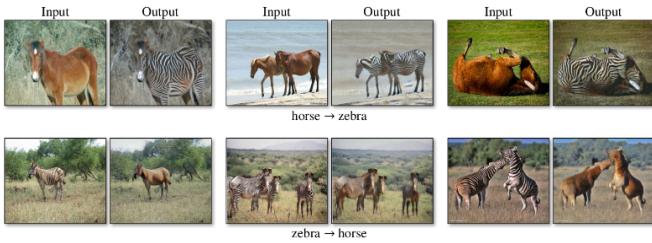
Siklus GAN dapat diterapkan pada masalah yang sama seperti Pix2pix. Beberapa gambar berikut menunjukkan hasil yang diperoleh.

- Perpindahan Gaya



Collection style transfer I: we transfer input images into the artistic styles of Monet, Van Gogh, Cezanne, and Ukiyo-e.

- Transfigurasi objek (kuda menjadi zebra)



- Kasus kegagalan

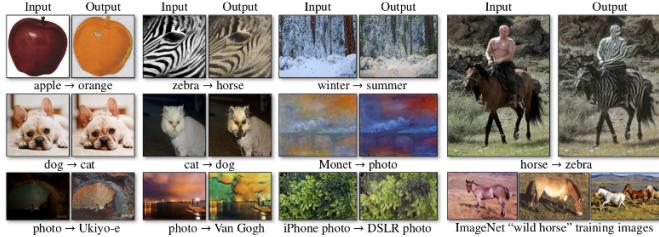


Fig. 14. Typical failure cases of our method. Left: in the task of dog→cat transfiguration, CycleGAN can only make minimal changes to the input. Right: CycleGAN also fails in this horse → zebra example as our model has not seen images of horseback riding during training. Please see our website for more comprehensive results.

V. CYCLEGAN WITH TENSORFLOW

<https://colab.research.google.com/drive/1nPN2Px6DkiEH2kQ1aDOfarcPhJH2H1xM>.

- Import Library yang diperlukan
- Input Pipeline

Tutorial ini melatih model untuk menerjemahkan dari gambar kuda, ke gambar zebra. Anda dapat menemukan dataset ini dan yang serupa di sini.

- Import and reuse the Pix2Pix models

Impor generator dan diskriminatur yang digunakan di Pix2Pix melalui paket tensorflow *examplesyangdiinstal*.

Loss functions

Mengenai Loss functions: Di CycleGAN, tidak ada data berpasangan untuk dilatih, oleh karena itu tidak ada jaminan bahwa input x dan pasangan target y bermakna selama pelatihan. Jadi untuk menegakkan bahwa jaringan mempelajari pemetaan yang benar, penulis mengusulkan hilangnya konsistensi siklus. Loss diskriminatur dan loss generator mirip dengan yang digunakan di pix2pix. Konsistensi siklus berarti hasilnya harus mendekati input asli. Misalnya, jika seseorang menerjemahkan kalimat dari bahasa Inggris ke bahasa Prancis, dan kemudian menerjemahkannya kembali dari bahasa Prancis ke bahasa Inggris, maka kalimat yang dihasilkan harus sama dengan kalimat aslinya.

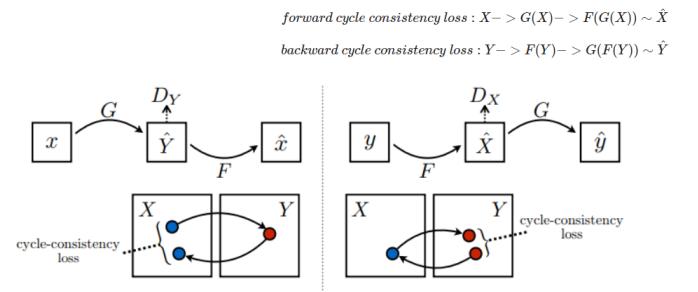


Fig. 15.

• Training Model

Pada bagian training modelnya, untuk EPOCH = 40 sedangkan di paper aslinya itu EPOCH = 200, yahh.. mungkin Predictions may be less accurate.



Fig. 16.

VI. CONCLUSIONS

Siklus GAN memiliki kemampuan untuk menghasilkan hasil yang sangat baik untuk data tidak berpasangan

dan berpasangan. Salah satu kemungkinan kasus penggunaan yang sangat relevan dengan kebutuhan praktis adalah pembelajaran model untuk menjembatani kesenjangan distribusi antara data sintetis dan nyata dalam situasi di mana data sintetis murah dan mudah didapat dibandingkan dengan data nyata. Menarik juga bagaimana mempelajari pemetaan terbalik dapat digunakan untuk memungkinkan model belajar pada kumpulan data yang tidak berpasangan.

Meskipun hasilnya sangat mengesankan, poin nyeri GAN reguler seperti mode collapse, pelatihan yang tidak stabil, dan sejumlah besar sumber daya komputasi yang diperlukan masih tetap ada.

REFERENCES

- [1] Jun-Yan Zhu,Taesung Park,Phillip Isola Alexei A. Efros, <https://arxiv.org/pdf/1703.10593.pdf>.
- [2] Jun-Yan Zhu,Taesung Park,Phillip Isola Alexei A. Efros, <https://arxiv.org/pdf/1611.07004.pdf>.
- [3] <https://github.com/tensorflow/docs/tree/master/site/en/tutorials/generative>
- [4] <https://engineering.nordeus.com/image-translation-pix2pix-cyclegan/#:~:text=In\%20Pix2pix\%2C\%20model\%20G\%20was,domain\%20Y\%20to\%20domain\%20X\%20.>