

Colorizarea pozelor alb-negru

Eric Petru Stavarache

Contents

0	Introducere	2
1	Munca Anterioara	2
2	Preliminarii	2
2.1	Notiuni utilizate	2
	Retele generative adversariale	2
2.2	Date de antrenare	3
3	Colorizare folosind GAN-uri	3
3.1	CycleGAN	3
4	Rezultate si posibile imbunatatiri	4
5	Bibliografie	5

Abstract

Aceasta lucrare trateaza colorizarea pozelor folosind retele adversariale, in particular arhitectura CycleGAN.

Voi compara rezultatele cu abordarile traditionale.

0 Introducere

Colorizarea pozelor alb-negru este o sarcina dificila pentru metodele traditionale de invatare.

Una dintre cele mai mari contributii la aceasta dificultate este faptul ca retele traditionale sunt specializate pe minimizarea unei pierderi de tipul $\log\text{-loss}$ sau $L2$. Considerand spatiul 3-dimensional al culorilor pixelilor RGB: $\{0, 1, \dots, 255\}^3$. Punctele care minimizeaza suma distantelor $L2$ tind sa fie distribuite in jurul culorii gri. Din aceasta cauza, pozele generate tind sa fie preponderent gri si lipsite de viata.

1 Munca Anterioara

Abordarea clasica de colorizare a pozelor este bazata pe retele convolutionale, iar rezultatele din 2016 ale lui Richard Zhang et al sunt impresionante [1][2]. Pentru a rezolva problema distributiei culorilor, ei modifica functia de pierdere pentru a penaliza culorile care nu apar frecvent in poze. i.e., desi culoarea gri minimizeaza $L2 - loss$, nu este intalnita foarte des in pozele adevarate, deci nu va mai fi culoare dominanta in pozele colorizate.

2 Preliminarii

2.1 Notiuni utilizate

Retele generative adversariale

Definiție 2.1. Retelele generative adversariale (eng. Generativ Adversarial Network) au fost inventate de catre Ian Goodfellow in 2014. Acestea sunt cea mai noua metoda de a rezolva problema esantionarii dintr-un spatiu al datelor unde abordarile traditionale au un bias catre medie. [5].

Aceste retele sunt compuse din doua parti: reseaua generatoare si reseaua discriminatoare. Reteaua generatoare G incearca sa genereze output din distributia O , bazandu-se fie pe input din distributia I . Vom nota distributia lui G cu G . Reteaua discriminatoare D primeste ca input i . D incearca sa distinga daca i a provenit din O , distributia adevarata a datelor, sau din G , distributia pe care o genereaza G .

Pierderea retelei G este data de cat de bine reuseasca sa "pacaleasca" reseaua D , in timp ce pierderea retelei D este cat de bine reuseste sa distinga intre datele adevarate si cele artificiale.

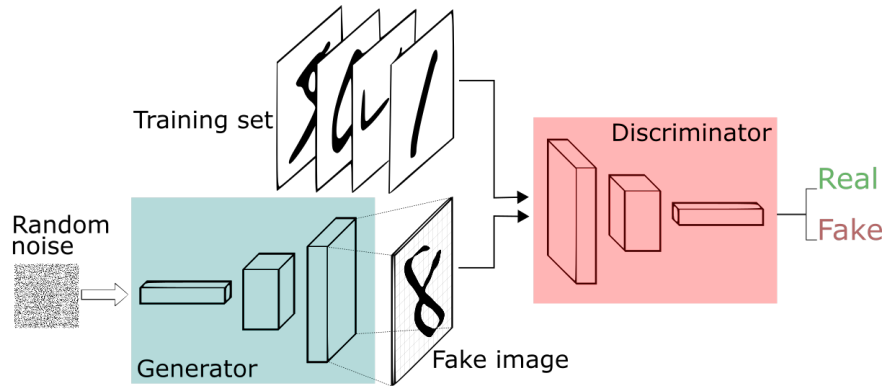


Figure 1: Retele Adversariale Generative

Ideea centrala a acestui tip de model se bazeaza pe teoria jocurilor: pe domeniul de min-max. G incearca sa maximizeze scorul, in timp ce D inceaca sa minimizeze scorul.

2.2 Date de antrenare

Pentru antrenarea retelei am combinat dataset-urile puse la dispozitie de autorii CycleGAN [6]. Aceste dataset-uri au fost menite pentru a transforma dintr-un domeniu specific in altul (de exemplu de a converti din cai in zebre). In abordarea mea, am creat reuniunea tuturor pozelor din dataset-uri, deoarece domeniul de poze color include toate celelalte poze. Pentru fiecare poza din reuniune, am generat varianta alb-negru.

3 Colorizare folosind GAN-uri

Aceasta lucrare incearca sa trateze problema colorizarii folosind Retele Adversariale Generative - GAN. Abordarea naiva ar fi formata din cele doua componente clasice: reseaua generatoare si cea discriminatoare. Reteaua generatoare G primeste o imagine alb-negru si genereaza una color. Reteaua discriminatoare D primeste o imagine color, si incearca sa distinga daca este o imagine naturala, sau daca a fost generata de catre D.

Problema cu aceasta abordare este ca reseaua generatoare descopera ca optim este sa genereze mereu aceeaasi poza colorata, fara a tine cont de input. Pentru a rezolva asta, am utilizat abordarea prezentata in *Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks*, sau CycleGAN [3][4].

3.1 CycleGAN

CycleGAN este un model care rezolva problema ca output-ul generatorului nu depinde de input. Ideea centrala este sa avem un generator G1 care ne trece din domeniul I in domeniul

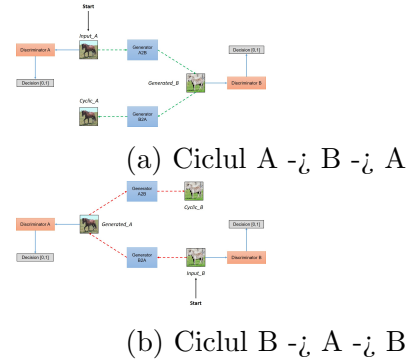


Figure 2: Modelul de functionare CycleGAN

G si un generator G_2 care ne trece din G in I . Functiile lor de pierdere vor fi corelate de faptul ca vrem ca transformarea inversa sa fie idempotentă. i.e., daca din alb-negru trecem o poza in color si apoi inapoi in alb-negru, vrem ca cele doua poze alb-negru sa fie cat de similare se poate.

4 Rezultate si posibile imbunatatiri

Pozele generate sunt destul de incetosate. Acesta fapt se genereaza cantitatii mici de date de antrenare folosite, iar o posibila imbunatatire ar fi utilizarea setului de antrenare imagenet



(a) Poza alb-negru



(b) Poza colorata

Figure 3: Colorizarea unei poze

5 Bibliografie

- [1] Richard Zang, Phillip Isola, Alexei A. Efros
Colorful Image Colorization, <https://arxiv.org/pdf/1603.08511.pdf>, ECCV, 2016.
- [2] Richard Zang *Colorful Image Colorization*, <http://richzhang.github.io/colorization/>
- [3] Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros, *Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks*, <https://arxiv.org/abs/1703.10593>
- [4] Jun-Yan Zhu, *CycleGAN Project Page*, <https://junyanz.github.io/CycleGAN/>
- [5] Ian J. Goodfellow et al
Generative Adversarial Networks, <https://arxiv.org/abs/1406.2661>
- [6] Taesung Park, *CycleGAN datasets*,
https://people.eecs.berkeley.edu/~taesung_park/CycleGAN/datasets/