**Aplicațiile GCC-GAN în îmbunătățirea realistă a semnelor de circulație sintetice**

**I. Schiță**

**0. Rezumat**

* Aici voi discuta de ce îmbunătățirea realistă a semnelor de circulație sintetice reprezintă un pas important către generarea de date cu semne de circulație „rare” și de ce acest lucru este esențial pentru mașinile autonome.

**1. Introducere**

* Aici voi vorbi mai întâi despre rețelele neuronale convoluționale profunde (Deep Convolutional Neural Networks) și despre impactul lor asupra viziunii computerizate, apoi despre GAN-uri [4] în general și de ce avem nevoie de un model mai bine ajustat în acest domeniu specific, în loc de modele generale. De asemenea, voi explica de ce generarea de date de calitate ajută algoritmii pentru mașinile autonome să performeze mai bine (pentru că distribuția datelor este dictată de modul în care sunt antrenate, iar un set de date suficient de divers va acoperi majoritatea cazurilor întâlnite de șoferi). În final, voi discuta pe scurt despre cauzele lipsei de date [3].

**2. Lucrări conexe**

* Aici voi discuta despre abordările actuale și limitele lor: o abordare simplă bazată pe GAN [4], GAN-uri condiționale [10] (pentru etichetarea tipului de semne de circulație pe care dorim să le generăm) și metodele propuse în [1]:
  + O **abordare Paste**, unde un semn de circulație vechi este retușat („inpainted”), iar imaginea generată este îmbunătățită cu un GAN.
  + O **abordare Cycle**, în care semnul vechi este retușat, iar rețeaua de retușare este antrenată simultan. Semnul de circulație este apoi plasat în imagine și comparat la pasul discriminator cu semnul original. Sunt aplicate anumite pierderi pentru a asigura similitudinea imaginilor.
  + O **abordare Style**, unde StyleGAN [8] este utilizat pentru a oferi stilul semnului de circulație sintetic, însă această metodă este dificil de antrenat și depinde în mare măsură de setul de date.

**3. Metoda propusă**

* În abordarea propusă, imaginile vor fi retușate ca în metodele anterioare, iar masca va fi aplicată pe imagine acolo unde dorim plasarea semnului sintetic. Un avantaj față de metodele precedente este că masca poate fi rotită sau pot fi aplicate diferite tipuri de transformări, iar modelul utilizat va ști, în majoritatea cazurilor, cum să le aplice corect. Aceasta, împreună cu semnul de circulație sintetic, va trece prin GAN-ul Geometric și Consistent din punct de vedere Coloristic [2]. Apoi se va explica de ce această metodă este superioară în păstrarea unor caracteristici pe care metodele vechi nu le luau în considerare.

**4. Rezultate**

* Aici voi prezenta câteva rezultate vizuale comparate cu cele obținute prin alte metode, subliniind avantajele majore ale acestui model. De asemenea, voi folosi un clasificator WideResNet pentru a încerca să prezic semnele de circulație rare generate de algoritmul propus și voi afișa rezultatele. Acestea vor fi comparate cu rezultatele prezentate în [1] și în unele din [13].

**5. Concluzii și considerații viitoare**

* Voi concluziona ceea ce am propus în lucrare și voi discuta despre direcțiile viitoare de cercetare.

**III. Plan de cercetare**

**Ipoteză:**  
Ipoteza acestei cercetări este că aplicarea GAN-ului geometric și consistent din punct de vedere coloristic (GCC-GAN) pentru generarea de îmbunătățiri realiste ale semnelor de circulație sintetice va depăși metodele existente, oferind date mai precise și mai diverse pentru antrenarea algoritmilor pentru mașini autonome. Această ipoteză se bazează pe convingerea că abilitatea GCC-GAN de a menține consistența geometrică și coloristică va genera semne de circulație sintetice mai convingătoare și mai informative.

**Metodologie:**

**Colectarea datelor:**

* Adunarea unui set de date diversificat cu imagini reale ale semnelor de circulație, asigurând reprezentarea semnelor rare și mai puțin frecvente.

**Arhitectura modelului:**

* Implementarea abordării propuse, integrând GCC-GAN [2] cu un mecanism de retușare a semnelor de circulație [14].
* Ajustarea arhitecturii pentru a gestiona variațiile în aspectul semnelor de circulație și scenariile de fundal.

**Antrenare:**

* Antrenarea modelului pe setul de date augmentat, optimizând pentru păstrarea caracteristicilor esențiale în semnele de circulație sintetice.
* Monitorizarea și ajustarea hiperparametrilor pentru a obține rezultate optime.

**Evaluare:**

* Evaluarea modelului folosind metode atât calitative, cât și cantitative.

**Comparație cu abordările existente:**

* Compararea abordării propuse cu metodele anterioare, în special cele menționate în [1].
* Evidențierea avantajelor, precum realismul îmbunătățit, diversitatea și aplicabilitatea pentru semnele de circulație rare.

**IV. Contribuții la stadiul actual al cercetării**

Această lucrare va contribui la stadiul actual al cercetării în domeniu prin:

**Avansarea realismului în date sintetice:**

* Extinderea limitelor realismului în generarea de date sintetice pentru semne de circulație, sporind autenticitatea conținutului generat. Acest progres este esențial pentru antrenarea algoritmilor care să recunoască și să răspundă unei game variate de scenarii reale de trafic.

**Depășirea limitărilor abordărilor anterioare:**

* Abordarea limitărilor metodelor anterioare, cum ar fi pierderea detaliilor geometrice sau coloristice, și oferirea unei soluții mai cuprinzătoare prin integrarea GCC-GAN. Această contribuție sprijină rafinarea tehnicilor de sintetizare a semnelor de circulație.

**II. Bibliografie**

[1] Anton Konushin, Boris Faizov, and Vlad Shakhuro. “Road images augmentation with synthetic traffic signs using neural networks”. In: preprint arXiv:2101.04927 (2021), pp. 1-15.

[2] Bor-Chun Chen and Andrew Kae. “Toward Realistic Image Compositing with Adversial Learning”. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (2019), IEEE pp. 8407-8416.

[3] Christian Ertler, Jerneja Mislej, Tobias Ollmann, Yubin Kuang, Lorenzo Porzi and Gerhard Neuhold. “The mapillary traffic sign dataset for detection and classification on a global scale”. In: Computer Vision–ECCV 2020 Proceedings, Part XXIII 16. Springer. 2020, pp. 68–84.

[4] Alec Radford, Luke Metz and Soumith Chintala. “Unsupervized Representation Learning with Deep Convolutional Generative Adversial Networks”. In: preprint arXiv:1511.06434 (2015), pp. 1–16.

[5] Raymond A. Yeh, Chen Chen, Teck Yian Lim, Alexander G. Schwing, Mark Hasegawa-Johnson, Minh N. Do. “Semantic Image Inpainting with Deep Generative Models”. In: CVPR. (2017), pp. 5485–5493.

[6] Sebastian Houben, Johannes Stallkamp, Jan Salmen, Marc Schlipsing, and Christian Igel. “Detection of traffic signs in real-world images: The german traffic sign detection benchmark”. In The 2013 international joint conference on neural networks (IJCNN), pages 1–8. IEEE, 2013.

[7] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14, pages 694–711. Springer, 2016.

[8] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 4401–4410, 2019.

[9] Yanghao Li, Naiyan Wang, Jiaying Liu, and Xiaodi Hou. Demystifying neural style transfer. arXiv preprint arXiv:1701.01036, 2017.

[10] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784, 2014.

[11] Eitan Richardson and Yair Weiss. On gans and gmms. Advances in Neural Information Processing Systems, 31, 2018.

[12] Vladislav Igorevich Shakhuro and AS Konouchine. Russian traffic sign images dataset. Computer optics, 40(2):294–300, 2016.

[13] Johannes Stallkamp, Marc Schlipsing, Jan Salmen, and Christian Igel. Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition. Neural networks, 32:323–332, 2012.

[14] Alexandru Telea. An image inpainting technique based on the fast marching method. Journal of graphics tools, 9(1):23–34, 2004.

[15] Leon A Gatys, Alexander S Ecker, and Matthias Bethge. Image style transfer using convolutional neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 2414–2423, 2016