

Redes Adversárias Geradoras Aplicadas no Incremento de Amostras de Imagens para Identificação de Lâmpadas de Iluminação Pública

Rodolpho M. B. Picolo, Francisco Fambrini, José H. Saito

Centro Universitário Campo Limpo Paulista (UNIFACCAMP)
13231-230 - Campo Limpo Paulista - SP - Brasil
rodolphopicolo@gmail.com, ffambrini@gmail.com, saito@cc.faccamp.br

Abstract. Generative adversarial networks (GANs) are a computational tool that enables the production of data sets with specific characteristics defined through examples, without necessity of a deep previous classification or labeling. They are, mainly, used in the context of images. In this work it is proposed the GANs use in the public illumination lights identification samples generation.

Resumo. Redes adversárias geradoras (GANs) são uma ferramenta computacional que viabiliza a produção de conjuntos de dados com características específicas e definidas através de exemplos, sem a necessidade de uma classificação ou rotulação prévia aprofundada. São utilizadas, principalmente, no contexto de imagens. Neste trabalho propõe-se a utilização de GANs para a geração de amostras de imagens utilizadas na identificação de lâmpadas de iluminação pública.

1. Introdução

Redes adversárias geradoras (GANs) são ferramentas computacionais de processamento de sinais utilizadas para a geração de réplicas de sinais similares aos de uma fonte original (Goodfellow et al. 2014). Em outras palavras, uma rede adversária geradora é uma ferramenta para a geração de uma probabilidade de dados similar a uma outra, apresentada como exemplo. Seu funcionamento é baseado em duas redes neurais treinadas simultaneamente, uma responsável por gerar sinais cada vez mais similares à fonte original, a rede geradora, e outra responsável por identificar se o sinal recebido é da fonte original ou obtida pela rede geradora, rede discriminatória. Recorrentemente, são apresentados, à rede discriminatória, sinais provindos da fonte original e sinais provindos da rede geradora. Os parâmetros da rede discriminatória são corrigidos de forma a melhorar a identificação de origem do sinal. Ao mesmo tempo, os parâmetros da rede geradora são corrigidos com o intuito de aumentar a similaridade do sinal gerado em relação ao original, baseados na resposta da rede discriminatória.

Segundo Creswell (2018), pode ser feita uma analogia considerando a rede discriminatória como um especialista em arte e a rede geradora como um falsificador. No processo de treinamento do conjunto, é apresentado recorrentemente, ao especialista, uma obra real e no próximo momento uma obra falsa. O especialista aprimora seu conhecimento com a comparação de sua avaliação com a informação de origem da obra, enquanto o falsificador aprimora suas obras falsas em função da opinião do especialista. Essa recorrência ocorre até que o especialista não tenha mais condições de saber a origem da obra, instante em que o falsificador tenha desenvolvido habilidades a ponto de conseguir produzir uma obra falsa que, submetida à avaliação do especialista, é tida como verdadeira. Inúmeras aplicações de redes GANs se encontram na literatura, dentre as quais a geração de imagens artificiais, com características específicas, por Kumar et al. (2020); tradução imagem-para-imagem de Pang et al. (2021); super-resolução de imagens para definição de texturas descritas por Ledig et al. (2016); edição semântica de imagens, aplicando filtros em áreas afetadas, desenvolvida por Yang et al. (2022); e transferência de estilo, que consiste em aplicar a textura de uma imagem fonte sobre a região de uma outra imagem, preservando o seu valor semântico, descrito por Bethge (2016).

Propõe-se neste trabalho a aplicação das redes GANs para a replicação de amostras de imagens no contexto de identificação de modelos de lâmpadas de iluminação pública. Essa identificação de modelos de lâmpadas são importantes para uma companhia de distribuição de energia elétrica, para calcular o consumo energético da iluminação pública e posterior cobrança de tarifas junto às prefeituras. Como as prefeituras são responsáveis pela troca das lâmpadas, que ocorrem frequentemente sem o conhecimento das concessionárias, essa identificação passa a ser importante. Existem vários tipos de lâmpadas de iluminação pública dentre as quais vapor de mercúrio, vapor de sódio, LED (Light Emitting Diode) e mistas com tubo de arco de vapor de mercúrio em série com um filamento incandescente de tungstênio (Li et al. 2009; Rodrigues et al. 2011). Uma técnica para a identificação dessas lâmpadas utiliza um conjunto de amostras de imagens filtradas de luzes, para a detecção do Índice de Reprodução de Cor (IRC). O presente trabalho resume o projeto, em andamento, que tem o objetivo de pesquisar a possibilidade de aumentar um conjunto reduzido de amostras dessas imagens, utilizando GANs, para serem utilizadas no aprendizado de máquinas para a identificação das lâmpadas.

Este texto é dividido nas seguintes seções além desta introdução. A Seção 2 refere-se a uma descrição de uma arquitetura de redes GANs e a Seção 3, descrição da aplicação proposta para as redes GANs. A Seção 4 refere-se à conclusão.

2. Arquitetura das redes GAN

De acordo com Goodfellow et al.(2014), as redes adversárias geradoras são formadas por duas redes neurais com funções distintas, a rede discriminatória e a rede geradora. As duas redes neurais são funções diferenciáveis e por conseguinte podem ser treinadas com o uso do algoritmo de retropropagação de erro para o aprimoramento dos resultados.

Considerando-se o diagrama da figura 1: \mathbf{x} : é o conjunto de dados para o qual se deseja gerar réplicas; \mathbf{z} : é um conjunto aleatório de distribuições de dados, que será transformado até adquirir as características de \mathbf{x} , chamado de espaço latente; $p_z(\mathbf{z})$: é uma função que define os valores iniciais dos elementos de \mathbf{z} , aleatoriamente; $D(\mathbf{x};\theta_d)$: um perceptron de múltiplas camadas que recebe \mathbf{x} como dados de entrada e θ_d como pesos dos neurônios, tem como saída, s , um escalar que representa a probabilidade da distribuição submetida ser de um elemento pertencente a \mathbf{x} ou a p_g ; $G(\mathbf{z};\theta_g)$ é um perceptron de múltiplas camadas que recebe \mathbf{z} como dados de entrada e θ_g como pesos dos neurônios, tem como saída, p_g , uma distribuição de probabilidades com as mesmas características estruturais de cada elemento de \mathbf{x} ; p_g é cada distribuição obtida como resultado de G ; s é escalar, entre 0 e 1, obtido como resultado de $D(\mathbf{x};\theta_d)$, sendo que quanto mais próximo de 0 maior a probabilidade do elemento submetido ao discriminador pertencer ao conjunto \mathbf{z} e quanto mais próximo de 1, maior a probabilidade do elemento submetido ao discriminador pertencer ao conjunto \mathbf{x} ; $M(\mathbf{z})$ é uma função que mapeia cada elemento de \mathbf{z} para um elemento de \mathbf{x} .

Figura 1 – Diagrama dos componentes de redes adversárias geradoras.

O processo de treinamento das redes GANs foi descrito por Goodfellow et al. (2014), como segue: (1) \mathbf{z} é submetido à p para a geração de valores iniciais, aleatoriamente; (2) a rede neural discriminatória D , recebe um elemento de \mathbf{x} , x_l , que tem como saída s_{xl} ; (3) é calculado o erro, e_{sxl} , de s_{xl} em função da origem do elemento, que, neste caso, é o conjunto \mathbf{x} , então, quanto mais próximo s_{xl} for de 1, menor o erro; (4) D é aprimorado com retropropagação de erro em função de e_{sxl} ; (5) a rede neural geradora G , recebe um elemento de \mathbf{z} , z_l , que gera a saída p_g ; (6) o elemento z_l é submetido à função M , que devolve o elemento de \mathbf{x} , x_2 , correspondente; (7) é calculado o erro, e_{pg} , de p_g em função da comparação com o elemento x_2 ; (8) G é aprimorado com retropropagação de erro em função de e_{pg} ; (9) O elemento z_l é modificado de forma a assumir os valores de p_g ; (10) p_g é submetido à D , produzindo uma saída s_{pg} ; (11) é calculado o erro, e_{spg} , de s_{pg} , que, como a origem da entrada é o conjunto \mathbf{z} , quanto mais próximo de 0, menor o erro; (12) D é aprimorado com retropropagação de erro em função do erro e_{spg} ; (13) a sequência de passos, de 2 a 12, é repetida até que os valores e_{sxl} e e_{spg} se aproximem o máximo possível de 0,5, obedecendo um número máximo de iterações ou tempo máximo de processamento pré-determinados.

3. Aplicação de GANs para o Incremento do Número de Amostras de Imagens

Propõe-se neste trabalho a aplicação das redes GANs na replicação de amostras de imagens no contexto de treinamento de máquinas para a identificação de modelos de lâmpadas de iluminação pública.

O Índice de Reprodução de Cor permite obter a correspondência entre a cor real da superfície de um objeto e sua aparência diante de uma fonte luminosa (Fambrini et al. 2022). Diferentes fontes de luz apresentam IRC diferentes e, portanto, com o treinamento de uma rede neural utilizando um conjunto suficiente de amostras de cada fonte de luz é possível identificar o tipo de cada lâmpada encontrada pelas concessionárias, desde que se obtenha os IRCs das imagens fotográficas de luzes correspondentes. A título de ilustração, é mostrado na figura 2, fotos de amostras de iluminação após aplicação de 24 filtros de cores, para as lâmpadas de vapor de mercúrio, LEDs, vapor metálico, luz mista, vapor de sódio, variando as potências conforme ilustrado, e uma foto de uma amostra sem nenhuma lâmpada.

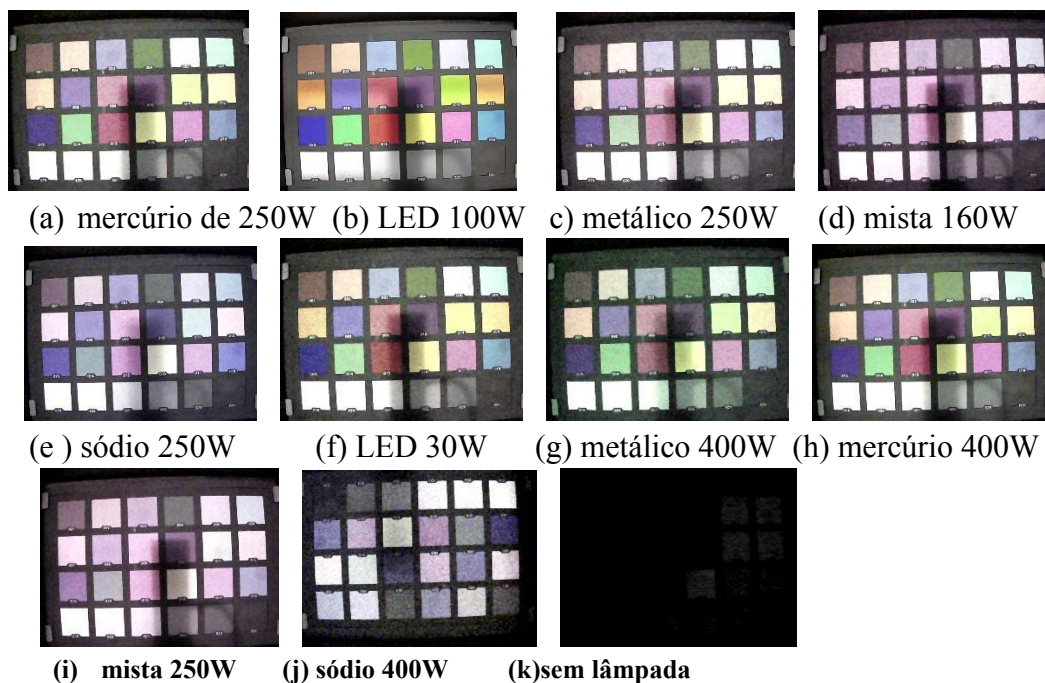


Figura 2. Exemplos de amostras de iluminação após a aplicação de filtros de cores, para as lâmpadas de vapor de mercúrio (a, h), LED (b, f), vapor metálico (c, g), mista (d, i), sódio (e, j) e sem lâmpada (k).

4. Conclusões

Para o trabalho proposto, as redes GANs devem reproduzir réplicas de cada uma das amostras de imagens como as da figura 2. Assim, com um número suficiente de amostras de imagens é possível o treinamento de redes neurais para o reconhecimento de tipos de lâmpadas. Essa abordagem de incremento do banco de amostras e seu uso no treinamento de redes neurais deve ser comparada com outras abordagens de identificação de objetos.

Nesta fase do projeto, uma rede GAN está sendo treinada sobre apenas uma amostra do conjunto de imagens para, através de transferência de conhecimento, com menor necessidade computacional, as imagens representativas de todo o conjunto serem geradas.

Referências

- Bethge M., Ecker A. S., Gatys L. A. (2016) “Image Style Transfer Using Convolutional Neural Networks”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Creswell A. et al. (2018) “Generative Adversarial Networks: An Overview”. In: *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 53-65. DOI: 10.1109/MSP.2017.2765202.
- Fambrini, F., Caetano D.C., Rangel A., Iano Y., Santos A.M., Rissi G.F. (2022) An Innovative Lighting Recognition System Based on Color Rendering Index and Computational Neural Networking. 31st IEEE International Symposium on Industrial Electronics (ISIE22). Anchorage, Alaska, USA, June 1-3, 2022.
- Goodfellow, I. et al. (2014) “Generative Adversarial Nets”. In: *Advances in Neural Information Processing Systems 27 (NIPS 2014)*.
- Kumar A., Bi L., Klim J., Feng D. (2020) “Chapter Five - Machine learning in medical imaging in Biomedical Information Technology (Second Edition)”. *Academic Press*, editor David Dagan Feng. DOI: 10.1016/B978-0-12-816034-3.00005-5.
- Ledig C. et al. (2016) “Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network”, DOI: 10.48550/arXiv.1609.0482.
- Li F., Chen D., Song X., Chen Y. (2009) Leds: A promising energy-saving light source for road lighting. In Power and Energy Engineering Conference, 2009. APPEEC 2009, Asia-Pacific, p.1-3, march, 2009.
- Pang Y., Lin J., Qin T., Chen Z. (2021). “Image-to-Image Translation: Methods and Applications”. In: *IEEE Transactions on Multimedia*. PP. 1-1. 10.1109/TMM.2021.3109419.
- Rodrigues C.R.B.S., Almeida P.S., Soares G.M., Jorge J.M., Pinto D.P., Braga H.A.C. (2011) An experimental comparison between different technologies arising for publica lighting: LED lumaires replacing high pressure sodium lamps. In 2011 IEEE International Symposium on Industrial Electronics (ISIE), p. 141-146, june, 2011.
- Yang S., Wang H., Long B., Zhan W. (2022) “Context-Consistent Semantic Image Editing with Style-Preserved Modulation”. DOI: 10.48550/arXiv.2207.06252.