

Redes Generativas Adversariais para Detecção de Câncer de Pele

William R. Massuda*, Sandra Avila.

Resumo

Esse projeto de pesquisa teve como objetivo o estudo de arquiteturas de redes generativas adversariais (GANs) para a geração de imagens sintéticas de lesões de câncer de pele, com foco principal na validação das diferentes métricas de avaliação para essas arquiteturas.

Palavras-chave:

Redes Generativas Adversariais, Aprendizado Profundo, Melanoma.

Introdução

Câncer de pele é o tipo de câncer mais comum no mundo. Entretanto, a taxa de sobrevivência cresce bastante caso haja uma detecção precoce (de 15% para cerca de 95%).

Nesse contexto, estudos têm sido feito na área de aprendizado de máquina com o objetivo de facilitar a identificação precoce das lesões. Porém, um dos maiores problemas existentes nessa área é a falta de imagens anotadas disponíveis. Assim, o surgimento das redes generativas adversariais (GANs) [1], arquiteturas com o objetivo principal de gerar imagens sintéticas, se mostraram como uma possível solução para amenizar esse problema. Compostas por uma rede conhecida como geradora e outra conhecida como discriminadora, as GANs são capazes de gerar imagens inéditas, aumentando assim a base de dados. Com o intuito de avaliar as imagens geradas por essas redes, esse projeto teve como objetivo o estudo de diferentes métricas de avaliação existentes, como forma de verificar a real utilidade das pontuações fornecidas e de encontrar uma possível forma de melhorar o treinamento utilizando-as.

Resultados e Discussão

Para a obtenção dos resultados foram utilizadas três diferentes métricas: Inception Score (IS) [2], Frechét Inception Distance (FID) [3] e Sliced Wasserstein Distance (SWD) [4]. As arquiteturas de GANs foram treinadas utilizando-se três diferentes estratégias, uma delas sendo o treinamento utilizando as imagens sem as labels, e as outras duas foi utilizado o treinamento com as labels com concatenação e com multiplicação. Utilizamos um conjunto de 10015 imagens da tarefa 3 da competição ISIC 2018 e outro sendo um agrupamento de vários conjuntos de dados totalizando 26847 imagens. Coletamos também os resultados das métricas aplicadas entre o conjunto real e ele mesmo para a obtenção de um limite inferior.

Tabela 1. Resultados da métrica IS.

| Base de dados | IS | |
|---------------------|------|-----------|
| | Real | Sintética |
| 10015/sem labels | 3.30 | 2.78 |
| 10015/multiplicação | 3.28 | 3.23 |
| 10015/concatenação | 3.28 | 3.49 |
| 26847/concatenação | 4.53 | 4.59 |

Tabela 2. Resultados da métrica FID

| Tubble 21 1 (Countados da motrica 1 15) | | | |
|---|------|-----------|--|
| Base de dados | FID | | |
| | Real | Sintética | |
| 10015/sem labels | 0 | 44.47 | |
| 10015/multiplicação | 0 | 68.77 | |
| 10015/concatenação | 0 | 78.99 | |
| 26847/concatenação | 0 | 58.86 | |

Tabela 3. Resultados da métrica SWD

| Tabola C: Necalitades da motifica CVID. | | | |
|---|------|-----------|--|
| Base de dados | SWD | | |
| | Real | Sintética | |
| 10015/sem labels | 2.02 | 17.15 | |
| 10015/multiplicação | 2.51 | 43.64 | |
| 10015/concatenação | 2.51 | 20.61 | |
| 26847/concatenação | 1.29 | 16.95 | |

A partir dos resultados, observamos uma grande discrepância, resultando em uma ausência de suporte base que ajudem a embasar uma conclusão com relação ao desempenho das GANs. Podemos notar também que a adição de label e o aumento no conjunto de imagens de treinamento nem sempre gera um resultado melhor nas pontuações, dificultando a identificação de fatores que contribuam para o aprendizado das redes.

Conclusões

Conseguimos observar a grande diferença nos valores e no comportamento das pontuações de cada métrica no decorrer do treinamento da GAN, evidenciando a grande dificuldade na confiabilidade e escolha de uma única métrica para a realização da avaliação das imagens. Com tantos trabalhos atualmente utilizando métricas como Inception Score para avaliar os resultados de diferentes arquiteturas de GANs, é preciso cautela, necessitando sempre de mais fatores que ajudem a fundamentar uma conclusão satisfatória.

Agradecimentos

Agradecemos ao PIBIC/CNPq, SAE/UNICAMP, FAEPEX (3125/17), FAPESP (2017/16246-0) e Google LARA 2018.



^[1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets", NIPS, 2014. [2] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved techniques for training gans," in NIPS, 2016.

^[3] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "Gans trained by a two time-scale update rule converge to a local nash equilibrium," in NIPS, 2017.

^[4] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of gans for improved quality, stability, and variation," ICLR, 2018.