

# UTILIZAÇÃO DO ALGORITMO DE GENERATIVE ADVERSARIAL NETWORK EM SUPER RESOLUÇÃO DE IMAGENS PARA O AUMENTO DE QUALIDADE DE MODELOS DE ELEVAÇÃO DIGITAIS BASEADOS EM DADOS SRTM

Leonardo Assumpção Moreira<sup>1</sup>, Lisandra Cristine Monteiro Blanco<sup>2</sup>, Nathan Damas Antonio<sup>3</sup>, Hideo Araki<sup>4</sup>

<sup>1</sup>UFPR, Curitiba-PR, leonardoamoreira@gmail.com; <sup>2</sup> UFPR, Curitiba-PR, lisandrac.blanco@gmail.com; <sup>3</sup> UFPR, Curitiba-PR, nathandamas@ufpr.br; <sup>4</sup> UFPR, Curitiba-PR, haraki@ufpr.br

## RESUMO

Aborda-se o aumento de resolução espacial de DEM pela utilização de algoritmos de aprendizado profundo acoplados a técnicas de Super-Resolução de Imagem Única em modelos digitais de elevação para obter versões de melhor qualidade espacial a partir de entradas de resolução mais baixa. O desenvolvimento de uma metodologia baseada em GAN permite melhorar a resolução espacial inicial de imagens de baixa resolução. A abordagem metodológica utiliza um conjunto de dados com modelos digitais de elevação SRTM (90 metros e 30 metros de resolução espacial), criados com o objetivo de permitir a realização do estudo. Verificou-se que ao aumentar o número de iterações o desempenho do modelo gerado foi melhorado e a qualidade da imagem gerada aumentou. Além disso, a análise visual da imagem gerada contra as de alta e baixa resolução mostrou uma grande semelhança entre as duas primeiras.

**Palavras-chave** — Modelos Digitais de Elevação; Generative Adversarial Network; Super Resolução de Imagens; Machine Learning; Deep Learning.

## ABSTRACT

We address the increase in spatial resolution of DEM by using deep learning algorithms coupled with Single Image Super-Resolution techniques in digital elevation models to obtain better spatial quality versions from lower resolution inputs. The development of a methodology based on GAN allows to improve the initial spatial resolution of low resolution images. The methodological approach uses a dataset with SRTM digital elevation models (90 meters and 30 meters of spatial resolution), created with the objective of allowing the study to be carried out. It was found that by increasing the number of iterations the performance of the generated model was improved and the quality of the generated image increased. Furthermore, the visual analysis of the generated image against the high and low resolution ones showed a great similarity between the first two.

**Key words** — Digital Elevation Model; Generative Adversarial Network; Image Super-Resolution; Machine Learning; Deep Learning.

## 1. INTRODUÇÃO

A SR (Super Resolução) de imagens é uma classe de técnicas que aprimoram a resolução de uma imagem, realizando o aumento da resolução espacial da mesma, aumentando o seu número de *pixels*, procurando diminuir a ocorrência de efeitos por conta da própria aquisição, como perda de detalhes, perda de qualidade, falta de nitidez, ocorrência de borrões, dentre outros. Ela baseia-se na ideia de que imagens de baixa resolução de uma determinada cena pode ser gerada uma imagem de alta resolução. Dessa forma, ela tenta reconstruir a imagem da cena original com alta resolução, dado um conjunto de imagens observadas em resolução mais baixa [1].

O mapeamento sistemático básico terrestre na escala 1:25.000 (mapa índice 1:25.000) possui um expressivo vazio cartográfico, onde a progressão do mapeamento de novas áreas ainda não mapeadas e a atualização temporal das já atualizadas segue de maneira lenta ou até mesmo nula. Desde 2008, por exemplo, tal escala não apresenta progressão significativa em sua situação. Existem diferentes fatores que influenciam esta questão, indo desde problemas políticos e institucionais, até questões de falta de recursos humanos, financeiros e técnicos.

Deseja-se desenvolver alternativas de geração de MDE com melhor resolução espacial por meio da utilização de técnica de SISR através de *Generative Adversarial Network* (GAN). O modelo que se pretende utilizar para a produção de um MDE com SR tem sido utilizado para obtenção de SR de fotografias coloridas e preto e branco, porém não foi encontrada nenhuma pesquisa no que se refere à aplicação do mesmo a dados espaciais altimétricos.

## 2. MATERIAL E MÉTODOS

A área de estudo da presente pesquisa foi o estado de Santa Catarina. A criação dos *datasets* se deu a partir do recorte de cada uma das imagens utilizadas dos MDEs de alta e baixa resolução espacial (MDEs SRTM de 90 metros e 30 metros de resolução espacial) em *tiles* de 156 por 156 *pixels* para as imagens do *ground truth*. Tais recortes foram organizados

em pares (ground truth e imagem de baixa resolução), sendo então, constituídos como os *datasets* de treino e de validação, onde foram procedidos os treinamentos dos modelos DEM-ESRGAN utilizados para a geração da imagem de super-resolução, assim como a realização das avaliações.

Para a solução do problema apresentado na presente tese, pretendeu-se utilizar as Redes Adversárias Generativas (GAN), que têm tido um grande destaque na área de inteligência artificial desde que foram introduzidas pelo artigo chamado Generative Adversarial Nets, publicado em 2014 por Ian J. Goodfellow e outros autores, pois a introdução do conceito de treinamento antagônico representou um progresso conceitual para as pesquisas realizadas com o uso do machine learning e do deep learning [2].

Uma das variações dos modelos GAN que pode ser utilizada para a obtenção de super-resolução de imagens é o Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN), que através de modificações realizadas no algoritmo originalmente proposto, alcançou consistentemente melhor qualidade visual com texturas mais realistas e naturais do que o GAN e o Super-Resolution Generative Adversarial Network (SRGAN).

Para a adaptação do algoritmo ESRGAN para o algoritmo DEM-ESRGAN, foram necessárias alterações nas entradas e nas saídas dos modelos. O ESRGAN tem como entrada e saída imagens .PNG, coloridas com 3 bandas (RGB). Para executar arquivos com informações altimétricas, foi necessário alterar as entradas e a saída para imagens .TIF, com apenas 1 banda.

Foi realizada a visualização lado a lado dos MDEs: o de baixa e o alta resolução, assim como o MDE gerado. O objetivo de tal avaliação é permitir a visualização e a comparação perceptiva da imagem gerada com a imagem original de baixa resolução e o alvo de alta resolução, verificando suas semelhanças e diferenças.

Para a realização da análise de curvas de nível geradas automaticamente, foi construído um algoritmo em Python para gerar curvas de nível nas imagens de alta e baixa resolução, assim como nas imagens geradas. Tal análise permite a comparação visual das curvas em diferentes altitudes, possibilitando verificar quais processamentos atingiram melhor conformação em comparação com a imagem de alta resolução.

A análise do Padrão de Exatidão Cartográfico Altimétrico teve por objetivo verificar o PEC dos produtos gerados. A avaliação da exatidão posicional altimétrica tem como base a análise dos resíduos entre as coordenadas de pontos retirados de um produto cartográfico e seus homólogos observados em outro produto.

Para a realização da análise do Padrão de Exatidão Cartográfico Altimétrico, foi calculada, inicialmente, a matriz de diferença dos pixels de cada par de imagens (ground truth e gerado). Calculou-se, em seguida, o Erro

Médio (EM) e o Erro Padrão (EP), para então verificar as escalas adequadas a serem utilizadas pelo produto avaliado.

### 3. RESULTADOS E DISCUSSÃO

#### 3.1 Análise dos processamentos DEM-ESRGAN P7, P8 e P9

Ao término dos processamentos DEM-ESRGAN, foram calculados os valores de suas referidas métricas PSNR. Em relação a este quesito, o primeiro processamento (P7) apresentou o PSNR de 39,103, o segundo processamento (P8) apresentou o PSNR de 41,317, enquanto o terceiro (P9) apresentou um PSNR de 42,049. Dessa forma, o processamento P6 apresentou maior medida quantitativa da qualidade de reconstrução no campo de compressão de imagem em relação aos demais processamentos analisados.

Durante os processamentos, verificou-se que ocorre uma variação oscilatória na progressão dos valores do PSNR, não ocorrendo dessa maneira um crescimento contínuo com o decorrer do tempo (Figura 1).



Figura 1 - Variação dos valores da métrica do PSNR dos processamentos DEM-ESRGAN executados (P7, P8 e P9) em função da progressão das épocas dos mesmos (onde k equivale a 1.000)

Ao se realizar a comparação do *loss l<sub>g</sub> perceptron* dos processamentos em questão, percebe-se que todos partem de um pico inicial e a partir daí passam a demonstrar uma tendência de aproximação do zero. Percebe-se que os processamentos com mais épocas se aproximaram mais do 0 do que os com menos épocas (Figura 2).



Figura 2 Variação dos valores apresentados pelas funções de perda *l<sub>g</sub> percep* de cada um dos processamentos DEM-ESRGAN executados (P7, P8 e P9) em função da progressão das épocas dos mesmos (onde k equivale a 1.000)

Com o intuito de verificar o melhor ajuste do número de

épocas aos processamentos DEM-ESRGAN realizados com o dataset 3, foram performados adicionalmente 2 (dois) novos processamentos de 75.000 e 125.000 épocas (valores circundantes aos de P6, responsável pela obtenção dos melhores desempenhos nas análises executadas). Dessa forma, verificou-se que o procedimento com 100.000 épocas se mostrou o mais adequado.

Conclui-se, parcialmente, que o 3º processamento do referido experimento apresentou melhor métrica final dentre as apresentadas durante o processamento em si. Já em relação à variação da função de perda  $L_{percep}$ , o 3º processamento também apresentou menor erro final.

### 3.2. Visualização dos MDEs de baixa resolução (esquerda), alta resolução (direita) e gerados pelo DEM-ESRGAN (centro)

A presente análise foi efetuada com os resultados dos 3 (três) processamentos realizados, resultando na visualização das imagens geradas pelo algoritmo DEM-ESRGAN, das imagens de baixa resolução e das imagens de alta resolução. Nessa visualização, colocou-se lado a lado, para comparação, as três imagens equivalentes, estando a de menor resolução à esquerda, a gerada ao centro e a de maior resolução à direita.

Dessa forma, a visualização permite a comparação perceptiva da imagem gerada com a imagem original de baixa resolução e o alvo de alta resolução. Pôde-se observar que todas as imagens geradas analisadas apresentaram grande semelhança com as respectivas imagens de alta resolução (Figura 3 e Figura 4).

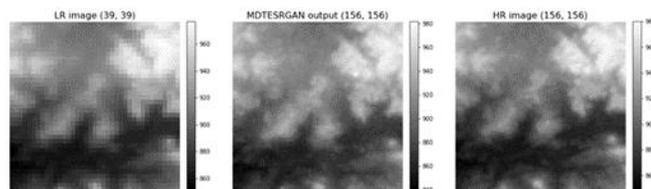


Figura 3. Exemplo da visualização da imagem gerada pelo algoritmo DEM-ESRGAN (centro), da imagem de alta resolução (direita) e da imagem de baixa resolução (esquerda) da imagem 2 do processamento P9

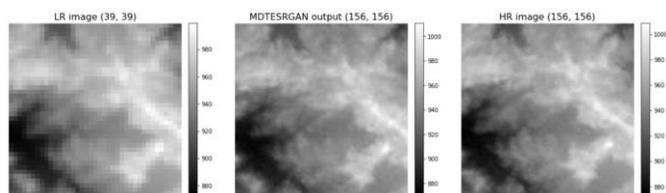


Figura 4. Exemplo da visualização da imagem gerada pelo algoritmo DEM-ESRGAN (centro), da imagem de alta resolução (direita) e da imagem de baixa resolução (esquerda) da imagem 3 do processamento P9

Como conclusão parcial, verificou-se que as imagens do processamento P9 apresentou maior proximidade com as

imagens de alta resolução do que as imagens dos demais processamentos do experimento.

### 3.3. Análise de curvas de nível geradas automaticamente

As seguintes figuras (Figura 5 e Figura 6) mostram a conformação das curvas de nível geradas automaticamente sem suavização, tanto para as imagens de alta resolução quanto para as imagens geradas pelos processamentos DEM-ESRGAN.

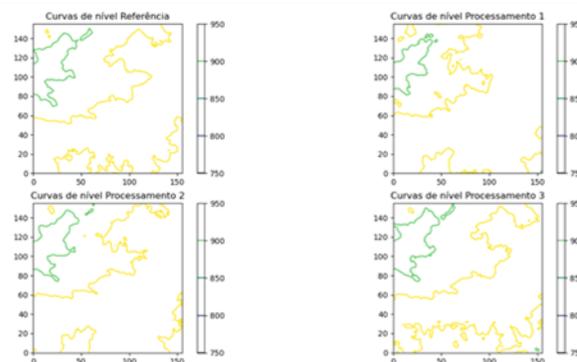


Figura 5. Visualização das curvas de nível da imagem de alta resolução e das imagens geradas pelo algoritmo DEM-ESRGAN da imagem 3 dos processamentos P7, P8 e P9

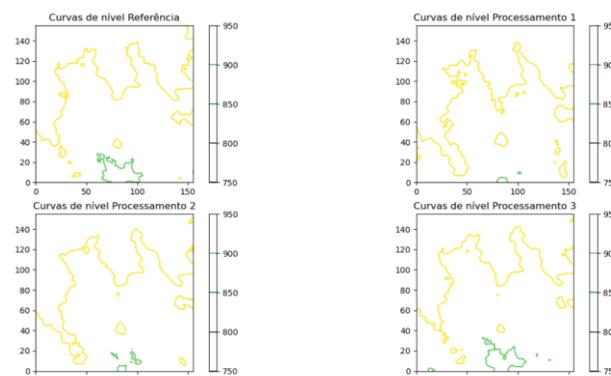


Figura 6. Visualização das curvas de nível da imagem de alta resolução e das imagens geradas pelo algoritmo DEM-ESRGAN da imagem 4 dos processamentos P7, P8 e P9

A partir desta análise, verificou-se que o processamento P9 apresentou melhor desempenho na comparação com as imagens de alta resolução, seguido pelo processamento P8.

### 3.4. Análise do Padrão de Exatidão Cartográfico Altimétrico

A avaliação da exatidão posicional altimétrica tem como base a análise dos resíduos entre as coordenadas de pontos retirados de um produto cartográfico e seus homólogos observados em outro produto. Para a realização da análise do Padrão de Exatidão Cartográfico Altimétrico foi calculada, inicialmente, a matriz de diferença dos pixels de

cada par de imagens (ground truth e gerada). Calculou-se em seguida o Erro Médio (EM) e o Erro Padrão (EP), para então verificar as escalas adequadas a serem utilizadas pelo produto avaliado (Tabela 1).

Process.	Img.	Res. Inicial	Res. final	EM	EP
P4	1	90 metros	30 metros	-1,326	8,947
	2			3,309	10,853
	3			-1,819	8,715
	4			1,373	10,806
P5	1	90 metros	30 metros	3,596	8,869
	2			6,694	9,982
	3			4,339	8,298
	4			4,210	9,229
P6	1	90 metros	30 metros	9,762	9,212
	2			13,029	10,326
	3			10,580	8,841
	4			10,624	9,483

**Tabela 1. Tabela dos valores calculados do Erro Médio e do Erro Padrão das imagens geradas nos respectivos processamentos**

Como conclusão parcial do experimento, verificou-se que o processamento P7 alcançou os menores erros médio e erros padrão dentre as imagens avaliadas. Comparando-se ainda o P8 e P9, verificou-se melhores valores calculados no segundo.

#### 4. CONCLUSÕES

Considerando as análises e comparações estatísticas apresentadas neste trabalho, verificou-se que aumentar o número de iterações é favorável ao desempenho do modelo gerado e à qualidade da imagem gerada. Os resultados obtidos apontam que a metodologia proposta é construtiva, cumprindo o fluxo de trabalho do SISR (*Single Image Super Resolution*) ao gerar uma imagem MDE de alta resolução a partir de uma única matriz MDE de baixa resolução. Além disso, uma análise de regressão indicou que uma equivalência total de pixels pode ser alcançada em um número finito de iterações.

O DEM-ESRGAN mostrou uma melhoria substancial nas métricas de qualidade em comparação com os métodos tradicionais de interpolação nas execuções de *benchmarks*. Portanto, o desenvolvimento de modelos digitais de elevação de super-resolução mostrou-se oportuno, dado os avanços tecnológicos nas áreas de inteligência artificial voltada para o processamento digital de imagens.

#### 5. REFERÊNCIAS

[1] A. C. C. Rézio, W. R. Schwartz, H. Pedrini. Super-resolução de imagens baseada em aprendizado utilizando descritores de

características. *X Congresso Brasileiro de Inteligência Computacional*, 2011.

[2] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, Warde-Farley, D. S. Ozair, A. E. Courville and Y. Bengio. Generative Adversarial Nets, *Veterinary Immunology and Immunopathology*, v. 155, pp. 1-10, 2013.