

Resumos das pesquisas de anterioridade

Eduardo Brasil Araujo

November 25, 2021

A Statistical Prediction Model Based on Sparse Representations for Single Image Super-Resolution

Um link para uma explicação do artigo no youtube

Link do issue

O artigo é fechado (sua distribuição não é gratuita), mas vi ser necessário demonstrar o mesmo pela tangente que o artigo escolhido para ser apresentado faz; por se tratar de um modelo de super-resolução em uma única imagem que foi considerado um dos melhores antes da publicação artigo sobre SRCNN (o artigo do repositório), e tendo em vista que ele tem o mesmo objetivo e mais ou menos a mesma metodologia. O presente consegue uma aproximação estatística com representação esparsa (que é uma representação dos dados com o mínimo de valores possíveis — uma melhor explicação) da imagem. Em suma, ele realiza predições sobre determinados atributos da imagem baseando-se em sua representação esparsa. Ele tem em sua implementação uma *Restricted Boltzman machine*, que consegue aprender distribuição probabilística com a entrada do programa.

Link para uma melhor aprofundamento sobre *sparse coding*

A Fully Progressive Approach to Single-Image Super-Resolution

Link do pdf

Link do Github

Link do issue

O paper consegue demonstrar uma super resolução em uma única imagem utilizando a técnica progressiva de geração. Em suma, o que a técnica significa é que ele vai progressivamente aumentando resolvendo a tarefa, de modo que

cada tarefa é algo fácil de ser resolvido, ou seja, para o nosso contexto, ela vai aumentando gradativamente a resolução até que consiga chegar no nosso resultado esperado. O artigo evidencia o uso de uma rede que consegue subtrair os serrilhados causados pelo escalonamento da imagem de baixa resolução para uma imagem maior (até 8 vezes maior), e depois que a imagem passa por essa rede, ela vai para uma rede adversária generativa que consegue implementar os detalhes menores, que dão mais resolução à imagem.

Learning Temporal Coherence via Self-Supervision for GAN-based Video Generation (TecoGAN)

[Link do pdf](#)

[Link do Github](#)

[Link do issue](#)

O presente artigo perfoma a façanha de aumentar a resolução (técnica de super resolução) com coerência temporal, ou seja, consegue performar aumento de resolução em imagens sequenciais, um vídeo, por assim dizer. Ele o realiza com a utilização de uma rede neural generativa, que diferente dos outros métodos que utilizam a mesma técnica de rede neural, consegue fazer a super resolução gerando muito poucos artefatos, ou então nenhum artefato, além de aprensetar maiores acréscimos de detalhes em comparação com outros métodos que realizam a mesma coisa. Além disso, ele consegue melhorar simulações física de partículas, em que outras técnicas de melhoramento dessas simulações precisavam ter um certo conhecimento de alguns parâmetros das partículas como velocidade, volume, esparsidade, porém, a presente técnica precisa somente das informações contidas na própria imagem do vídeo e já consegue performar melhor que a majoritariedade das outras técnicas anteriores.

RAISR: Rapid and Accurate Image Super Resolution

[Link do pdf](#)

[Link do video](#)

[Link do issue](#)

O paper utiliza técnicas de aprendizado de máquina (sem ser a rede neural) para conseguir realizar o upscaling (aumentar a imagem), para uma resolução maior, e, talvez, conseguir adicionar mais detalhes a elas, uma vez que elas se apresentam em um estado empobrecido devido ao escalonamento anterior que degradou parte da informação da imagem. Não é para a técnica ser confundida com *Image Inpainting*, que é uma técnica que substitui uma parte da imagem

baseada nos arredores da mesma. O algoritmo do artigo é altamente rápido e apresentou resultados superiores aqueles do artigo que estamos apresentando, além de ser melhor que os métodos manuais, para citar alguns: Nearest-Neighbor, SRCNN, A+.

Generative Adversarial Networks: An Overview

[Link do pdf](#)

[Link do issue](#)

O artigo explica o que é uma rede neural adversarial generativa, e suas principais aplicações; no final ela mostra vários outros *papers* que utilizam a técnica para as mais diversas aplicações. O que é uma rede neural generativa adversarial? Basicamente é uma rede que contém uma parte que gera (a rede generativa), e a parte da rede que diz se a determinada saída da rede generativa é real ou não, ou seja se foi gerada artificialmente ou não (essa parte da rede é chamada de rede discriminadora). Com esse empasse entre gerar uma saída e dizer se ela é real ou não, a rede vai aprendendo cada vez mais, tanto a como detectar uma imagem falsa, quanto a gerar uma imagem cada vez mais real.

Accelerating the Super-Resolution Convolutional Neural Network

[Link do pdf](#)

[Link do issue](#)

O presente artigo é um estudo sobre as redes neurais convolucionais que performam super resolução. Ele compara alguns modelos e tem como objetivo demonstrar quais as técnicas que devem ou não serem feitas para conseguir a melhor performance e eficiência. O que me chamou atenção é a comparação que ele fez com o artigo que nós escolhemos com uma arquitetura chamada FSRCNN, esta que é uma rede que possui uma rede mais profunda e filtros com o núcleo menor, ela consegue performar mais rápido tanto em tempo de execução quanto em tempo de treino por possuir menos parâmetros. Ela também tem a parte de escalonamento treinável, o que permite que consigamos uma melhor acurácia por termos mais filtros de treino significantes.

A tabela a seguir é um bom resumo dos resultados encontrados no artigo:

Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network

[Link do pdf](#)

Table 1. The transitions from SRCNN to FSRCNN.

	SRCNN-Ex	Transition state 1	Transition state 2	FSRCNN (56,12,4)
First part	Conv(9,64,1)	Conv(9,64,1)	Conv(9,64,1)	Conv(5,56,1)
Mid part	Conv(5,32,64)	Conv(5,32,64)	Conv(1,12,64)- 4Conv(3,12,12)- Conv(1,64,12)	Conv(1,12,56)- 4Conv(3,12,12)- Conv(1,56,12)
Last part	Conv(5,1,32)	DeConv(9,1,32)	DeConv(9,1,64)	DeConv(9,1,56)
Input size	S_{HR}	S_{LR}	S_{LR}	S_{LR}
Parameters	57184	58976	17088	12464
Speedup	1×	8.7×	30.1×	41.3×
PSNR (Set5)	32.83 dB	32.95 dB	33.01 dB	33.06 dB

Figure 1: Tabela de comparação

[Link do issue](#)

O artigo apresenta uma rede neural adversarial generativa com a intenção de realizar super resolução de uma única imagem de forma que o resultado seja o mais fotorrealista possível. Faz-se uso de uma rede residual com vários caminhos de pulo de conexão (típico de redes residuais). Ela faz uma comparação com o nosso artigo escolhido, dizendo que caso a rede neural fosse mais profunda, ela conseguiria um resultado melhor. Não só a rede generativa utiliza uma rede neural mais profunda, como também tem mais parâmetros de treinos em outras formas, como por exemplo no filtro de escalonamento. Além disso, ela utiliza uma função de perda (*loss function*) diferente a da MSE (Mean Squared Error), que é a *Perceptual loss function* que ela opera sobre o erro das partes características perceptíveis das imagens, ao invés do erro por pixel feita pela MSE. Os resultados são impressionantes e define um novo estado da arte com relação à super resolução de imagem.

Fast and Accurate Single Image Super-Resolution via Information Distillation Network

[Link do pdf](#)

[Link do issue](#)

O artigo tem como foco principal demonstrar uma super resolução de imagens de modo rápido e com grande acurácia com uma rede de distilação de informação. O presente consegue o feito com uma arquitetura um tanto diferenciada de rede, em que se detém um bloco de extração de características, e depois múltiplas camadas de distilação de informação para extrair informação residual, e depois um bloco de reconstrução que agrega as representações residuais das imagens de alta resolução e então uma operação de adição dessas informações com uma imagem maior que é a imagem de baixa resolução escalonada. O resultado demonstrou-se muito satisfatório, uma vez que consegue ser bem mais rápidos

que as técnicas anteriores, podendo se tornar até uma de uso em tempo real, além de demonstrar um melhor resultado.

Coupled Deep Autoencoder for Single Image Super-Resolution

[Link do pdf](#)

[Link do issue](#)

O paper tem como objetivo realizar a super resolução de imagens utilizando um par de autoencoders. Ele explica que vai utilizar dois autoencoders (que é uma rede neural que é melhor descrita por conseguir reconstruir uma entrada baseada na codificação da rede neural), é uma rede que consegue uma representação diferenciada de uma dada entrada. Com essas duas redes, ele consegue distilar a representação intrínseca das imagens de baixa resolução e de alta resolução, e então, executa o mapeamento entre essas duas partes da rede. É explicado também que esse processo do autoencoder de baixa resolução e de alta resolução são processo sem dependência, portanto são treinados em paralelo.

Video Super-Resolution With Convolutional Neural Networks

[Link do pdf](#)

[Link do issue](#)

Explica como aplicar super resolução em vídeo, e para isso demonstra três arquiteturas diferentes. Tem uma parte do início que explica como que uma representação esparsa da imagem pode ser vista como uma camada numa rede neural e cada átomo da representação esparsa pode ser visto como um filtro na rede neural. Primeiramente, pretreinam a rede neural convolucional para consegue atingir super resolução em únicos quadros. Como dito anteriormente, são apresentados três tipos de arquiteturas de redes neurais para super resolução do vídeo; todas elas utilizam a rede pretreinada de super resolução de imagem. Para a parte do vídeo, o que é feito é a concatenação dos quadros da vizinhança, ou seja, caso queiramos a super resolução do quadro t , teremos que por na rede o quadro $t-1$ e $t+1$, assim, teremos uma melhor coerência da imagem com relação ao tempo, uma coerência temporal.

Computed tomography super-resolution using convolutional neural networks

[Link do pdf](#)

[Link do issue](#)

Existe um dilema muito grande nas imagens de tomografia computadorizada, ter que balancear a quantidade de raios-x que o paciente leva, com a qualidade da imagem. O que ocorre é que para uma melhor imagem é necessário uma maior exposição aos raios-x, com isso, quando se necessita de maiores especificações das imagens do procedimento, põe mais em risco a vida do paciente. Uma das soluções encontradas é a utilização de técnicas de super resolução de imagens, e o artigo demonstra isso utilizando uma rede neural convolucional. O artigo propõe dois métodos, um que utiliza uma única imagem e consegue as informações residuais que são somadas à imagem escalonada para criar o output de super resolução, e outro que toma do mesmo processo mas que utilizando-se várias fatias de fotos, e este é o que apresentou melhor resultado. Ambas são bem promissoras por terem apresentado resultados bem satisfatórios.

Video Satellite Imagery Super Resolution via Convolutional Neural Networks

[Link do pdf](#)

[Link do issue](#)

Tratando-se de imagens de satélite, o que mais ocorre com essas imagens é a degradação e compressão da imagem devido à natureza dos processos de aquisição. Para melhor resolver isso, e conseguir fotos com mais informações espaciais de maiores resoluções com o uso de um método de super resolução utilizando uma rede neural convolucional. A arquitetura da rede neural utilizada é uma rede convolucional de 20 camadas VDSR (Very Deep Super Resolution — Super Resolução Muito Profunda) para super resolução, para conseguir extrair os detalhes mais característicos, e utiliza as métricas MSE e PSNR. Os resultados são bem satisfatórios, sendo superiores tanto às técnicas de interpolação bicúbica, SRCNN (o nosso artigo apresentado) e VDSR.

Super-Resolution Using Convolutional Neural Networks Without Any Checkerboard Artifacts

[Link do pdf](#)

[Link do issue](#)

A intenção do artigo é produzir super resolução sem gerar artefatos de xadrez nas imagens. Existem dois tipos de redes neurais para super resolução, um deles utiliza imagens escalonadas com interpolação, portanto não produzem os artefatos que estamos a estudar, e outro que utilizam de camadas de deconvolução para aumentar as imagens. São dispostos dois tipos de algoritmos em que se pode ser aplicado os métodos de remoção de artefatos xadrez, são eles: camadas

de deconvolução e convolução sub-pixel. O artigo propõe três tipos de uso para o kernel de interpolação $H0$, que é treinar sem utilizar o $H0$, treinar utilizando o $H0$ e treinar utilizando o $H0$ dentro das camadas de escalonamento. Melhores definições para o kernel $H0$ estão presentes no artigo.