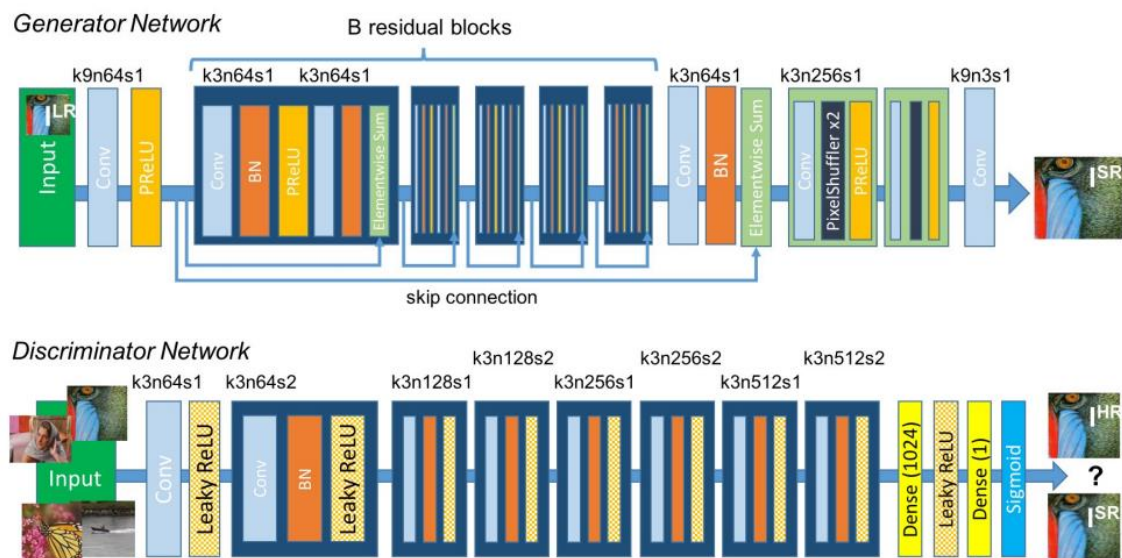


Aluno: Rafael Monteiro Laranjeira

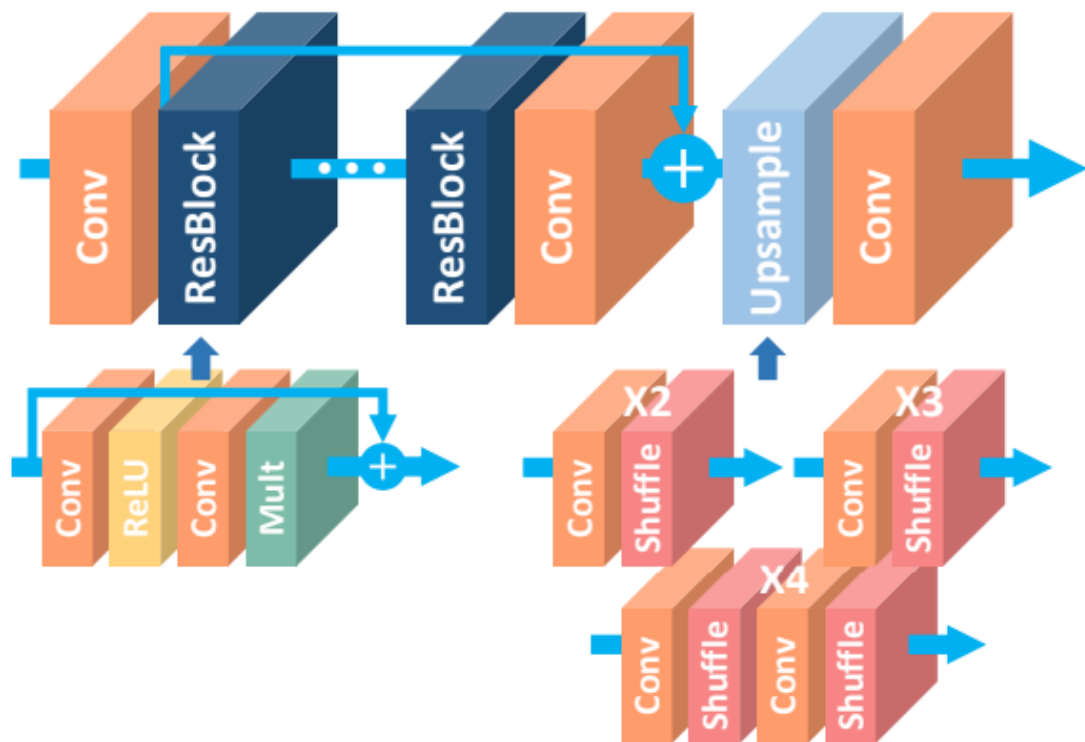
Conferences

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



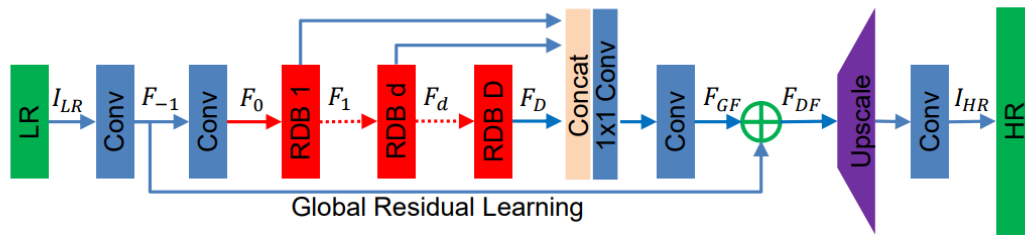
O artigo propõe uma rede adversária generativa para super resolução de imagens. É o primeiro framework capaz de inferir imagens foto realistas com um fator de aumento de 4 vezes. Foi proposta uma função de perda perceptiva composta de uma perda adversarial e uma perda de conteúdo. A perda adversarial nos leva em direção à solução através de uma rede discriminadora treinada para diferenciar as imagens com super resolução e as imagens foto realistas originais. Além disso, a perda de conteúdo é motivada pela similaridade perceptiva no lugar da similaridade no espaço de pixels. Essa rede profunda residual foi capaz de restaurar texturas foto realistas de imagens com resoluções extremamente baixas.

Enhanced Deep Residual Networks for Single Image Super-Resolution



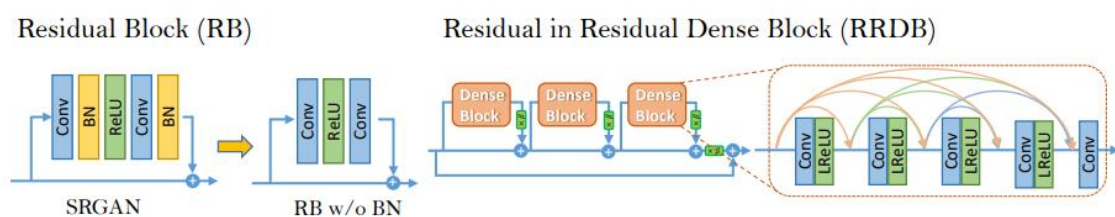
Esse paper propõe uma rede aprimorada de super resolução: *Enhanced Deep Super-Resolution network* (EDSR). Essa rede possui um desempenho superior aos métodos de super resolução mais utilizados pelo estado da arte. Essa melhora se deu pela otimização que foi feita removendo módulos desnecessários de redes residuais convencionais. As redes residuais tem tido resultados muito promissores no período recente e esse artigo conseguiu aprimorar mais ainda os resultados obtidos por redes residuais. Além disso, a performance também melhorou devido à expansão do tamanho do modelo e estabilização da metodologia para treinar a rede. Um sistema de múltiplas escalas de super resolução também é proposto, sendo capaz de reconstruir imagens de alta resolução usando diversas escalas em um mesmo modelo.

Residual Dense Network for Image Super-Resolution



Esse paper introduz uma nova rede residual para tentar resolver uma falha comum em diversos artigos que fazem uso das redes neurais convolucionais para super resolução de imagens: A maioria dos modelos propostos não faz uso completo das features hierárquicas decorrentes da imagem original de baixa resolução, o que acaba limitando o desempenho destas. Esse artigo propõe um bloco residual denso pra extrair as features locais mais abundantes e ainda permitir a conexão direta do estado anterior de outros blocos residuais densos com todas as camadas do bloco atual, criando um mecanismo memória contígua. A fusão das features no bloco residual denso é usada para aprender de forma adaptativa features mais eficientes do bloco anterior e do atual. No fim, é feita a fusão das features globais e aprende-se de maneira adaptativa as features hierárquicas.

ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks



Esse artigo introduz uma melhoria à rede adversária generativa de super resolução (SRGAN). Um grande problema da SRGAN é que ela consegue gerar texturas realistas em altas resoluções, porém esse resultado vem acompanhado de artefatos não desejados. A ESRGAN é uma rede que busca melhorar a arquitetura, a perda adversarial e a perda perceptiva da SRGAN. Para isso, a ESRGAN introduz um bloco RRDB (Residual-in-Residual Dense Block) sem normalização em batch. A ideia é advinda de uma rede adversaria generativa relativista, onde o discriminador prevê a fidelidade

relativa no lugar da fidelidade absoluta. A perda perceptiva é melhorada usando as features obtidas antes da ativação.

Image Super-Resolution Using Very Deep Residual Channel Attention Networks

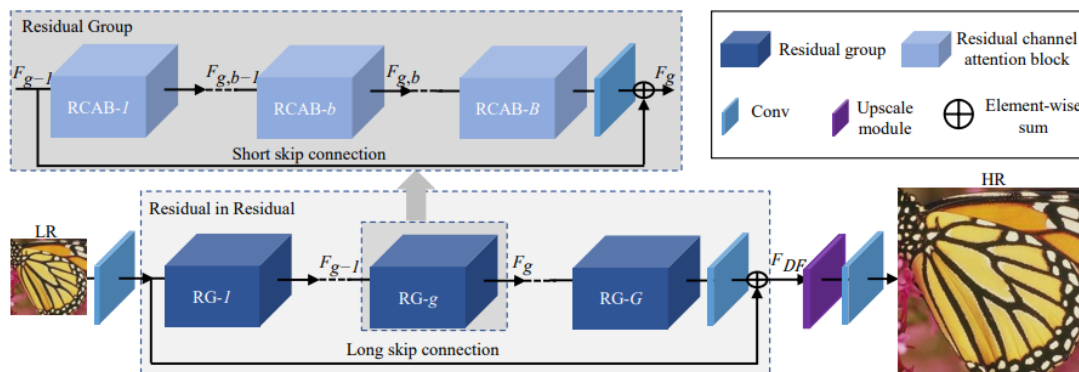
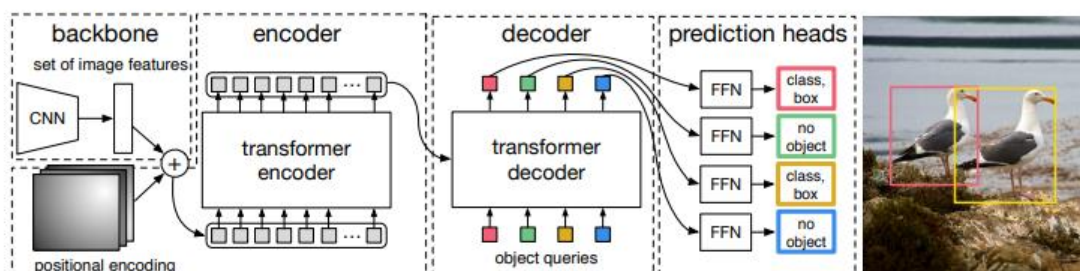


Fig. 2. Network architecture of our residual channel attention network (RCAN)

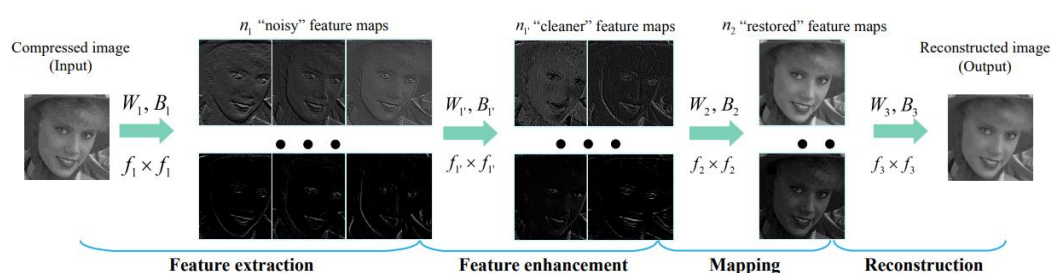
Esse artigo partiu do fato de que as redes neurais convolucionais profundas são mais difíceis de treinar para propor uma rede residual profunda de atenção de canal. É proposta uma estrutura resíduo dentro do resíduo (RIR) pra formar uma rede bastante profunda composta de vários grupos residuais com longos saltos de conexões. Essa estrutura permite que as informações de baixa frequência, que são mais abundantes, pulem diversas conexões e permitam que a rede foque nos detalhes (informações de alta frequência). Além disso, o artigo propõe um mecanismo de atenção que considera as interdependências entre os canais e permite o escalonamento das features num nível de canal.

End-to-End Object Detection with Transformers



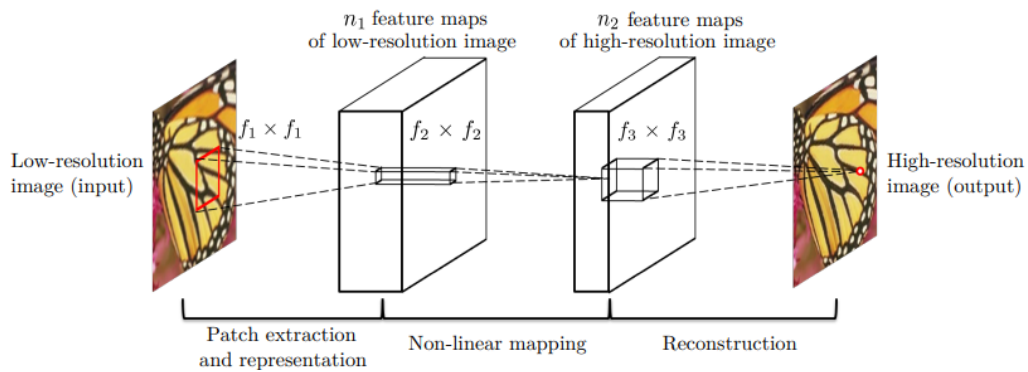
O artigo propõe um novo método que vê a detecção de objetos como um problema de predição de conjuntos. Os principais pontos desse método são: definir um cálculo de perda global para o conjunto, forçando previsões únicas através de um emparelhamento de grafos bipartidos, além de uma arquitetura de transformer composta por decodificador e codificador. Dado um pequeno conjunto de objetos, o método pondera sobre as relações entre os objetos e do contexto global da imagem pra retornar um conjunto de previsões em paralelo. O modelo tem a vantagem de ser simples e não requer biblioteca especial para funcionar. Foi possível atingir resultados próximos aos do estado da arte e até melhores em alguns casos.

Compression Artifacts Reduction by a Deep Convolutional Network



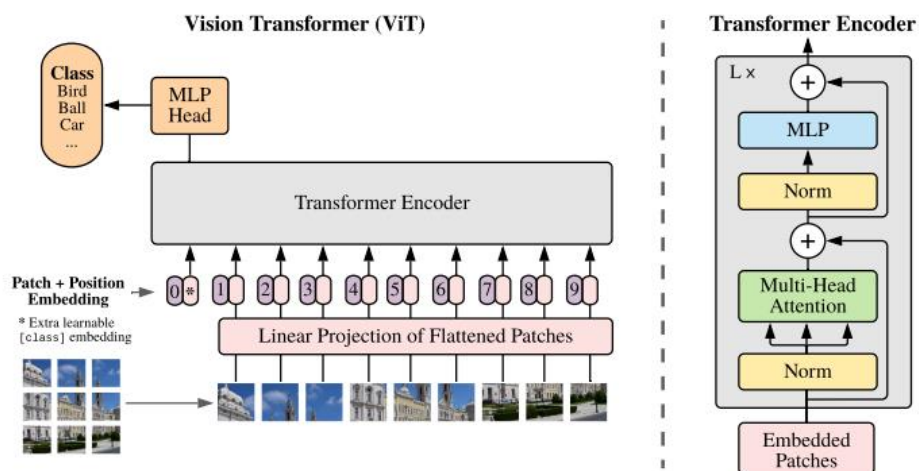
Inspirado pelas redes neurais convolucionais profundas, esse artigo formulou uma rede compacta e eficiente para reduzir os efeitos de diferentes artefatos que foram gerados na compressão de imagens no formato JPEG. Em contrapartida, os algoritmos existentes focam majoritariamente em remover artefatos bloqueantes e produzir uma saída borrada ou restaurar as imagens nítidas que vieram com efeitos indesejáveis. O artigo mostra que um modelo mais profundo pode ser treinado com as features obtidas por uma rede mais rasa. Ademais, é feita uma análise diversas configurações de transferência de aprendizagem em problemas de visão computacional de baixo nível.

Image Super-Resolution Using Deep Convolutional Networks



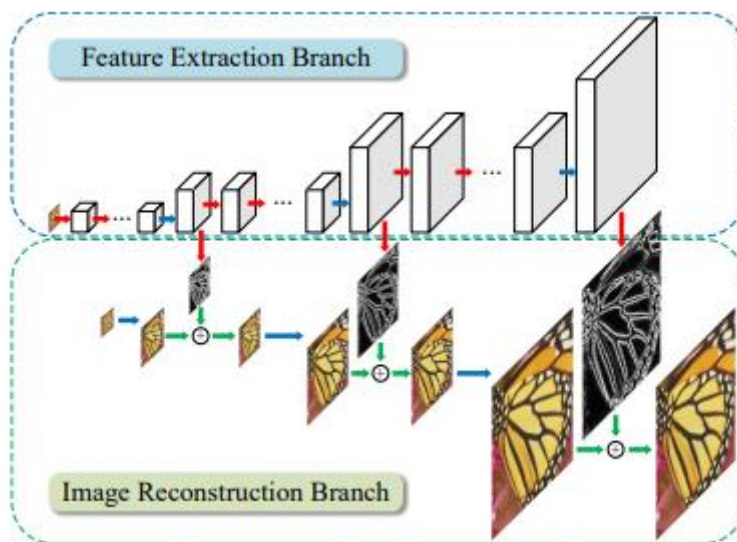
O artigo propõe um método que usa aprendizado profundo para a super resolução de imagens. O método tenta mapear diretamente as imagens de baixa resolução com as imagens de alta resolução. O mapeamento é feito usando redes neurais convolucionais profundas, onde a entrada é uma imagem de baixa resolução e a saída é uma imagem de alta resolução. Diferentemente dos métodos tradicionais, o método proposto por esse artigo otimiza todas as camadas de uma vez só. Além disso, a estrutura é leve e consegue obter resultados tão bons quanto os do estado da arte de forma mais rápida. Uma análise é feita para descobrir quais parâmetros conseguem conciliar melhor o dilema entre velocidade e performance, sem perder muita qualidade.

AN IMAGE IS WORTH 16X16 WORDS: TRANSFORMERS FOR IMAGE RECOGNITION AT SCALE



Esse artigo mostra que a dependência das redes neurais convolucionais para problemas de visão computacional não é necessária. O que geralmente vinha acontecendo na prática é aplicar o mecanismo de atenção junto das redes neurais convolucionais ou substituir algum componente dessas redes por esse mecanismo. No artigo, o método usou um transformer puro para classificação de imagens e teve resultados excelentes. Ele foi treinado com uma grande quantidade de dados e comparado com diversas redes neurais convolucionais do estado da arte, e além dos bons resultados, ele precisou de bem menos recursos computacionais para treinar.

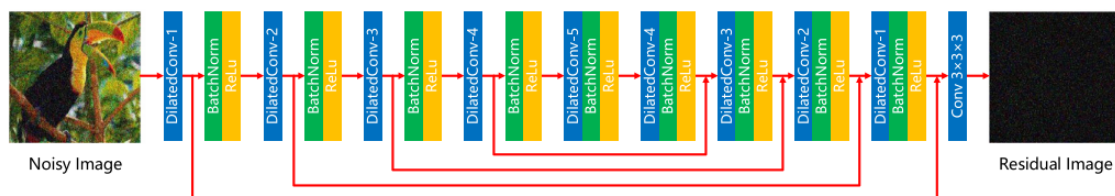
Deep Laplacian Pyramid Networks for Fast and Accurate Super-Resolution



Esse paper introduz a *Laplacian Pyramid Super-Resolution Network* (LapSRN), uma rede que reconstrói progressivamente os resíduos das imagens de alta resolução a partir das imagens de baixa resolução. Em cada nível da pirâmide laplaciana, o modelo pega o mapeamento das features de resolução mais espessa como entrada, prevê o resíduo de alta frequência e usa as convoluções transpostas até o nível mais fino. Esse método não requer nenhum tipo de interpolação bicúbica e por isso precisa de menos recursos computacionais.

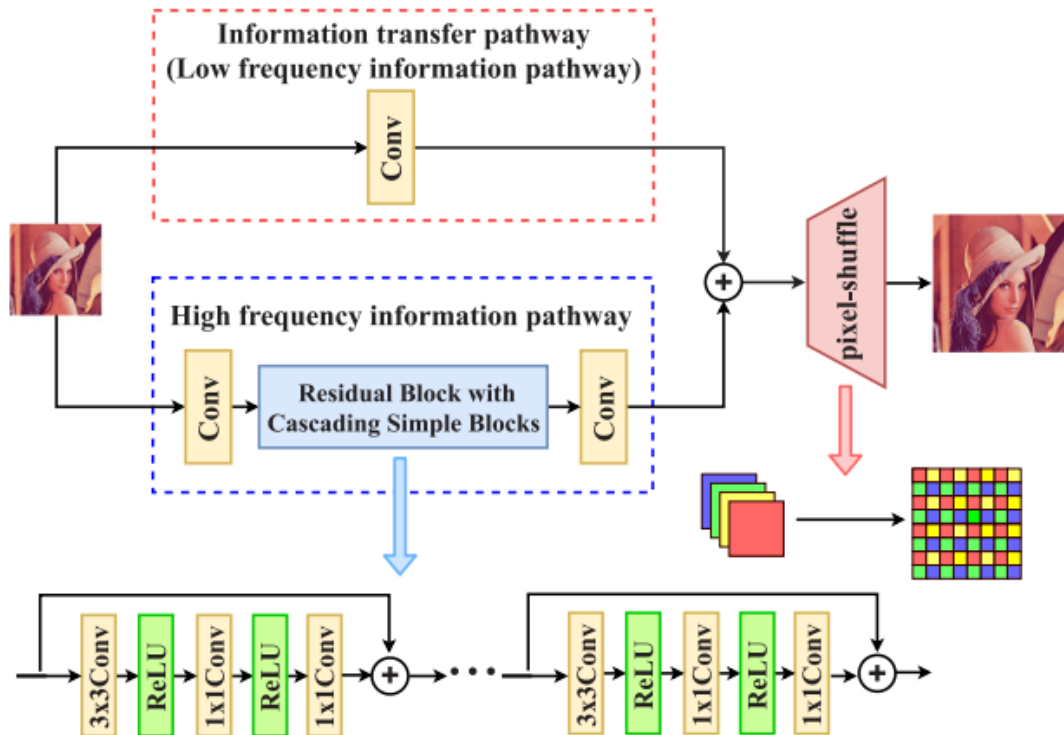
Journals

Dilated residual networks with symmetric skip connection for image denoising



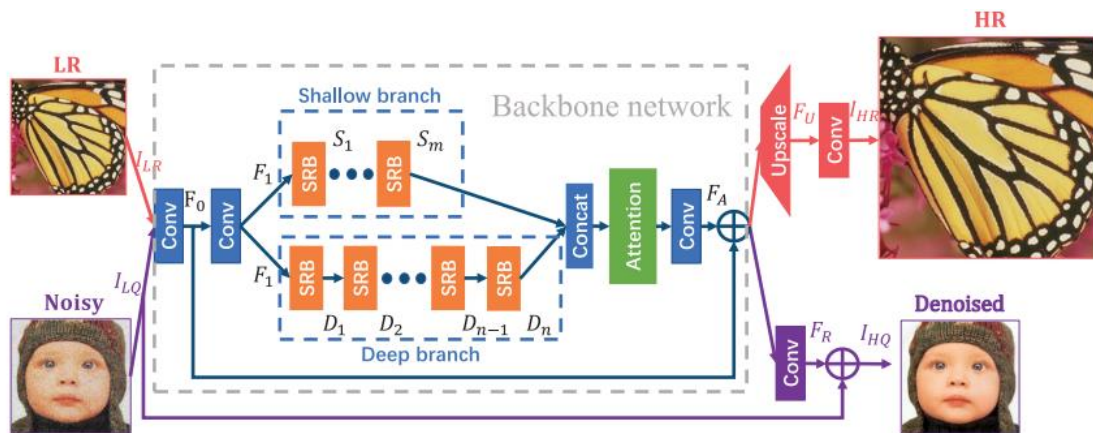
Esse paper introduz uma nova estrutura de redes residuais que consegue bom desempenho sem a utilização de normalização em batch. Essa estrutura, denominada *Dilated Residual Networks with Symmetric Skip Connection* (DSNet), depende da combinação de saltos de conexões simétricas e convolução dilatada. Essa nova abordagem é computacionalmente mais eficiente no treinamento, pois temos uma menor quantidade de camadas e parâmetros da rede. A rede tem vantagem quando se trata de ruídos gaussianos, mais especificamente. Por fim, essa arquitetura ultrapassa o desempenho de métodos do estado da arte e possui a possibilidade ser implementada se beneficiando do uso de GPUs.

Residual network with detail perception loss for single image super-resolution



Nesse paper, é apresentada uma proposta de rede neural residual com introdução de blocos simples em cascata para a super resolução de imagens. Esses blocos de multi-layer perceptron possuem a capacidade de extrair features e fazer um mapeamento complexo com menos parâmetros. A utilização de redes residuais permite pular conexões e reduzir o problema do vanishing-gradient das redes neurais profundas. Essa rede possui 2 ramificações, uma para prever as informações de alta frequência (detalhes) da imagem de alta resolução e outra para prever as informações de baixa frequência da imagem de alta resolução.

Two-stream deep sparse network for accurate and efficient image restoration



Esse artigo tenta resolver um problema recorrente nas técnicas de restauração de imagem que usam redes neurais convolucionais: ignorar a capacidade das features de alta frequência e as features de baixa frequência se complementarem, assim limitando a qualidade da restauração da imagem. O artigo propõe uma rede esparsa para aprender explicitamente as features rasas e as features profundas e avaliar a contribuição de cada uma na reconstrução. As features rasas geralmente são as bordas e formas da imagem, enquanto as features profundas são os detalhes mais específicos de cada imagem. Em cada ramificação da rede esparsa há um bloco responsável pra agregar de maneira eficiente as features hierárquicas através da construção de conexões esparsas entre as camadas.