

# Pesquisa Exploratória

## 1. Cross-Scene Crowd Counting via Deep Convolutional Neural Networks

- a. **Resumo:** Novas cenas de multidões que não estavam presentes no conjunto de treinamento, são o principal desafio na contagem de multidões. O desempenho da maioria dos métodos existentes de contagens de multidões diminuem significativamente quando aplicados a uma cena nova. Para resolver esse problema, foi proposto o uso de uma rede **neural convolucional profunda**, para contagem de multidões, e ela é treinada alternativamente com dois objetivos de aprendizado: densidade e contagem de multidões. Essa abordagem de aprendizado comutável proposta, capacita uma otimização para os dois objetivos.
- b. **Proposta:** Para lidar com uma cena nova, é apresentado um método orientado a dados para ajustar o modelo CNN treinado para a cena alvo.
- c. **Link:**[https://www.cv-foundation.org/openaccess/content\\_cvpr\\_2015/html/Zhang\\_Cross-Scene\\_Crowd\\_Counting\\_2015\\_CVPR\\_paper.html](https://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Zhang_Cross-Scene_Crowd_Counting_2015_CVPR_paper.html)

## 2. Learning To Count Objects in Images

- a. **Resumo:** Foco em imagens de treinamento anotadas com pontos (um ponto por objeto). Evita-se a tarefa de aprender a detectar e localizar instâncias de objetos individuais. Em vez disso, apresenta-se o problema de como estimar uma densidade de imagem cuja integralidade sobre qualquer região da imagem fornece a contagem de objetos na região. Aprender a inferir essa densidade pode ser formulado como uma minimização de uma função quadrática de risco regularizado.
- b. **Proposta:** Nova estrutura de aprendizado supervisionado para tarefas de contagem de objetos visuais. Como estimar um número de células em uma imagem microscópica ou o número de humanos em quadros de vídeo de vigilância.
- c. **Link:**<http://papers.nips.cc/paper/4043-learning-to-count-objects-in-images>

### 3. Single-Image Crowd Counting via Multi-Column Convolutional Neural Network

- a. **Resumo:** Foi proposto uma arquitetura simples, mas eficaz, da rede neural convolucional com várias colunas para mapear a imagem ao seu mapa de densidade de multidões. Essa rede neural permite que a imagem de entrada seja de tamanho ou resolução arbitrária. Ao utilizar filtros com campos receptivos de diferentes tamanhos, os recursos aprendidos por cada coluna da CNN são adaptáveis a variações no tamanho das pessoas devido ao efeito de perspectiva ou resolução da imagem. Além disso, o mapa de densidade real é calculado com precisão com base em núcleos adaptativos à geometria e não é necessário conhecer o mapa da perspectiva da imagem de entrada. Experimentos mostram que esse modelo, uma vez treinado em um conjunto de dados, pode ser prontamente transferido para um novo conjunto de dados.
- b. **Proposta:** Desenvolver um método que possa estimar a contagem de multidões a partir de uma imagem individual com densidade de multidão arbitrária e perspectiva arbitrária.
- c. **Link:** [https://www.cv-foundation.org/openaccess/content\\_cvpr\\_2016/papers/Zhang\\_Single-Image\\_Crowd\\_Counting\\_CVPR\\_2016\\_paper.pdf](https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Zhang_Single-Image_Crowd_Counting_CVPR_2016_paper.pdf)

### 4. Switching Convolutional Neural Network for Crowd Counting

- a. **Resumo:** A análise de multidões é composta por uma infinidade de fatores como interoclusão entre pessoas devido a aglomeração, alta similaridade de aparência entre pessoas, elementos de fundo e grande variabilidade de pontos de vista da câmera. As principais abordagens de ponta, abordam esses fatores usando arquiteturas CNN de várias escalas, redes recorrentes e fusão tardia de recursos da CNN de várias colunas com diferentes campos receptivos. Propôs-se a troca de rede neural convolucional que aproveita a variação da densidade de multidão dentro de uma imagem para melhorar a precisão e a localização da contagem prevista. Patches de uma grade dentro de uma cena de multidão são retransmitidos para regressores independentes da CNN com base na qualidade de previsão de contagem de multidão da CNN estabelecida durante o treinamento. Os regressores da CNN independentes são projetados para ter campos receptivos diferentes e um classificador de comutador é treinado para retransmitir o patch da cena da multidão para o melhor regressor CNN. Observa-se que a troca retransmite um patch de imagem para uma coluna da CNN específica, com base na densidade da multidão.
- b. **Proposta:** Foi proposto um novo modelo de contagem de multidões que mapeia uma dada cena de multidões para sua densidade.
- c. **Link:** <https://ieeexplore.ieee.org/abstract/document/8099912>

# Conceitos Básicos

## 1. Deep Neural Network

Uma rede neural profunda (deep neural network, DNN) é uma rede neural artificial, com várias camadas entre as camadas de entrada e saída. A rede neural encontra a manipulação matemática correta para transformar a entrada na saída, seja uma relação linear, ou uma relação não linear. A rede se move pelas camadas, calculando a probabilidade de cada saída. Cada manipulação matemática é considerada uma camada.

A princípio, o DNN cria um mapa de neurônios virtuais e atribui valores numéricos aleatórios, ou pesos, “às” conexões entre eles. Os pesos e entradas são multiplicados e retornam uma saída entre 0 e 1. Se a rede não reconhece com precisão um padrão específico, um algoritmo ajusta os pesos, até encontrar a manipulação matemática correta para processar completamente os dados.

## 2. Convolutional Neural Network

Em deep learning, uma rede neural convolucional é uma classe de redes neurais profundas, mais comumente aplicada à análise de imagens visuais. O nome indica que a rede emprega uma operação matemática chamada convolução. As redes convolucionais são redes neurais que usam a convolução no lugar da multiplicação geral da matriz em pelo menos uma de suas camadas.

Uma rede neural convolucional, consiste em uma camada de entrada, uma de saída, além de várias camadas ocultas. As camadas ocultas fazem uma convolução. A função de ativação é geralmente uma camada RELU (retificador)

As camadas convolucionais envolvem a entrada e passam seu resultado para a próxima camada. Isso é semelhante à resposta de um neurônio no córtex visual a um estímulo específico. Nas redes neurais, cada neurônio recebe entrada de algum neurônios locais na camada anterior. Em uma camada totalmente conectada, cada neurônio recebe informações de todos os elementos da camada anterior. Em uma camada convolucional, os neurônios recebem entradas apenas de uma sub área restrita da camada anterior. Em uma camada totalmente conectada, o campo receptivo é toda a camada anterior. Em uma camada convolucional, a área receptiva é menor do que toda a camada anterior.

Cada neurônio em uma rede neural, calcula o valor de uma saída aplicando uma função específica aos valores de entrada provenientes do campo receptivo na camada anterior.

## 3. Convolução

No campo matemático análise funcional, a convolução é uma operação matemática em duas funções que gera uma terceira função que retorna a influência de uma das funções da entrada com a outra. Algumas características da convolução são semelhantes à correlação cruzada.