

Processamento de Imagem

Advanced Institute for Artificial Intelligence – AI2

<https://advancedinstitute.ai>

Agenda

- ☐ Visão computacional
- ☐ Processamento de Imagem
- ☐ Análise de Imagem
- ☐ Representação de imagem
- ☐ Desafios

A visão computacional procura auxiliar a resolução de problemas altamente complexos, buscando imitar a cognição humana e a habilidade do ser humano em tomar decisões de acordo com as informações contidas na imagem.

- O processamento digital de imagens consiste em um conjunto de técnicas para capturar, representar e transformar imagens com o auxílio de computador.
- O emprego dessas técnicas permite extrair e identificar informações das imagens e melhorar a qualidade visual de certos aspectos estruturais, facilitando a percepção humana e a interpretação automática por meio de máquinas.

- A análise/interpretação de imagens visa obter uma descrição que contenha informação suficiente para distinguir entre diferentes objetos de interesse, de forma confiável e requerendo o mínimo de intervenção humana.
- Certos passos relevantes que envolvem o processamento e a análise de imagens, muitas vezes, é realizada por um operador humano que detem o conhecimento ou a experiência sobre o domínio da aplicação.

- Análise de imagens capturadas por meio de microscópios ópticos ou eletrônicos beneficia áreas que variam diversas como biologia por exemplo.
- Exemplos de aplicação:
 - Análise de estruturas em cristalografia
 - Análise e identificação de espécies de insetos

A análise de fotografias aéreas ou imagens de satélite permite uma melhor compreensão da superfície terrestre, auxiliando tarefas como:

- ☐ Estudo de fenômenos naturais como vulcões
- ☐ Padrões obtidos por copas de árvores
- ☐ Identificação de áreas de plantio, floresta
- ☐ Obtenção de características diversas por meio de imagens aéreas
- ☐ Estimativa de características de uma propriedade avaliando área construída

Processo de aquisição de uma imagem (Digitalização)

- ☐ Diversos dispositivos podem ser utilizados para realizar aquisição de imagens
 - Cameras digitais, scanners, satellite, ressonância magnética, etc
- ☐ As imagens obtidas por tais dispositivos são representadas por meio de um conjunto de pixels
- ☐ Os dispositivos podem possuir um número variável de canais para obtenção de uma mesma imagem

A representação de uma imagem é feita por meio dos seguintes valores: Altura, Largura e canais.

- ☐ A altura e largura definem a quantidade de itens que compõem a imagem que é chamada de pixel
- ☐ Cada pixel é uma unidade mínima da imagem e pode ser descrito de acordo com diferentes informações chamadas canais

Uma imagem pode ser representada como uma matriz de 3 dimensões sendo altura x largura x canais.

Resolução Espacial

- ❑ A quantidade de pixels presente em uma mesma região da imagem define a resolução.
- ❑ Quanto mais pixels em uma mesma área da imagem, maior a resolução e portanto, maior o nível de detalhe da imagem
- ❑ Uma imagem contendo um grande número de pixels não necessariamente possui resolução maior que outra contendo menor número de pixels.
- ❑ A resolução de uma imagem deve ser escolhida de modo a atender ao grau de detalhes que devem ser discerníveis na imagem

Exemplos:

- Uma área de 200 cm²
 - Se usarmos 10 pixels na dimensão x e 10 pixels na dimensão y, cada pixel vai corresponder a 2 cm²
 - Se usarmos 20 pixels na dimensão x e 20 pixels na dimensão y, cada pixel vai corresponder a 1 cm²
- A resolução influencia a capacidade humana de visualização
- A resolução influencia na capacidade de algoritmos realizarem funções automáticas diversas na imagem

Profundidade da imagem

- ☐ Quantidade de variação de valores para cada pixel na representação da imagem
- ☐ Se utilizamos um número alto de profundidade a imagem possui mais detalhes
- ☐ Com um número baixo de profundidade a imagem vai possuir menos detalhes
- ☐ O Processamento da imagem depende de um nível adequado de detalhe
- ☐ Diminuir a profundidade pode auxiliar em facilitar a manipulação da imagem

- ❑ Uma biblioteca muito popular para processamento de imagem é o opencv
- ❑ O matplotlib pode ser usado para plotar um array como uma imagem
- ❑ A seguir mostramos um código elementar para manipulação de imagens

```
import cv2 as cv
import matplotlib.pyplot as plt
img = cv.imread('ab5.jpg')
plt.imshow(img)
```

Desafios relacionados ao processamento de imagens

- ☐ Aumentar ou diminuir
- ☐ Rotacionar
- ☐ Correções
- ☐ Omitir ou realçar detalhes

Processamento de Imagem

Aumentar ou diminuir uma imagem



- ☐ Quando mudamos as medidas da imagem, pode ocorrer distorções
- ☐ Nesse sentido, é necessário considerar as proporções para fazer um redimensionamento correto

Rotacionar uma imagem



- ☐ Ao rotacionar uma imagem, é possível que parte da imagem seja cortada
- ☐ É necessário considerar um novo tamanho que permita incluir a nova imagem completamente

Outras manipulações comuns em uma imagem

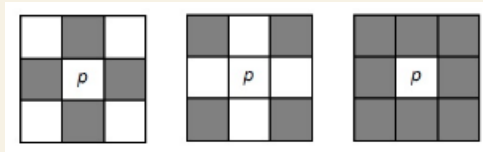
- ☐ Recortar uma área de interesse
- ☐ Incluir um elemento na imagem
 - Incluir texto
 - Incluir forma
 - Incluir outra parte da mesma imagem

Outras operações podem ser realizadas em imagens por meio de filtros:

- ☐ Detectar borda, contornos
- ☐ Clarear, escurecer
- ☐ Eliminar ruídos
- ☐ Transformações morfológicas
- ☐ Borrar (Blur)

Tais operações auxiliam na identificação de padrões.

- ❑ Filtro é uma operação que é aplicada em todos os pixels de uma imagem, alterando alguns pixels de acordo com certos critérios
- ❑ Normalmente, a análise de pixels é feito por regiões chamadas vizinhança
 - A vizinhança de um pixel é definido como o pixel imediatamente acima, abaixo ou aos lados
 - Normalmente as operações levam em conta a variação de valores da vizinhança do pixel e raramente consideram os pixels isoladamente
- ❑ O processo de filtragem é feito utilizando matrizes denominadas máscaras, as quais são aplicadas sobre a imagem



blur (borrar)

- ☐ Filtro para remover pixels considerados "outlier"
- ☐ Tais pixels normalmente representam ruídos na imagem
 - Ruídos podem ser gerados por diversos fatores, como o próprio processo de aquisição

Thresholding

- ☐ Filtro para transformar imagem em imagens binárias
- ☐ Auxilia em encontrar padrões de contorno na imagem

Filtro de Sobel

- ☐ Calcula o gradiente da intensidade da imagem em cada ponto
- ☐ Obtendo-se assim uma indicação se a variação de luminosidade em um ponto ocorreu de forma abrupta ou suave
- ☐ Essa transição claro-escuro ajuda na identificação de contornos

Transformações Morfológicas

- ❑ Erosão: elimina elementos da imagem
- ❑ Dilatação: aumenta elementos da imagem
 - Abertura e Fechamento: combina os efeitos da Erosão e Dilatação

Efeitos da Erosão

- ☐ Diminuir partículas
- ☐ Eliminar componentes menores que o elemento estruturante
- ☐ Permitir a separação de componentes conectados

Efeitos da dilatação

- ☐ Aumentar partículas
- ☐ Preencher buracos
- ☐ Conectar componentes próximos

Identificação de padrões como linhas ou círculos:

- ☐ Hough Lines
- ☐ Hough Circles

Filtros do tipo black-hat

- ☐ Retorna elementos com padrão de tamanho menor e mais escuros que pixels da vizinhança
- ☐ O tamanho dos elementos extraídos podem ser controlados pelo elemento estruturador

Convolução

- ☐ O OpenCV fornece uma função, `cv2.filter2D ()`, para aplicar uma operação de convolução em um kernel com uma imagem.
- ☐ Esse processo pode ser generalizado para obter qualquer tipo de padrão em uma imagem

Camada de convolução em Keras:

```
model.add(layers.Conv2D(32,(5,5),activation='relu', input_shape=(28, 28,1)))
```

Camada de Pooling: filtro para redução de dimensionalidade

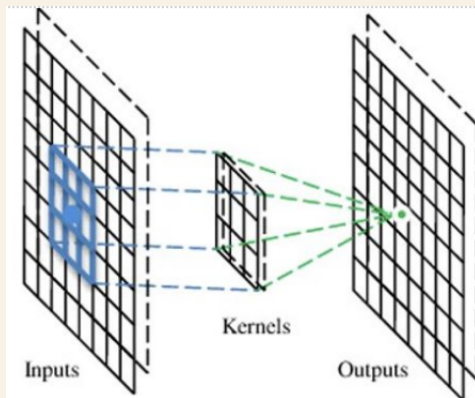
```
model.add(layers.MaxPooling2D((2, 2)))
```

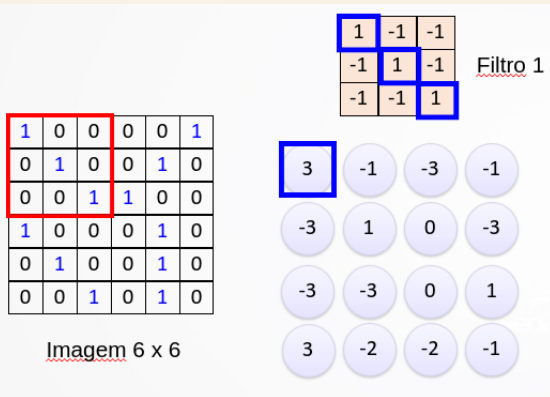
- Normalmente tais camadas são combinadas para realçar padrões que definem características da imagem
- Dessa forma, o classificador da rede Neural atua apenas nos padrões encontradas, e não em todos os detalhes

- ❑ camadas densas aprendem padrões globais em seu espaço de features de entrada
- ❑ camadas de convolução aprendem padrões locais: no caso de imagens, padrões encontrados em pequenas janelas da entrada Baseados em filtros (processamento de imagem)
- ❑ Os padrões aprendidos em uma convolução são invariantes à translação.
 - Os mesmos padrões podem aparecer várias vezes na mesma entrada em diferentes posições

- ❑ Depois de aprender um determinado padrão no canto inferior direito da imagem, uma CNN pode reconhecê-lo em qualquer lugar: por exemplo, no canto superior esquerdo.
- ❑ Uma MLP precisaria de amostras para cada possibilidade de ocorrência do padrão na imagem
- ❑ CNNs precisam de menos amostras de treinamento para aprender representações que têm poder de generalização.

- ☐ Uma Matriz de Entrada (input), passa por um filtro (matriz Menor) e gera uma Matriz de Saída
- ☐ A aplicação desse filtro é chamado convolução
- ☐ Cada filtro é capaz de detectar um padrão em uma imagem





1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

Imagem 6 x 6

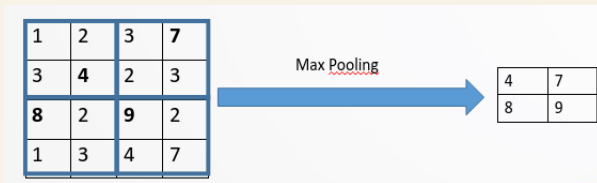
1	-1	-1
-1	1	-1
-1	-1	1

Filtro 1

3	-1	-3	-1
-3	1	0	-3
-3	-3	0	1
3	-2	-2	-1

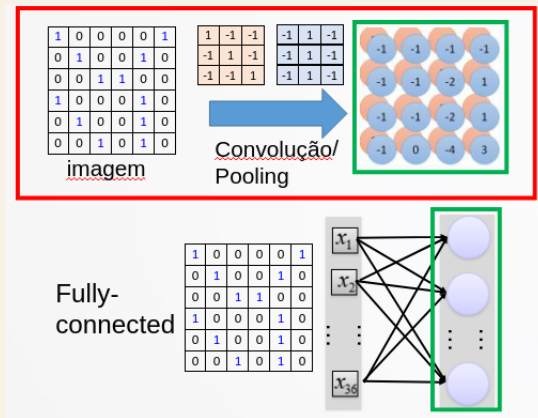
- ❑ Como resultado da operação de Convolução uma matriz com um tamanho menor é gerada
- ❑ Para impedir que a matriz perca informação nesse proceso é usado o recurso de Padding
- ❑ Uma borda é inserida na matriz original, de forma que a matriz resultante é do tamanho da original
- ❑ Tipos de padding
 - valid
 - zero
 - same

- ☐ As camadas do tipo pooling são utilizadas para simplificar a representação de uma imagem.
- ☐ Dois métodos comuns de pooling são:
- ☐ pool médio : substituir todos os pixels de uma região pela média do valor dos pixels
- ☐ pool máximo: substituir todos os pixels de uma região pelo pixel de maior do valor no intervalo



□ Rede Totalmente Conectada (MLP) × Convolução

- No processo de Convolução filtros reduzem a imagem a matrizes menores
- São encontrados padrões comuns em diversas partes de uma mesma imagem
- Na rede MLP a entrada é avaliada de modo único
- Para MLP aprender que um mesmo padrão aparece em duas ou mais partes da imagem, são necessários muitos exemplos



- Uma Rede Neural Convolucional é montada tipicamente da seguinte forma:
 - Conjuntos de camadas de convolução seguida por uma camada pool
 - Tais camadas correspondem a identificação de padrões diversos na imagem
 - Ao final é introduzida uma camada densa para classificação

- ❑ Uma rede convolucional treinada pode ser utilizada como base para montar atacar outro problema
- ❑ Tal abordagem é feita reutilizada uma topologia treinada e substituindo e retreinando apenas a camada de classificação
- ❑ A vantagem dessa abordagem, é que todas as capacidades de detecção de padrão se mantêm para a nova rede.