

Visão Computacional e Aprendizado de Máquina

Breve introdução

Vitor Greati¹ Vinícius Campos¹

¹Universidade Federal do Rio Grande do Norte

Sumário

O gap semântico

Visão Computacional

Para além de *pixels*

Aplicações

Técnicas

Aprendizagem de Máquina

O que você percebe nestas imagens?



O que você percebe nestas imagens?



A facilidade com que respondemos a essa pergunta se deve ao nosso sistema visual **nativo** extremamente poderoso!

O que o computador percebe nessas imagens

As matrizes de *pixels*

À primeira vista...

```
[[ 42  23  19 ...,  21  29  25]
 [ 40  40  36 ...,  24  24  21]
 [ 28  30  36 ...,  30  13  27]
 ...,
 [115  78  45 ...,  28  36  17]
 [ 67  78 192 ...,  35  31  36]
 [ 67  79 104 ...,  34  32  31]]
```

```
[[138 137 137 ..., 107 107 107]
 [135 134 134 ..., 107 107 107]
 [130 129 129 ..., 107 107 107]
 ...,
 [145 145 146 ..., 142 142 142]
 [146 145 144 ..., 144 144 145]
 [147 146 144 ..., 145 145 146]]
```

```
[[222 224 224 ..., 204 201 200]
 [223 225 223 ..., 201 203 204]
 [226 226 226 ..., 204 202 205]
 ...,
 [210 203 208 ..., 192 188 189]
 [206 206 207 ..., 190 188 189]
 [210 208 210 ..., 191 193 185]]
```

```
[[ 48  45  40 ...,  28  29  31]
 [ 45  46  43 ...,  28  29  30]
 [ 41  43  43 ...,  27  27  29]
 ...,
 [101 101 103 ...,  64  51  32]
 [ 98  97  99 ...,  63  71  57]
 [ 97  97  97 ...,  38  57  65]]
```

O que o computador percebe nessas imagens

As matrizes de *pixels*

À primeira vista...

```
[[ 42  23  19 ...,  21  29  25]
 [ 40  40  36 ...,  24  24  21]
 [ 28  30  36 ...,  30  13  27]
 ...,
 [115  78  45 ...,  28  36  17]
 [ 67  78 192 ...,  35  31  36]
 [ 67  79 104 ...,  34  32  31]]
```

```
[[138 137 137 ..., 107 107 107]
 [135 134 134 ..., 107 107 107]
 [130 129 129 ..., 107 107 107]
 ...,
 [145 145 146 ..., 142 142 142]
 [146 145 144 ..., 144 144 145]
 [147 146 144 ..., 145 145 146]]
```

```
[[222 224 224 ..., 204 201 200]
 [223 225 223 ..., 201 203 204]
 [226 226 226 ..., 204 202 205]
 ...,
 [210 203 208 ..., 192 188 189]
 [206 206 207 ..., 190 188 189]
 [210 208 210 ..., 191 193 185]]
```

```
[[ 48  45  40 ...,  28  29  31]
 [ 45  46  43 ...,  28  29  30]
 [ 41  43  43 ...,  27  27  29]
 ...,
 [101 101 103 ...,  64  51  32]
 [ 98  97  99 ...,  63  71  57]
 [ 97  97  97 ...,  38  57  65]]
```

Imagens digitais monocromáticas

Matrizes $I_j \in \mathbb{M}_{w_j \times h_j}([0, \dots, 255])$ ou funções

$f_j : \{1, \dots, w_j\} \times \{1, \dots, h_j\} \rightarrow [0, 255]$, onde w_j é a largura e h_j é a altura da imagem j .

O *gap* semântico

Percepção humana × Percepção da máquina

Gap semântico

Diferença entre a maneira como o ser humano **percebe** o conteúdo de uma imagem e como a imagem pode ser **representada** de forma manipulável no computador.

Sumário

O gap semântico

Visão Computacional

Para além de *pixels*

Aplicações

Técnicas

Aprendizagem de Máquina

Visão Computacional

O que é?

Visão Computacional

Visão Computacional é uma área da Ciência da Computação cujo propósito é capacitar os computadores para extraírem informações de imagens, ou seja, permitir que tenham um entendimento visual do mundo.

Visão Computacional

O que é?

Visão Computacional

Visão Computacional é uma área da Ciência da Computação cujo propósito é capacitar os computadores para extraírem informações de imagens, ou seja, permitir que tenham um entendimento visual do mundo.

Algumas das práticas mais comuns estão ligadas à tarefa de **reconhecimento**, nas formas de:

Visão Computacional

O que é?

Visão Computacional

Visão Computacional é uma área da Ciência da Computação cujo propósito é capacitar os computadores para extraírem informações de imagens, ou seja, permitir que tenham um entendimento visual do mundo.

Algumas das práticas mais comuns estão ligadas à tarefa de **reconhecimento**, nas formas de:

- ▶ Classificação de objetos
- ▶ Identificação
- ▶ Detecção

Visão Computacional

Desafios

Variação de ponto de vista

Não importa sob qual ângulo se fotografe um gato: ele continuará sendo um gato.

Variação de escala

Não importa a que distância o gato estará da câmera: ele continuará sendo um gato.

Deformação

Um gato pode estar esticando suas pernas ou/e contorcendo seu pescoço para se lamber, e isso não o faz ser outro ser além de um gato na imagem.

Visão Computacional

Desafios

Oclusão

Um gato pode estar espiando o mundo ao redor de dentro de uma caixa, apenas com a cabeça de fora, e ele continuará sendo um gato à lente de uma câmera em frente à caixa.

Iluminação

Um gato num estacionamento mal iluminado ainda é um gato.

Ruído de fundo

Um gato em frente à uma tela de TV repleta de ruído ainda é um gato.

Variações intra-classe

Gatos de diversas raças, cores e tamanhos serão sempre gatos.

Sumário

O gap semântico

Visão Computacional

Para além de *pixels*

Aplicações

Técnicas

Aprendizagem de Máquina