

# Inteligência Artificial na Visão Computacional

Revisado 03/03/2024

[PSI3471 aula 0 – início] [PSI5790 aula 1 parte 1 - início]

## I. Áreas envolvidas neste curso

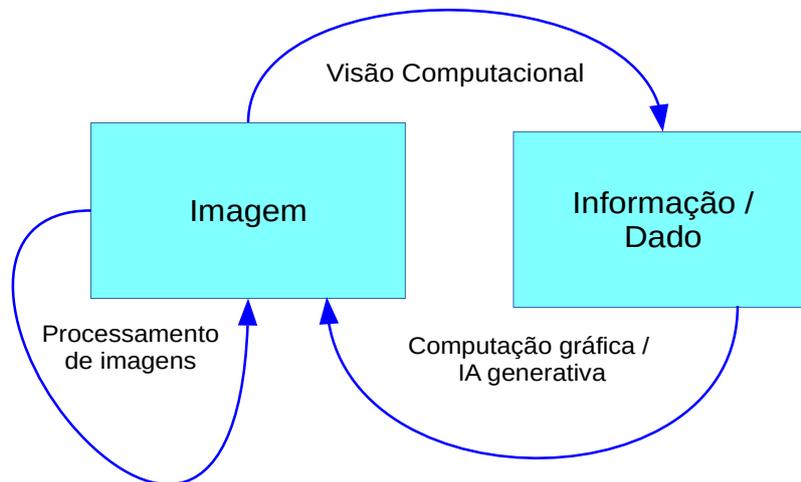


Figura 1: Processamento de imagens e visão computacional.

Processamento de imagens manipula imagens, de modo que a entrada e a saída do processo são imagens (figura 1). Por exemplo, reduzir ruído da imagem, ajustar brilho/contraste, aumentar resolução, rotacionar, etc.

Visão computacional busca extrair informações relevantes de imagens, de modo que a entrada do processo são imagens mas a saída são informações contidas na imagem (figura 1). Por exemplo, ler o número da placa de um automóvel, verificar se duas fotos se referem à mesma pessoa, descobrir o paciente tem Covid analisando radiografia, etc.

IA generativa gera imagens a partir de alguma informação fornecida. Não veremos em detalhes neste curso.

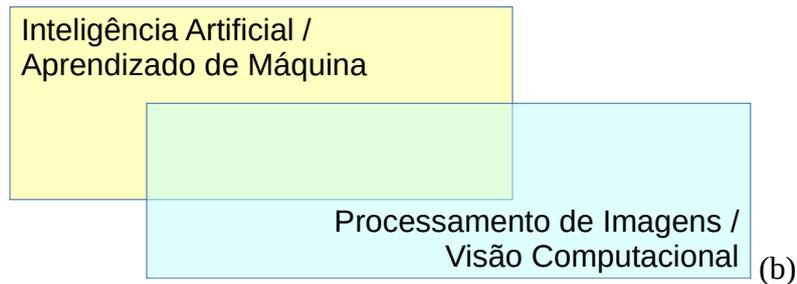
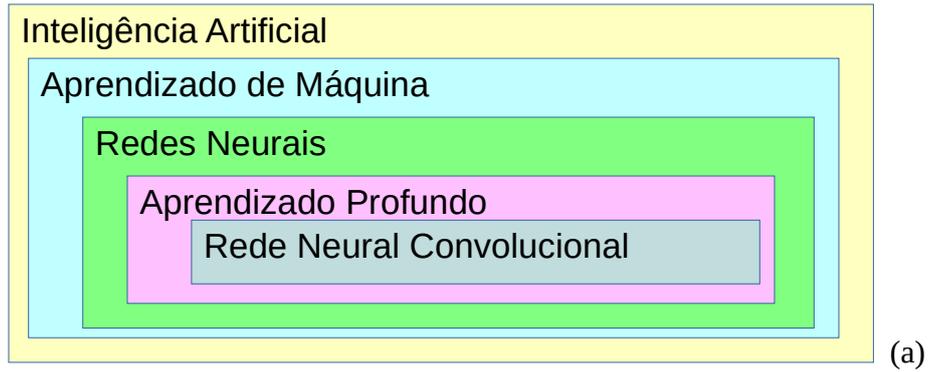


Figura 2: Relações entre diferentes áreas envolvidas neste curso.

## II. Conteúdo do curso

Neste curso, vamos estudar:

1. Processamento de imagens e visão computacional “convencional” (sem usar aprendizado de máquina – figura 2b ciano sem intersecção).
2. Aprendizado de máquina “clássico”, profundo e rede neural convolucional (figura 2a).
3. Aprendizado de máquina usando rede neural convolucional, com aplicações em processamento de imagens e visão computacional (figura 2b, área em intersecção).

### **III. Inteligência artificial, aprendizado de máquina e aprendizado profundo**

Inteligência artificial (IA) está realizando tarefas que pareciam impossíveis há apenas 10 anos atrás.

Estamos dando passos largos para chegar a um computador que:

- Entende o que vê;
- Lê e entende textos científicos, médicos e jurídicos;
- Compreende a fala humana e conversa em vários idiomas;
- Consegue gerar textos, imagens, músicas, vídeos e programas computacionais de acordo com a descrição fornecida;
- Aprende sozinho, a partir das experiências.

Praticamente todos os dias aparecem várias notícias sobre IA na grande mídia. É só colocar “inteligência artificial” num site de busca de notícias. Por exemplo:

- 28/02/2024 BBC Brasil: “‘Eram meu rosto e minha voz, mas era golpe’: como criminosos 'clonam pessoas' com inteligência artificial”
- 02/03/2024 R7: “Inteligência Artificial nas Eleições: Mesmo com resolução do TSE, tecnologia ainda ameaça processo eleitoral”
- 01/03/2024 ValorInveste: “Lá vem a IA: após resultados fortes, ações da Dell sobem 30%, e puxam Nvidia e rival”







## VI. Algumas aplicações que veremos no curso

Abaixo, apresento algumas aplicações de IA em visão computacional que veremos nos cursos PSI3471/PSI3472 ou PSI5790.

### 1) Classificação de imagens.

Como desenvolver um sistema que reconhece que na foto tem chimpanzé, orangotango, etc. (figura 5)? Antes do aprendizado profundo, este impossível resolver este problema. Vamos estudar no curso como isto pode ser feito.



Figura 5: Classificar imagens pelo conteúdo.

Vários problemas de classificação de imagens foram objeto de exercícios programas (EPs) deste curso. Num deles, pedi para classificar imagens de rostos em masculino ou feminino (figura 4). As boas soluções cometiam somente 1% de erro.

Num outro EP, pedi para classificar a variedade de grãos de arroz (figura 6).

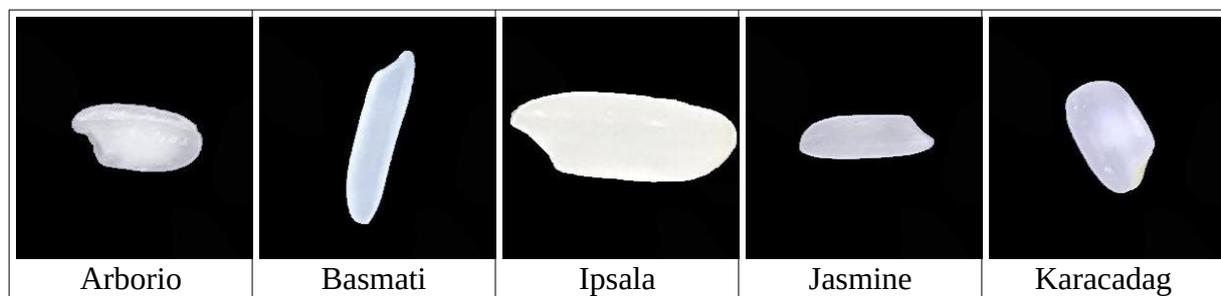


Figura 6: Os 5 tipos de grãos de arroz do conjunto de dados.

Um outro EP consistiu em classificar radiografias de pulmão em “normal” (sadio), com “co-vid-19” e “non-covid” (está com alguma infecção viral ou bacteriana mas não é covid) – figura 7. A taxa de acerto foi aproximadamente 95%.

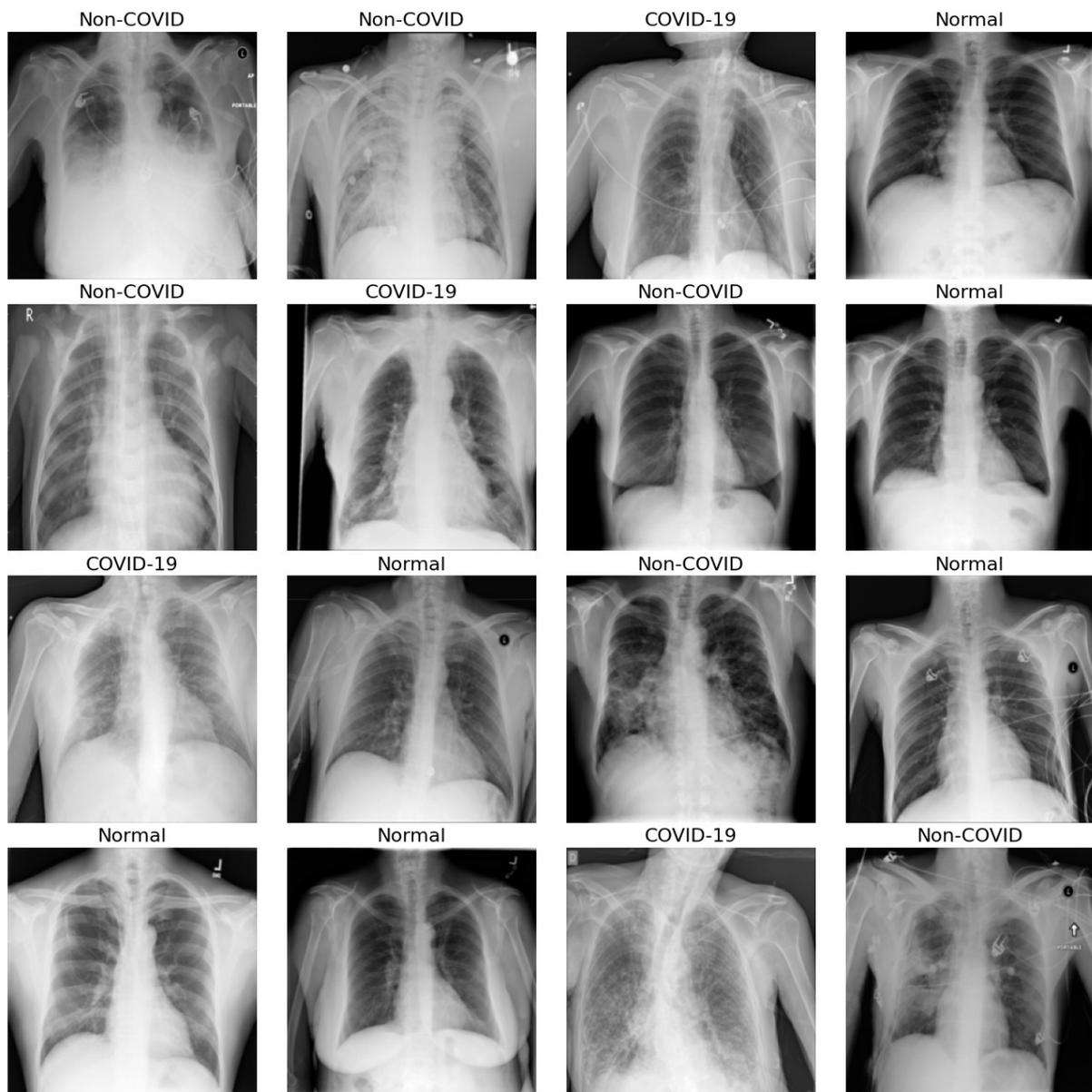


Figura 7: Exemplos de imagens do conjunto de imagens COVID-QU-Ex.

## 2) Detecção de objetos:

Também vamos estudar a detecção de objetos (figura 8): Localizar na imagem onde está cada objeto.

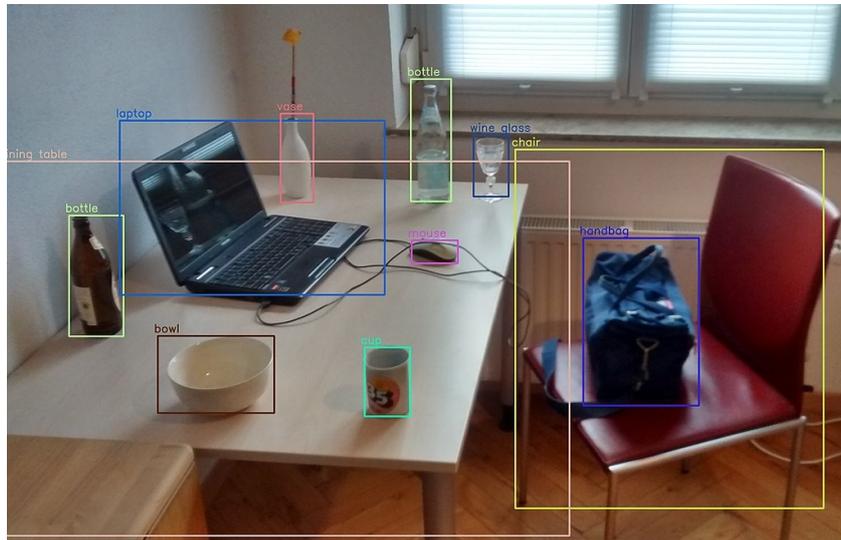


Figura 8: Detecção de objetos.

Na figura 9, além de classificar mamografia (raio-x da mama) em câncer/não-câncer, é possível criar “heatmap” indicando onde provavelmente está a lesão.

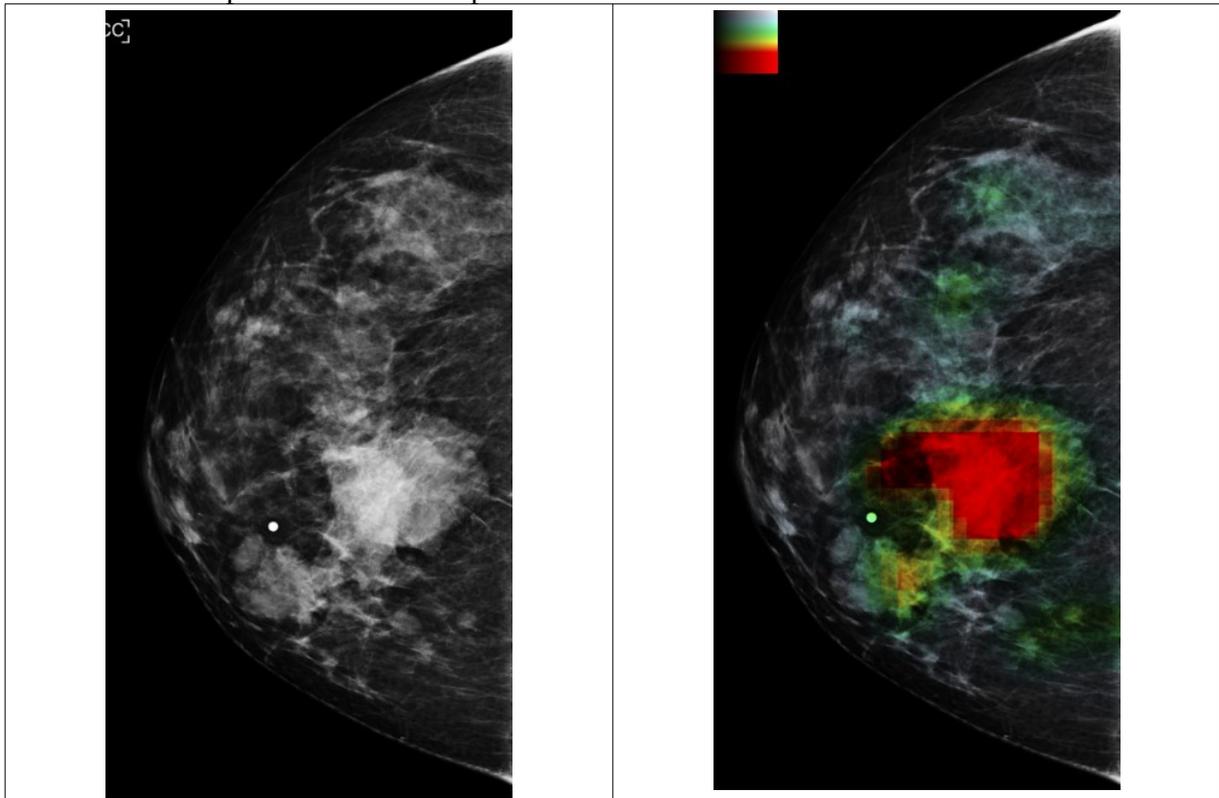


Figura 9: Computador consegue classificar mamografia em lesão maligna, lesão benigna ou sem lesão. Também pode gerar “heatmap” indicando o local provável de câncer.

### 3) Segmentação semântica:

Segmentação semântica consiste em classificar cada pixel da imagem em classes. Figura 10 mostra a segmentação da imagem em objeto de interesse (chimpanzé) e o fundo da imagem. Figura 11 segmenta as membranas das células, pintando-as de preto.

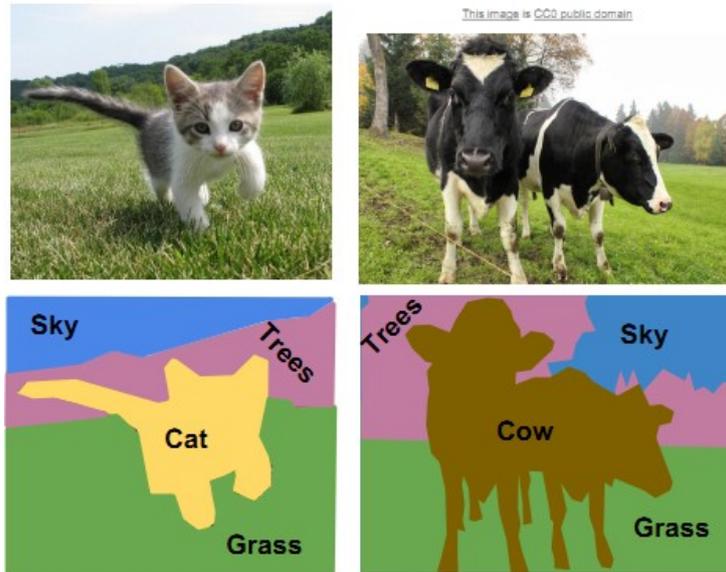


Figura: Segmentação semântica



Figura 10: Segmentação da imagem em objeto de interesse (chimpanzé) e fundo da imagem.

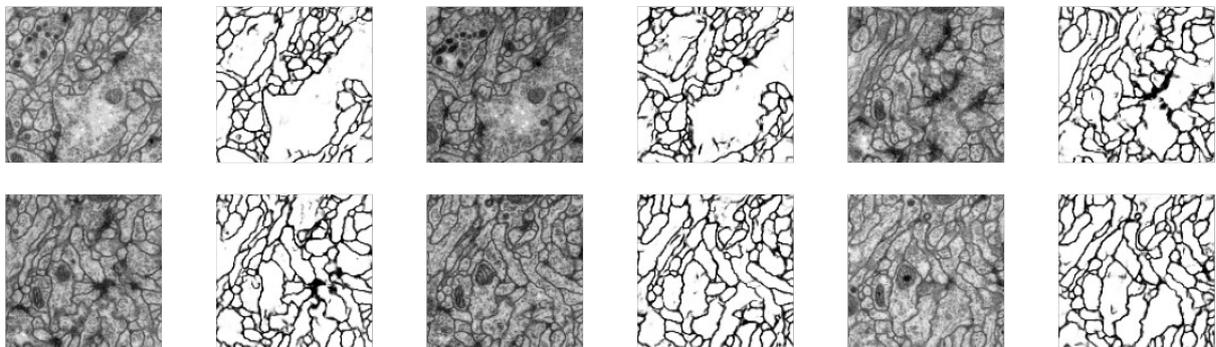


Figura 11: Algumas células (imagens da esquerda) e as membranas das células segmentadas.

#### 4) Colorir fotos preto-e-branco.

É possível usar aprendizado profundo para colorir automaticamente fotos preto-e-branco (figuras 12 e 13). Na internet, há vários documentários coloridos sobre a segunda guerra mundial, criados a partir de filmes originalmente gravados em preto-e-branco.



Figura 12: Exemplos de imagens coloridos artificialmente.

Como colorir uma foto preto-e-branco? Mostrar ao sistema IA muitas imagens coloridas. A partir de uma imagem colorida, o sistema gera imagem em níveis de cinzas correspondente. Tendo muitos pares de imagens de treino (níveis de cinza, coloridas), o sistema aprende a converter foto em tons de cinza para colorida.



Figura 13: Outros exemplos de colorir automaticamente imagens preto-e-branco.

Figura 14 mostra um EP desta disciplina para colorir imagens. Há casos em que é impossível descobrir a cor original de uma certa região da imagem, por exemplo, a cor do cachecol da moça da última linha. A original era verde mas o sistema pintou de marrom, pois não é possível descobrir a cor original a partir da imagem em níveis de cinzas.



Figura 14: Colorir imagens automaticamente.

## 5) Super-resolução:

Uma outra aplicação é aumentar a resolução das imagens, para obter imagens mais nítidas. As figuras 15 e 16 mostram o aumento da resolução de imagens 8x8 para 32x32. A figura 16 corresponde a um EP desta disciplina.

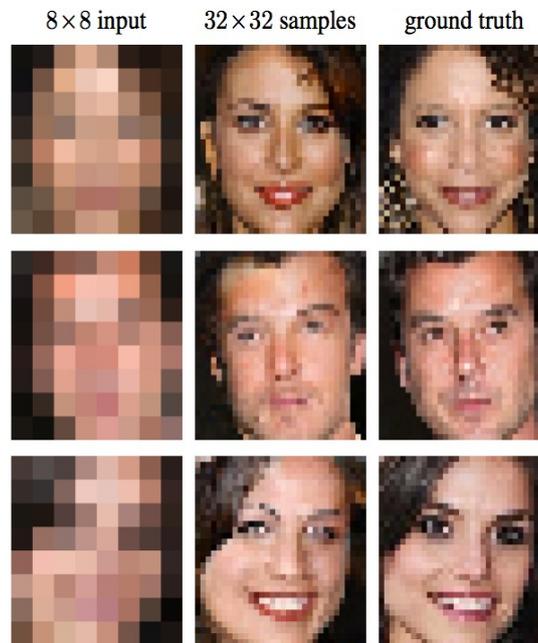


Figura 15: Super-resolução de imagens. Esquerda: Imagem em baixa resolução. Centro: Imagem gerada pela rede neural. Direita: Imagem em alta resolução ideal.



Figura 16: Imagens de super-resolução de um EP desta disciplina. Da esquerda para direita: Imagem original 8x8; interpolação bilinear para 32x32; interpolação bicúbica para 32x32; super-resolução com aprendizado profundo; saída ideal 32x32.

## 6) Reconhecimento facial:

Um problema comum é o reconhecimento facial, isto é, dada uma imagem facial, achar a pessoa correspondente num banco de dados (BD) de faces ou dizer que a pessoa não está no BD. Aqui, não podemos usar técnicas tradicionais de classificação, pois não é possível construir um BD com centenas de fotos de cada pessoa.

Num EP da disciplina, dada uma imagem facial, pedia-se que classifique a pessoa num dos 5 “vingadores”. Foi possível atingir taxa de erro de 1,2%.



Figura: Exemplos de imagens faciais dos 5 atores de vingadores.



### 3) Transferir estilo de pintura

Nesta aplicação, uma foto é redesenhada de acordo com o estilo de pintura de diferentes artistas (figura 19).

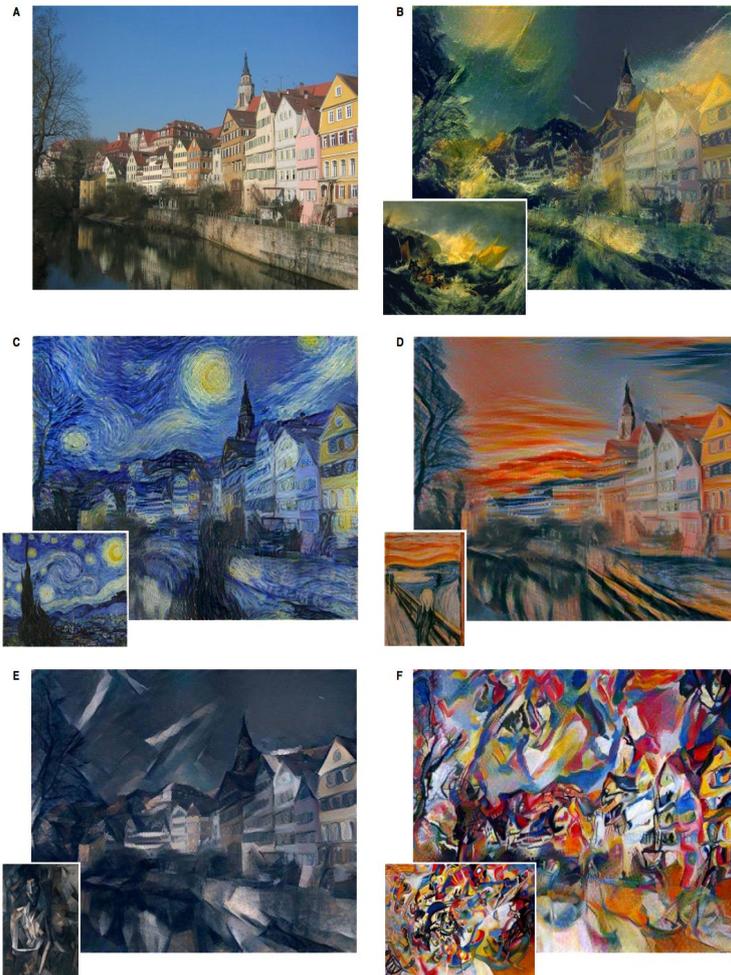


Figura 19: Transferência de estilo.

#### 4) Gerar imagens de objetos, animais ou rostos de pessoas inexistentes:

As pessoas da figura 20 não existem. Essas fotos foram geradas pelo aprendizado profundo. Na verdade, IA juntou as características de milhares rostos reais para criar um modelo que gera novos rostos a partir de números aleatórios, indistinguíveis de rostos verdadeiros.



Figura 20: Rostos gerados artificialmente.

<https://techxplore.com/news/2018-12-nvidia-face-making-approach-genuinely-gan-tastic.html>

O site:

<https://thisxdoesnotexist.com/>

gera versões fake de vários objetos, animais e rostos. Figura 21 mostra exemplos dessas imagens.

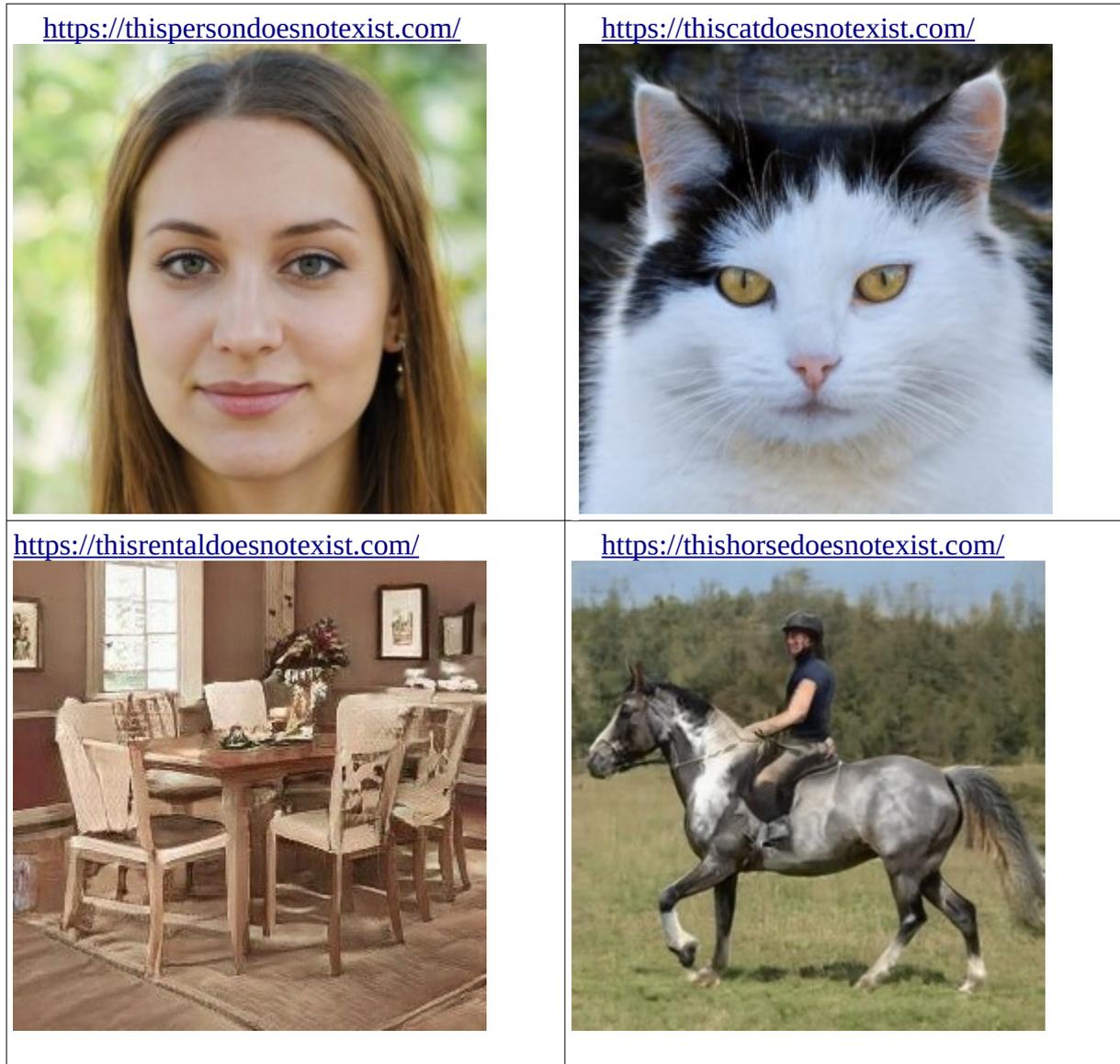


Figura 21: As imagens geradas artificialmente.

## 5) Aprender a jogar um jogo

O aprendizado por reforço é uma das sub-áreas do aprendizado de máquina que estuda como um agente inteligente deve realizar ações em um ambiente dinâmico para maximizar a recompensa. Aprendizado por reforço pode ser usada para fazer o computador a aprender jogar jogos, fazendo-o jogar várias vezes e informando-o se ganhou ou perdeu (figura 22).

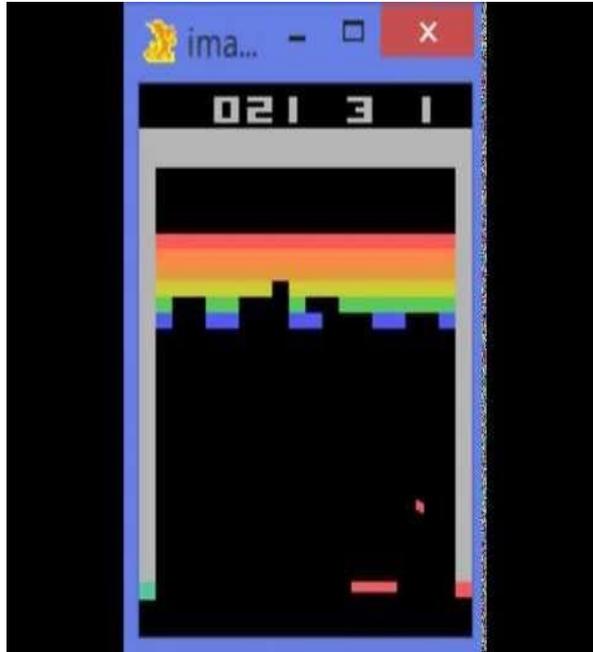


Figura 22: Aprendizado por reforço pode ser usada para ensinar o computador a jogar jogos.

## 6) Notícia sobre “leitura da mente”

Uma notícia de 09/abril/2023

<https://universoracionalista.org/experiencia-incrive-mostra-que-a-inteligencia-artificial-pode-ler-mentes-para-visualizar-nossos-pensamentos/>

e o artigo

[Chen2023] <https://arxiv.org/pdf/2211.06956.pdf>

dizem que o computador consegue “ler a mente”. Computador tenta “adivinhar”, a partir de sinais de fMRI (ressonância magnética funcional), a imagem que o paciente está vendo.

Em um experimento, a equipe coletou dados da atividade cerebral dos participantes via fMRI enquanto viam algumas imagens (figura 23). A partir dos sinais de fMRI, rede neural conseguiu “adivinhar” a imagem que os participantes estavam vendo.

Daqui a pouco, o computador pode conseguir decodificar os nossos sonhos.

As figuras 23 e 24 mostram a imagem que o paciente estava vendo e a correspondente imagem “adivinhada” pelo computador.

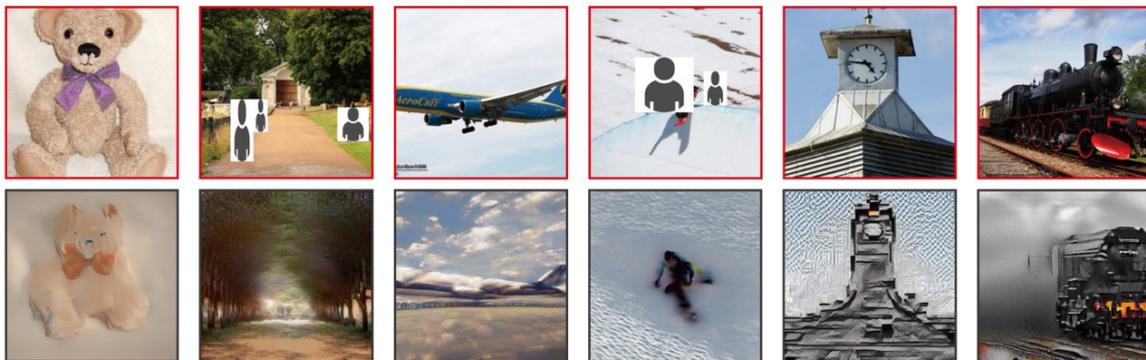


Figura 23: Na linha de cima, a imagem que o paciente estava vendo. Na linha de baixo, a imagem “adivinhada” pelo computador.

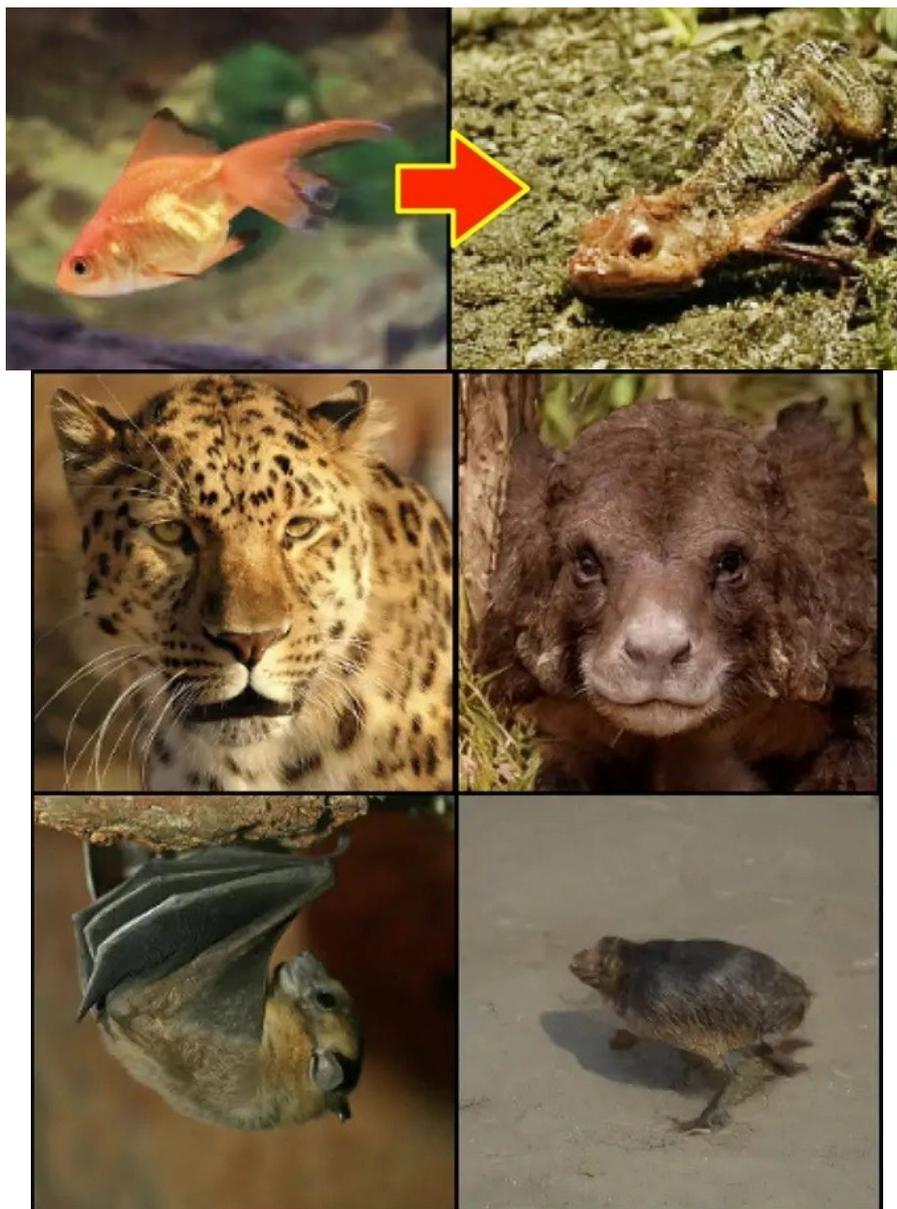


Figura 24: Esquerda, a imagem que o paciente estava vendo. Direita, a imagem “adivinhada” pelo computador.

## 7) Deep fake

“Deep fake” usa IA para trocar o rosto de pessoas em vídeos, sincronizar movimentos labiais, expressões e demais detalhes.

Há dois anos atrás, aparecia notícias sobre um vídeo “deep fake” de Zelenski, presidente da Ucrânia, pedindo que os soldados e a população se rendam aos russos. Vídeos “deep fake” são criados com técnicas de aprendizado profundo.



Figura 25: Deep fake usa IA para trocar o rosto de pessoas.

## 8) IA gera imagens e vídeos a partir da descrição em português/inglês

Pedindo para <https://labs.openai.com/> em português:

“Cachorro com óculos espelhado andando de moto na avenida paulista, digital art.”

Obtive as imagens da figura 26. Isto mostra que o sistema entendeu o pedido e tentou atendê-lo.



Figura 26: Imagens geradas pela frase “cachorro com óculos espelhado andando de moto na avenida paulista, digital art.”

