

# Relatório LAMIA 19

## Vídeo: Como Estamos Ensinando os Computadores a Compreender Fotos (III)

Kaique Medeiros Lima

### 1 Introdução

Esse card apresenta o vídeo "*How we teach computers to understand pictures*", que irá explorar o avanço da visão computacional e o uso de inteligência artificial para ensinar máquinas a interpretar imagens.

### 2 Descrição da atividade

#### 2.1 Vídeo – How we teach computers to understand pictures

A palestrante apresenta os avanços da visão computacional, uma área da inteligência artificial que busca ensinar máquinas a interpretar imagens como humanos. Ela explica que, enquanto a sociedade desenvolveu tecnologias avançadas, como carros autônomos e drones, as máquinas ainda têm dificuldade na identificação de conteúdo em imagens, algo que uma criança já consegue realizar.

Para enfrentar esse desafio, ela liderou o projeto ImageNet, que buscou replicar a maneira como crianças aprendem através de exemplos visuais. A equipe coletou e classificou 15 milhões de imagens organizadas em 22 mil categorias. Com a ajuda de milhares de trabalhadores pelo Amazon Mechanical Turk, construíram um banco de dados enorme que permitiu treinar algoritmos de deep-learning especialmente CNNs, para reconhecer objetos em fotos com mais precisão.

Esses algoritmos evoluíram para identificar não apenas objetos isolados, mas também cenas completas, relacionando elementos visuais a descrições em linguagem natural. São mostrados exemplos de como os computadores são capazes de descrever fotos com frases simples, como "um homem está perto de um elefante". Apesar dos avanços, os modelos ainda cometem erros, confundindo objetos e não entendendo contextos mais complexos ou emoções humanas.

A visão computacional já começa a gerar impactos sociais positivos. Com máquinas que enxergam, hospitais podem melhorar diagnósticos, carros autônomos se tornam mais seguros e robôs podem ajudar em situações de resgate. A colaboração entre humanos e máquinas permitirá novas descobertas, desde a exploração de territórios desconhecidos até o uso em pesquisas científicas e sociais.

Por fim, ela reflete sobre o futuro em que as máquinas terão visão para colaborar com humanos em tarefas que exigem percepção visual. Sua missão é tornar essa tecnologia acessível, criando um mundo onde máquinas ajudem pessoas a ver melhor e entender o

mundo com mais profundidade, trazendo benefícios para a sociedade e para as próximas gerações.

### 3 Conclusão

Nesse card, é apresentado o vídeo "*How we teach computers to understand pictures*", o qual explora os avanços da visão computacional e o uso de inteligência artificial para ensinar máquinas a interpretar imagens. A palestrante destaca seu projeto ImageNet, que reuniu milhões de imagens para treinar algoritmos de redes neurais convolucionais (CNNs) no reconhecimento de objetos. A tecnologia já vem trazendo benefícios, como diagnósticos médicos mais precisos e carros autônomos mais seguros. A colaboração entre humanos e máquinas promete impulsionar descobertas e gerar impactos positivos na sociedade.