

Buscador de imagens - Parte 1

Visão Computacional - 2018/2

Igor Montagner, Fábio Ayres, Raul Ikeda, Andrew Kurauchi

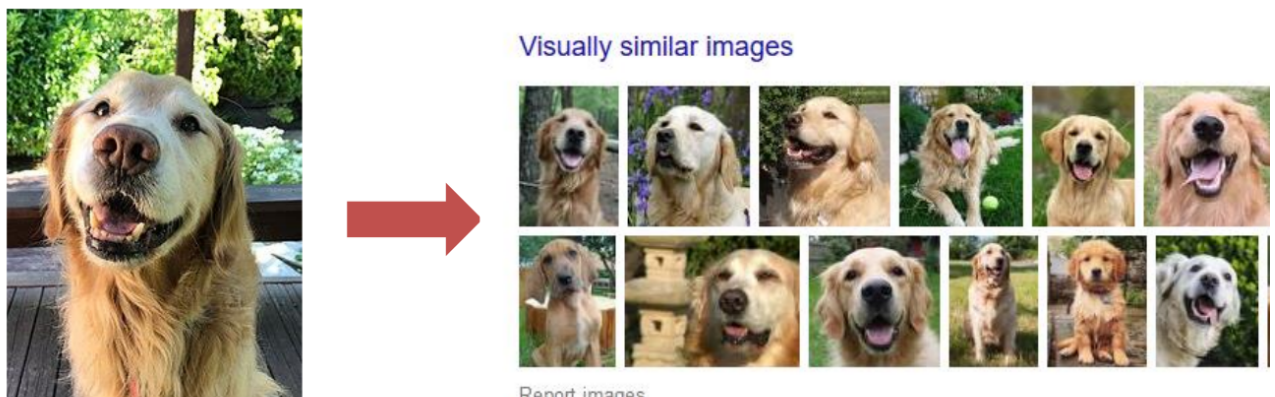
Aplicativos como o Google Photos e o iCloud oferecem funções que organizam a galeria de fotos usando critérios como similaridade e presença de pessoas, locais ou objetos nas imagens. Estas funções são implementadas usando técnicas de Visão Computacional que visam adicionar **semântica** às imagens. Ou seja, além de identificarmos características de baixo nível em imagens (como bordas ou segmentações por cores) essas técnicas também procuram adicionar **significado** às imagens. Por exemplo, podemos criar um programa que responde a perguntas como “A imagem apresentada contém um gato?”.



Figure 1: A imagem acima contém um gato? Fazer perguntas sobre o conteúdo de imagens é o tema deste módulo.

Similaridade entre imagens

Na primeira parte deste projeto focaremos em um problema de *Recuperação de Informação*: a busca por similaridade de imagens. Neste tipo de busca o usuário envia uma imagem e o sistema retorna imagens visualmente parecidas por ordem de similaridade.



{400px}

Seu programa deverá receber uma imagem como entrada, pesquisar em um banco de imagens próprio e retornar uma lista com as 5 imagens visualmente mais parecidas. De preferência seu programa deve ser capaz de fazê-lo

em menos de 1 segundo para milhares de imagens, de modo a tornar a experiência parecida com a obtida em um sistema de busca real.

Restrições tecnológicas

Apesar do projeto ser focado em construir um sistema de busca, faremos algumas limitações em relação às ferramentas usadas no projeto.

1. Se você já implementou uma técnica básica na atividade e agora deseja usá-la no projeto, você pode usar a função pronta do OpenCV;
2. Bibliotecas de aprendizado de máquina estão liberadas desde que elas resolvam somente um problema de aprendizado de máquina.
3. A não ser que seja explicitamente permitido, você não deve usar modelos pré-treinados que já fazem todo o pipeline de reconhecimento/localização/classificação de objetos em imagens.
4. O conceito **C** pode ser obtido combinando ferramentas do OpenCV, mas os conceitos mais avançados pedirão um esforço de implementação própria maior.

Na dúvida pergunte ao professor.

Entregas

Você poderá fazer duas entregas desta parte. A primeira entrega será para o dia 15/10 e deve ser acompanhada das atividades 1 e 2 desenvolvidas em sala de aula. Se você alcançar a rubrica **C** poderá fazer uma segunda entrega no dia 31/10 junto com a parte 2 do projeto.

Avaliação

Cada parte do projeto será avaliada de maneira independente e receberá um conceito entre **I** e **A** dependendo da sofisticação da solução implementada. Para ser aprovado é necessário obter conceito *ge* **C** em ambas as partes.

Importante: entregar uma documentação explicando o funcionamento do software é obrigatório e não fazê-lo resulta em conceito **D**. Seu texto deverá explicar como montar o banco de imagens, treinar o(s) modelo(s) e realizar buscas.

- Conceito **D**:
 - o software tem algum erro conceitual grave **OU**;
 - o software não funciona para uso basico **OU**
 - o software não está acompanhado de documentação explicando seu funcionamento.
- Conceito **C**:
 - o software é capaz de calcular a similaridade de imagens usando o modelo Bag of Visual Words e faz buscas em um conjunto de imagens no disco;
 - o software **NÃO** faz indexação nas imagens, calculando as representações das imagens a cada busca;
- Conceito **B**:
 - o software faz a indexação das imagens e guarda os resultados em um arquivo no disco.
 - o software está arquitetado corretamente e só faz as operações de treinamento do modelo quando solicitado;
 - o software implementa mais de uma métrica de similaridade e permite que o usuário escolha qual métrica será usada;
 - o software é acompanhado de um relatório mostrando seu desempenho em um dos datasets padrão usados no curso e comparando os resultados obtidos com outros da literatura.
- Conceito **A**:
 - o software é capaz de calcular similaridade usando uma rede pré-treinada de sua escolha;
 - o relatório do conceito **B** compara os dois tipos de representações de imagens extraídas e seu desempenho nos datasets da literatura;
 - o software armazena as imagens e suas representações em um banco de dados
 - o software permite adicionar imagens no banco de dados e retreinar o modelo de maneira simples.

Rubrica A+

Se o seu conceito for $\geq \mathbf{B}$ você pode investir na criação de uma interface gráfica para seu sistema de buscas. Você pode criar um sistema web, uma aplicação desktop ou mobile. Este item valerá no máximo 1,0 ponto.

Trabalhos em duplas

O projeto poderá ser feito em duplas, mas seguirá regras de aprovação diferentes: para ser aprovada a dupla precisa entregar um trabalho que obtenha ao menos o conceito **B** e precisa fazer a rubrica A+ detalhada acima.