

Processamento de Imagens utilizando técnicas de Aprendizado de Máquina

Aluno: Heitor Barroso Cavalcante

Orientadora: Nina S. T. Hirata

22 de maio de 2023

Resumo

Este texto descreve um projeto de iniciação científica proposto no contexto do Programa PICME-USP. O estudante procurou a orientadora, do Departamento de Ciência da Computação do IME-USP, interessado em reconhecimento de padrões em plantas usando Visão Computacional. O texto apresenta inicialmente algumas considerações sobre reconhecimento de padrões em imagens e fundamentos necessários para a realização de tal tarefa por meio de Visão Computacional: as áreas de conhecimento de Processamento de Imagens e de Aprendizado de Máquina (*Machine Learning*). Em seguida descreve os objetivos desta iniciação científica e um plano de trabalho.

1 Introdução

Com o avanço tecnológico, câmeras digitais tornaram-se bastante acessíveis, facilitando a aquisição de imagens no dia-a-dia. Com isso, surge também o interesse em organizar essas imagens ou extrair delas informações que possam ser úteis para algum fim.

Com respeito à extração de informação das imagens, dentre as várias possibilidades, podemos citar por exemplo o reconhecimento de espécies de plantas, ou ainda o reconhecimento de padrões tais como alterações de cores em folhas que podem indicar a presença de alguma doença.

Um sistema computacional para esse tipo de problema deve ser capaz de “ler” uma imagem e construir uma representação que possa ser mapeada para a informação de interesse. Por exemplo, se o objetivo é verificar se a coloração das folhas está fora do padrão esperado, o sistema deve segmentar a região na imagem que corresponde às folhas, e em seguida extrair as informações de cor dessa região, para finalmente compará-las com o padrão esperado. Soluções desse tipo enquadram-se na área de Visão Computacional.

O desenvolvimento de soluções de Visão Computacional [Sze11] requer conhecimentos sobre representação de imagens e processamento de imagens [GW02]. Ao longo dos anos, várias representações e famílias de processamentos foram desenvolvidas. Soluções clássicas em geral consistem em construir *pipelines* de processamento, isto é, uma sequência de transformações da imagem original até a informação final esperada. A criação desses *pipelines* requer não só conhecimentos sólidos sobre processamento de imagens, mas também um esforço considerável de tentativa-e-erro para o ajuste da sequência de processamentos, assim como dos parâmetros associados a cada processamento no *pipeline*.

A automatização de partes de *pipelines*, principalmente as relacionadas ao ajuste de parâmetros, foi uma evolução natural no desenvolvimento da área de Visão Computacional.

Para tanto, passaram a ser cada vez mais empregadas as técnicas de Aprendizado de Máquina.

Atualmente vários problemas na área de Visão Computacional tais como classificação de imagens, segmentação de imagens e detecção de objetos em imagens, são realizados de ponta-a-ponta por meio de redes neurais artificiais, especificamente as *deep neural networks*.

Portanto, o desenvolvimento de sistemas de Visão Computacional moderno requerem não somente o conhecimento de processamento de imagens, mas também de aprendizado de máquina.

2 Objetivos

Os objetivos deste projeto de iniciação científica são:

1. Estudo de fundamentos teóricos e práticos de Processamento de Imagens
2. Estudo de fundamentos teóricos e práticos de Aprendizado de Máquina
3. Aplicação dos conhecimentos adquiridos no desenvolvimento de soluções de Visão Computacional, bem como a avaliação e melhorias subsequentes da solução.

3 Atividades a serem desenvolvidas

Para cada um dos objetivos listados acima, detalhamos as atividades planejadas.

3.1 Fundamentos de Processamento de Imagens

Em um momento inicial será buscada uma familiarização com a área de processamento de imagens. As atividades planejadas envolvem o estudo de tutoriais que introduzem conceitos, terminologias e definições, e simultaneamente a implementação e experimentação de algoritmos simples para manipulação de imagens. Em seguida, alguns problemas específicos de processamento de imagens serão escolhidos e soluções existentes serão estudadas e implementadas. Ao final deste processo, espera-se que o estudante tenha compreendido como as imagens são representadas em computador e como elas são processadas, e tenha experimentado alguns exemplos de *pipelines* simples de processamento.

Especificamente, a familiarização inicial será baseada em

- Leitura e estudo do tutorial disponível em <https://www.tutorialspoint.com/dip/index.htm>
- Atividade prática de implementação, baseada no material [Image Processing Using Numpy: With Practical Implementation And Code](#)

Os problemas de processamento tais como binarização de imagens ou segmentação de objetos em imagens servirão para experimentar o ajuste de parâmetros em algoritmos de processamento de imagens e a construção de *pipelines* de processamentos. Para essa atividade poderão ser usadas as bibliotecas de processamento de imagens tais como [scikit-image](#) ou [OpenCV](#). Além disso, à medida que soluções estudadas envolvam algoritmos ou conceitos novos, os mesmos serão estudados e discutidos. Algumas referências bibliográficas adequadas para esses estudos são [GW02, Pri12, Sze11].

3.2 Fundamentos de Aprendizado de Máquina

Aprendizado de Máquina [AMLM12, Mur22] refere-se a uma área de estudo no qual parâmetros de modelos genéricos de processamento (família de funções que mapeiam um certo tipo de entrada para outro certo tipo de saída) são ajustados a partir de exemplos entrada-saída.

Dois tipos de problemas comumente tratados em aprendizado de máquina são a classificação e a regressão; no primeiro a saída-alvo é um valor que indica a categoria (classe) ao qual a entrada pertence (tais como a espécie de uma planta), e no segundo a saída-alvo é um valor contínuo (tais como temperatura, peso, etc).

Tipicamente assume-se que existe uma relação entre a entrada e saída e que a mesma pode ser expressa por uma função. Para o ajuste dos parâmetros, em geral busca-se a minimização de uma função de perda, isto é, o ajuste é feito de forma que a diferença entre o resultado produzido pelo modelo sendo ajustado e a saída-alvo esperada seja o menor possível. Esse ajuste de parâmetros é realizado num processo conhecido por treinamento e tanto o treinamento como o resultado podem ser afetados pela qualidade, características e quantidade de dados, assim como a família de funções utilizada para aproximar a relação entrada-saída.

O estudo de fundamentos irá cobrir os conceitos mencionados acima, alguns algoritmos bem conhecidos, assim como estratégias comumente empregadas para se alcançar um treinamento bem-sucedido. Será dado foco ao estudo das redes neurais artificiais [Nie15] pois elas constituem a base dos modernos algoritmos de *deep learning* [GBC16] utilizados na área de Visão Computacional.

Especificamente,

- Os fundamentos serão estudados utilizando a referência [AMLM12], assim com o site correspondente [Learning from Data \(Caltech\)](#)
- A parte prática poderá envolver inicialmente a implementação de algoritmos básicos de aprendizado de máquina, porém deverá em sua maior parte ser realizada utilizando-se bibliotecas como [scikit-learn](#), [Keras/TensorFlow](#) ou [PyTorch](#).
- Os problemas a serem abordados envolverão dados do tipo imagem.

3.3 Aplicações

Pretendemos explorar problemas relacionados ao processamento de imagens de plantas. Uma possibilidade é a utilização de conjunto de imagens publicamente disponíveis, muitas vezes já prontas para serem utilizadas por algoritmos de aprendizado de máquina. Alguns exemplos de *datasets* são:

- <https://www.kaggle.com/kaustubhb999/tomatoleaf>
- <https://data.mendeley.com/datasets/hb74ynkjc/4>
- <https://www.plant-phenotyping.org/datasets-home>

Dependendo do andamento, podemos também considerar a aquisição de imagens, por meio de celulares, ou por meio de *webcams* instaladas em um ambiente controlado, para por exemplo acompanhar o crescimento de uma planta.

3.4 Cronograma aproximado

O planejamento acima foi feito considerando-se uma duração de 2 anos. O primeiro ano será dedicado ao estudo dos fundamentos, tanto de processamento de imagens como de aprendizado de máquina, sob a perspectiva teórica e prática. Ao final do ano é esperado que o estudante seja capaz de desenvolver soluções baseadas em aprendizado de máquina para alguns tipos de problemas, tais como classificação de imagens. É importante salientar que hoje em dia é prática comum os códigos computacionais estarem disponíveis publicamente e que, em princípio, é possível reproduzir os resultados ou adaptar a solução para problemas similares mesmo sem se ter quase nenhum conhecimento sobre as técnicas empregadas. Porém, nesta iniciação científica almejamos que o estudante alcance ao final do primeiro ano de estudo conhecimentos suficientes para entender o processamento sendo realizado, avaliar o desempenho das soluções e, principalmente, discutir criticamente deficiências ou limitações das soluções desenvolvidas e de seu próprio conhecimento.

O segundo ano será dedicado para o desenvolvimento da aplicação descrita acima. Naturalmente, a complexidade do problema a ser tratado será maior, e também servirão como oportunidade para o exercício de práticas relacionadas à metodologia científica tais como caracterização do problema e questões de pesquisa, planejamento de experimentos, comparação de métodos, análise de resultados e escrita científica.

Referências

- [AMLMI12] Yaser S. Abu-Mostafa, Hsuan-Tien Lin, and Malik Magdon-Ismael. *Learning From Data*. AMLBook, 2012.
- [GBC16] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [GW02] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Addison-Wesley Publishing Company, second edition, 2002.
- [Mur22] Kevin P. Murphy. *Probabilistic Machine Learning: An introduction*. MIT Press, 2022.
- [Nie15] Michael A. Nielsen. *Neural Networks and Deep Learning*. Determination Press, 2015.
- [Pri12] Simon D. J. Prince. *Computer Vision – Models, Learning and Inference*. Cambridge, 2012.
- [Sze11] R. Szeliski. *Computer Vision – Algorithms and Applications*. Springer, 2011.