

A era digital e suas implicações sociais: Desafios e contribuições

ESTUDO SOBRE UMA PIPELINE GENÉRICA PARA SEGMENTAÇÃO DE IMAGENS DIGITAIS

**Fabio Nikaido, Felipe Alexandre, Rodrigo Azevedo, Matheus Mascarenhas,
Amita Muralikrishina, Carlos Eduardo Oliveira da Silva.**

Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, Rodovia Presidente Dutra Km
145 Jardim Diamante - 12223-201 - São José dos Campos-SP, Brasil,
fabio.nikaido@aluno.ifsp.edu.br, alexandre.felipe@aluno.ifsp.edu.br,
rodrigo.braz@aluno.ifsp.edu.br, mascarenhas@ifsp.edu.br, amita@ifsp.edu.br,
carlossilva@ifsp.edu.br.

Resumo

Atualmente as técnicas de aprendizado de máquinas vêm sendo amplamente aplicadas em diversos campos do conhecimento e o processamento de imagem é uma dessas áreas que vêm se destacando. A *pipeline* de aplicação de uma técnica de aprendizado de máquina voltada a imagens necessita de uma etapa de pré-processamento, sendo uma das partes mais importantes, otimizando o processamento ou até mesmo viabilizando-o. Uma imagem se trata de um conjunto de dados em formato de matriz com três dimensões, logo trabalhar com imagens significa trabalhar com grandes conjuntos de dados e a etapa de pré-processamento pode ser utilizada para se separar as regiões de interesse, as quais podem conter dados relevantes para o treinamentos, dos componentes ao fundo, otimizando a seleção de dados. Neste trabalho será feito um estudo sobre a etapa de pré-processamento dentro de uma *pipeline* de aprendizado de máquina voltada a visão computacional.

Palavras-chave: Visão computacional, aprendizado de máquina, processamento de imagens.

Área do Conhecimento: Engenharias - Ciência da Informação.

Introdução

O pré-processamento de dados se mostra uma etapa importante dentro de um *pipeline* de machine learning (ML) voltado a visão computacional (VC). Na maioria dos casos os dados recolhidos apresentam algum tipo de ruído ou distorção, principalmente quando trabalha-se com imagens, sendo que a ferramenta de captura influencia diretamente na qualidade, pois ruídos, desfoques ou mudanças repentinas das cores podem atrapalhar as etapas de segmentação e classificação de objetos (COSTA, 2022).

O processo de segmentação se trata da seleção de um elemento dentro de imagens digitais cujos dados têm relevância, sendo esse passo de suma importância para reconhecimento de padrões na VC. O elemento relevante em uma imagem digital tem o nome de região de interesse (ROI, do inglês Region of Interest). O objetivo da aplicação de técnicas de pré-processamento é a melhoria dos dados, suprimindo ruídos indesejados ou melhorando características chave para futuras etapas em um processo de segmentação. Neste trabalho serão estudadas técnicas de pré-processamento voltadas a imagens digitais demonstrando uma *pipeline* para a segmentação de imagens.

Metodologia

A aquisição de imagem consiste em técnicas de captura de imagens, sendo a primeira etapa do processamento de imagem, visando a captura de imagens e sua digitalização para uma imagem digital, a fim de encontrar a melhor forma de se extrair as informações relevantes de uma imagem. Nesta etapa, diversos fatores podem afetar a qualidade da captura de imagem, tais como, iluminação, ruídos e distorções, podendo afetar as demais etapas do processo (GONZALEZ e WINTZ, 1987).

A suavização e redução de ruído das imagens são técnicas importantes para a área de processamento de imagem. Segundo Fan (2019), essas técnicas consistem em remover os ruídos

A era digital e suas implicações sociais: Desafios e contribuições

presentes na imagem capturada, a fim de restaurá-las ao seu formato original, tendo em vista que essas perturbações podem afetar a qualidade da imagem e como consequência, a sua análise.

As técnicas de *blur* são utilizadas para executar essa suavização e redução de ruídos na imagem. Essa ferramenta consiste em desfocar a imagem, reduzindo sua nitidez e detalhes, melhorando a qualidade visual da imagem por meio da remoção de ruídos e artefatos indesejáveis.

Existem diversas técnicas de *blur*, tais como, *averaging*, *gaussian*, *blurring* e filtro de média. Neste artigo iremos aprofundar nos conceitos de filtro de média, tendo em vista ser um dos mais utilizados na etapa de suavização de imagens.

O filtro de média consiste em substituir cada *pixel* da imagem pela média dos valores dos *pixels* vizinhos dentro de uma área específica. Essa área pode ter diversos tamanhos e formas, que variam em relação à distância e ao peso dos *pixels* vizinhos (SANCHES, et al., 2015).

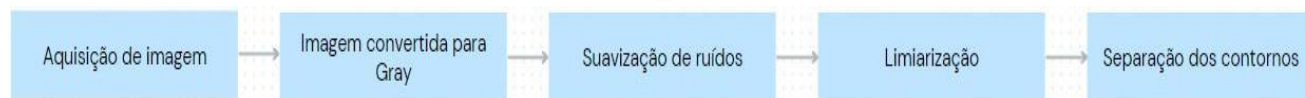
Limiarização é um procedimento de segmentação de imagens que consiste em separar os *pixels* de uma imagem em dois grupos: os que pertencem ao objeto de interesse (*foreground*) e os que pertencem ao fundo (*background*). Para isso, utiliza-se um valor de limiar (*threshold*) que define a intensidade de cinza que separa os dois grupos. Existem diferentes métodos para determinar o valor de limiar, dependendo das características da imagem e do objetivo da segmentação (GONZALEZ e WINTZ, 2018).

Alguns tipos de limiarização incluem a limiarização global simples, em que é utilizado um valor T escolhido pelo usuário para que a imagem seja segmentada, sendo esta dividida em $G1$ e $G2$, sendo $G1$ com *pixels* com intensidade maior que T e $G2$ com *pixels* com intensidade menor ou igual a T (SONKA, HLAVAC, BOYLE, 2014). Outro tipo de limiarização é a adaptativa, onde o valor da limiar é definido apenas para pequenas áreas da imagem ao invés de um valor único, funcionando melhor para imagens que possuem diferentes iluminações comparado ao método global simples e obtendo vários *thresholds* (BRADSKI, 2000).

O método de Otsu possui a vantagem de automaticamente determinar *thresholds* em vez de depender de um valor previamente definido pelo usuário. Através de um histograma gerado pela imagem, o computador consegue gerar um valor de limiar otimizado ao escolher um valor no vale do histograma, minimizando a variância dos níveis de cinza e deixando os grupos de *pixels* mais homogêneos (OTSU, 1979).

O processo de segmentação segue uma linha de raciocínio começando pela aquisição das imagens. Nesta etapa o instrumento de captura, que pode ser uma câmera digital ou uma câmera de celular sendo que fatores como a qualidade e resolução das imagens são definidas pelo dispositivo de captura. A etapa de suavização existe para se diminuir os efeitos do ruído associado às imagens, na etapa de limiarização é feito a separação dos contornos transformando as imagens originais em imagens binárias, por fim a ROI das imagens é separada. O fluxograma da Figura 1 mostra as etapas de uma segmentação.

Figura 1 - Etapas do processo de seleção de objetos dentro de uma imagem.



Fonte: o autor

A era digital e suas implicações sociais: Desafios e contribuições

Resultados

A *pipeline* apresentada pela Figura 2 mostra o quanto o processamento de imagens consegue selecionar vários objetos de uma vez separando informações pertinentes da ROI em relação às informações menos pertinentes classificadas como fundo. A Figura 2 mostra o processo descrito na Figura 1 aplicado a uma imagem de componentes eletrônicos. Da esquerda para a direita se tem a imagem original, filtro *gray*, filtro de suavização, no caso o aplicado foi o filtro de média, limiarização aplicando o método de Otsu e por fim os contornos selecionados são desenhados na imagem original. Todo processo foi aplicado utilizando Python com a biblioteca OpenCV.

Figura 2 - Processo de seleção de contornos da esquerda para direita, imagem original, filtro *gray*, filtro de média, limiarização e imagem original com os contornos desenhados.



Fonte: autor.

Discussão

Como pode ser visto na Figura 2, os componentes são diferenciados do fundo branco pelo algoritmo, separando informações pertinentes dos componentes, como sua posição e sua forma, do fundo - nesse caso, a parte branca da imagem. Essa separação se mostra muito útil porque, dentro de um processo de coleta de grandes dados, as informações importantes são separadas, poupando processamento e armazenamento devido ao fato de que uma imagem é uma matriz, logo, quanto menores e mais significativos os dados a serem analisados, melhor.

Na imagem original, 57 componentes no total foram distribuídos e 53 componentes foram detectados. Apesar disso, 6 componentes não foram detectados. Isso se deve ao fato de um componente ter sido detectado como três contornos diferentes no resultado final. Considerando o resultado final como 51 componentes devido ao fato de um único componente ter sido contado como três, temos uma porcentagem de detecção por volta de 89%.

Analisando mais a fundo os resultados obtidos, foi possível perceber alguns pontos nos quais o algoritmo não selecionou os componentes - por exemplo, os três LEDs, nos quais somente as partes com cores vermelhas foram identificadas. Isso acontece porque o algoritmo trabalha com o contraste existente entre o componente e o fundo, ignorando os segmentos nos quais as cores são semelhantes.

O filtro de média, ao suavizar a imagem, acabou perdendo detalhes. O resultado disso foi a melhora da seleção dos contornos dos componentes, entretanto, alguns componentes pequenos foram perdidos durante o processo.

A Figura 3 demonstra um exemplo de uma imagem cuja seleção de contornos se mostrou ineficiente. Os contornos selecionados revelam o contraste da iluminação, não os objetos em si, como por exemplo, traços do rosto, chapéu ou fundo. Isso mostra a interferência da iluminação na seleção dos contornos, logo o controle de iluminação é importante para a etapa de pré-processamento.

A era digital e suas implicações sociais: Desafios e contribuições

Figura 3 - A figura mostra a *pipeline* do processo aplicada à imagem.



Fonte: autor.

Conclusão

O trabalho mostra que a *pipeline* aplicada entrega resultados efetivos para um processo de seleção de ROI. O processamento de imagem é uma tecnologia que se encontra em crescimento e técnicas aplicadas a pré-processamento de imagens são imprescindíveis para melhoria do processo que utiliza essas imagens, sendo ele *machine learning* ou *data mining*, pois tais processos melhoram a seleção de contornos, melhorando a qualidade das informações coletadas e diminuindo as dimensões das matrizes imagens selecionadas, por conseguinte a capacidade de processamento necessária para se interpretar os dados.

Por meio dos resultados obtidos foi possível realizar todas as etapas do processamento por meio da biblioteca OpenCV, na qual foram selecionados os contornos em uma imagem de componentes eletrônicos separando assim as informações pertinentes, que seriam os componentes, do fundo da imagem.

A *pipeline* demonstrou resultados satisfatórios para o problema proposto, na qual foi possível comparar a imagem original e processada. Sendo assim, o método apresentado neste trabalho pode ser exportado e utilizado em outros processos.

Referências

BRADSKI, Gary et al. **OpenCV. Dr. Dobb's journal of software tools**, v. 3, n. 2, 2000.

COSTA, Victor L. et al. **Pré-processamento de imagens de baixa resolução utilizando Deep Learning baseado em um Autoencoder**, 2022.

FAN, Linwei et al. **Brief review of image denoising techniques. Visual Computing for Industry, Biomedicine, and Art**, v. 2, n. 1, p. 1-12, 2019.

GONZALEZ, Rafael C.; WINTZ, Paul. **Digital image processing**. Addison-Wesley Longman Publishing Com., Inc., 1987.

OTSU, Nobuyuki. **A threshold selection method from gray-level histograms. IEEE transactions on systems, man, and cybernetics**, v. 9, n. 1, p. 62-66, 1979.

SANCHES, Carlos H. et al. **Técnicas de suavização de imagens e eliminação de ruídos. Anais do EATI-Encontro Anual de Tecnologia da Informação. Frederico Westphalen-RS**, p. 21-30, 2015.

SONKA, Milan; HLAVAC, Vaclav; BOYLE, Roger. **Image processing, analysis, and machine vision**. Cengage Learning, 2014.