

Algoritmo Viola-Jones na detecção e reconhecimento facial

Drielle Viana Vieira

Resumo—As técnicas de detecção e reconhecimento facial têm chamado atenção ao longo dos anos, como uma alternativa de segurança para pessoas e até mesmo empresas. Com o passar do tempo foram criadas bibliotecas específicas em visão computacional voltadas para a detecção e o reconhecimento de faces. Nesse resumo será citado, em especial, o algoritmo de Viola-Jones e seu uso para esse propósito. O programa será executado na Raspberry Pi e as faces serão detectadas pela Pi Camera.

Palavras-chave: Reconhecimento facial, OpenCV, Raspberry Pi, Picamera, Programação, Detecção de faces, Python.

I. INTRODUÇÃO

Seres humanos são capazes de identificar faces e emoções desde bebês. O desenvolvimento de modelos computacionais capazes de imitar essa capacidade humana têm sido de interesse de muitas pessoas para aplicações em diversos ramos, como proteção de dados, identificação de criminosos, biometria e interação humano-máquina, por exemplo. O desenvolvimento de bibliotecas computacionais como o OpenCV para Python, junto com a evolução das câmeras e os hardwares de alto desempenho, tornou capaz essa aproximação entre o reconhecimento humano e o reconhecimento pelo computador.

A. Objetivos

O objetivo do trabalho estudado foi fornecer uma base teórica e prática de como construir um sistema de detecção e reconhecimento facial que será utilizado em estabelecimentos comerciais. O sistema de detecção será executado em uma Raspberry Pi, instalada em um ponto estratégico do estabelecimento. As capturas de imagens serão feitas continuamente e o sistema deverá identificar em quais imagens contém faces para depois fazer a extração das faces e enviar para o servidor onde os algoritmos de reconhecimento facial serão executados.

II. ALGORITMO VIOLA-JONES

O algoritmo de Viola-Jones é uma estrutura de detecção de objetos que foi proposta por Paul Viola e Michael Jones em 2001, por mais que esse algoritmo possa ser treinado para a detecção de diversos objetos, sua criação foi motivada para resolver problemas com detecções de faces humanas. Como dito anteriormente, o ser humano é capaz de detectar rostos desde bebê, porém, um computador necessita de instruções precisas para como fazê-lo. Viola-Jones propôs uma visão completa de faces verticais, ou seja, todo o rosto deve apontar para a câmera e não deve estar inclinado para nenhum dos lados. Encontra-se implementado na biblioteca OpenCV e pode ser utilizado para detectar outras partes do corpo além do rosto. O treino do algoritmo é um processo lento e demorado, porém,

a detecção é bastante rápida. O detector Viola-Jones possui quatro estágios: seleção de características Haar, criação de uma imagem integral, treinamento AdaBoost e os classificadores em cascata.

A. Etapas

1) *Características Haar-like*: As características Haar-like são recursos de imagens digitais utilizadas para reconhecimento de objetos, tiveram seu uso no primeiro detector facial em tempo real. Viola e Jones adaptaram essas características para um trabalho semelhante, que considera regiões retangulares adjacentes em um local específico de uma subjanela, soma as intensidades dos pixels em cada uma dessas regiões e depois calcula a diferença entre essas somas. Cada característica resulta em um valor único. A utilização de características se faz vantajosa pois o processamento computacional é menor do que se trabalhar pixel a pixel, além de conseguir reconhecer padrões previamente determinados. Na estrutura de detecção de imagens Viola-Jones, as características Haar-like são organizadas em classificadores de cascata para formar um classificador forte.

2) *Criação de uma Imagem Integral*: Proposto por Corw em 1984, o algoritmo de imagem integral calcula rapidamente a soma dos valores em um subconjunto retangular de uma matriz. Ela vem a ser uma tabela bidimensional do tamanho original da imagem, onde cada elemento equivale à soma de todos os níveis de cinza dos pixels à esquerda e acima do pixel atual.

3) *Adaptive Boosting*: Adaptive Boosting ou AdaBoost, foi inventado por Freund e Shapire, e é um método de aprendizado de máquina que consiste em combinar de forma ponderada vários classificadores fracos para obtenção de um classificador forte, podendo ser utilizado para aumentar a performance de outros algoritmos de aprendizagem. Se traduzido ao pé da letra, temos que ele é um estímulo adaptativo, ou seja, as classificações subsequentes feitas são ajustadas e escolhidas de forma favorável as instâncias classificadas negativamente em classificações anteriores. O algoritmo Viola-Jones utiliza uma variação do AdaBoost para selecionar as características e treinar o classificador. O AdaBoost funciona da seguinte forma. Ele chama um classificador fraco repetidas vezes para realizar iterações. Para cada chamada ocorre uma atualização na distribuição de pesos, indicando a importância do exemplo no conjunto de dados que vai ser utilizado para classificação. A cada iteração, cada exemplo classificado incorretamente é aumentado, para que assim, o novo classificador trabalhe com mais amostras.

4) *Classificadores em Cascata*: São classificadores que tendem a direcionar a atenção do algoritmo para locais onde há uma maior suspeita de haver faces, sendo uma concatenação de classificadores mais complexos na forma de cascata. Cada estágio terá um classificador forte construído pelo método AdaBoost, e ele vai servir como um filtro para determinar se uma determinada subjanela não possui ou talvez possua uma face. Uma janela é descartada de imediato se falhar em qualquer um dos estágios. A ideia chave em classificadores em cascata é a de colocar um classificador com uma alta taxa de detecção no início da cascata para que tenha uma rejeição de janelas negativas logo de início, evitando assim uma sobrecarga de processamento. Quanto mais característica um classificador tiver maior é sua taxa de detecção, menor sua taxa de falso-positivo, porém, seu tempo de cálculo aumenta.

III. DETECÇÃO FACIAL COM OPENCV

Lançada oficialmente pela Intel Corporation em 1999, a biblioteca OpenCV é uma biblioteca multiplataforma de código aberto voltada para aplicações de visão computacional em tempo real em constante desenvolvimento. É toda escrita em C++ porém existem interfaces em diversas outras linguagens de programação, como Python, por exemplo.

A. Treino do Classificador

Para realizar o treino de um classificador é necessário um banco de imagens, sendo esse separado em dois conjuntos de imagens, um conjunto com amostras positivas (imagens que contém faces) e outro um conjunto com amostras negativas (imagens de fundo). Após a aquisição das imagens positivas, é necessário identificar onde as faces estão localizadas. O OpenCV possui uma ferramenta própria que permite marcar com o cursor do mouse o local que contém alguma face. O OpenCV também possui uma ferramenta para o treinamento do classificador. Primeiro é criado um arquivo de vetor, onde você terá a quantidade de exemplos, a largura e a altura como parâmetros daquele vetor. Para treinar o classificador em cascata é utilizado um comando do OpenCV que recebe como parâmetros um diretório onde o classificador será salvo, o arquivo do vetor, a lista de imagens de fundo, e a quantidade de imagens negativas e positivas. Como resultado do treinamento, um arquivo é criado em um diretório, sendo esse arquivo o próprio classificador em cascata já treinado.

IV. RECONHECIMENTO FACIAL

Diferente da detecção, o reconhecimento facial é uma tecnologia capaz de identificar de forma automática uma pessoa através de uma imagem ou de um vídeo. O reconhecimento facial pode ser dividido em três tarefas diferentes. A primeira é a verificação, onde vai ser informado se a face reconhecida pertence ou não ao indivíduo; a segunda tarefa é a identificação, onde se compara a imagem de uma face desconhecida com um banco de dados; e a terceira, mas não menos importante, é o agrupamento, onde faces de uma mesma pessoa vão ser agrupadas mesmo que o indivíduo não esteja no banco de dados. O reconhecimento facial é uma ótima solução

para biometria, por se tratar de um método rápido e não intrusivo. O processo de reconhecimento facial é dividido em três etapas: a detecção da face, a extração de características e a identificação/verificação.

A. Extração de Características

Os métodos para a extração de características consiste em encontrar as informações mais relevantes da imagem original e representá-la em um espaço de menor dimensionalidade.

V. CONCLUSÃO

Mediante a leitura e o embasamento teórico necessário para a pesquisa e a realização desse resumo, foi possível concluir que o algoritmo Viola-Jones é válido para ser usado como um método para detecção e reconhecimento facial quando aliado a bibliotecas, como o OpenCV.

REFERÊNCIAS

- [1] SILVA, Júlio Batista. Explorando o algoritmo de Viola-Jones na detecção e reconhecimento facial. 2018. Tese de Conclusão de Curso (Graduação em Engenharia da Computação) - 2018.
- [2] MARCHIORO, Gessé; ANTONELLO, Ricardo. Desenvolvimento de um sistema de supervisão e controle residencial utilizando processamento de imagens. Semana da Ciência e Tecnologia, [s. l.], 5 jul. 2017. Disponível em: <https://secitec.luzerna.ifc.edu.br/wp-content/uploads/sites/30/2017/03/DOMOTICASECITEC-133.pdf>. Acesso em: 22 set. 2021.
- [3] Porto, Lucas Faria. Uso de Visão Computacional para Interface Homem-Computador. 2011. Tese de Conclusão de Curso (Graduação em Ciência da Computação) - 2011.