

Uso de Deep Learning aplicado no Reconhecimento de ações humanas a partir de vídeos em alta resolução visando identificar movimentos suspeitos

Henrique K. Secchi¹, Silvio A. Carro²

1. Discente do Curso Ciência da Computação – UNOESTE; 2. Docente – UNOESTE – Faculdade de Informática de Presidente Prudente.

E-mails: krupck@outlook.com, silvio@unoeste.br

(Área: A – Ciências Exatas e da Terra)

Introdução

A segurança pública é um tema frequente e recorrente e está em pauta diária na imprensa em geral. Diante disso, este trabalho propõe, por meio da combinação de diferentes algoritmos, o desenvolvimento de um sistema de identificação e rastreamento de pessoas que auxilie a área de segurança.

Para a identificação do indivíduo foi utilizado a CNN YOLO (IAEXPERT, 2020), que está em bastante evidência atualmente, por ser bastante eficiente e com ótima performance nesse contexto. Para o rastreamento foi utilizado outro tipo de algoritmo, o filtro de Kalman (WELCH AND BISHOP, 2006), e mostrou-se bastante eficiente para o rastreamento de objetos em vídeos.

Ainda em andamento, a técnica proposta para a identificação dos movimentos dos indivíduos é o algoritmo *Blazepose*, uma solução de *Machine Learning* capaz identificar pontos de referência do corpo humano (BAZAREVSKY, 2020).

Metodologia

A primeira etapa do algoritmo é submeter o vídeo para que possa ser realizada a identificação, nessa etapa foi utilizado o algoritmo YOLO. O YOLO utiliza uma única rede neural para realizar a classificação e a previsão de caixas delimitadoras para objetos detectados. O algoritmo divide a imagem em um *grid* de $S \times S$ células, onde cada célula dessa *grid* é responsável por fazer a predição de x caixas delimitadoras, caso haja mais de um objeto na mesma célula também é retornado a pontuação de confiança que informa o quanto de certeza o algoritmo tem que a caixa delimitadora contenha um objeto (IAEXPERT, 2020).

O YOLOv3 possui uma versão mais leve que preserva somente as principais camadas, chamada YOLOv3 *tiny*. Esta versão possui a vantagem de performance já que se trata de uma versão com menos camadas, consequentemente com menor processamento. Para este trabalho, foi utilizado o YOLO em sua versão *tiny*, uma vez que, essa arquitetura representou um ganho significativo de performance e o ganho de performance justificou a perda de precisão do modelo.

A partir do objeto detectado, a próxima etapa é realizar o recorte do objeto e selecionar somente o pedaço da imagem que corresponde ao objeto identificado, esse processo se faz necessário para as etapas seguintes, tanto a parte de rastreamento quanto a parte de identificação dos movimentos.

Com a imagem recortada, foi aplicado o algoritmo do filtro de Kalman, para conseguir identificar a variância na matriz de pixels. O Filtro de Kalman consiste em um conjunto de equações que possibilitam a implementação recursiva de um estimador, gerando predição ótima dos estados futuros de um sistema linear a partir de uma observação presente (WELCH AND BISHOP, 2006).

Para a parte de identificação dos movimentos a imagem que foi recortada após a detecção do YOLO é submetida ao algoritmo *Blazepose*, responsável por detectar os pontos de referência do corpo humano. O algoritmo é capaz de identificar 33 pontos do corpo distribuídos nos seguintes locais: nariz, olho direito, olho esquerdo, orelhas, boca, ombro, cotovelo, pulso, dedo mindinho, dedo indicador, polegar, quadril, joelho, tornozelo, calcanhar e pé.

Com o objeto contendo os pontos, e cada ponto contendo a posição X, Y dentro da matriz de pixels, a próxima etapa é calcular os ângulos formados pelos pontos, e dessa forma, identificar qual é a posição de cada ponto em relação a outro, possibilitando identificar qual é o movimento do indivíduo na cena. Caso o movimento esteja dentro das possibilidades de movimentos suspeitos, o algoritmo emitirá um alerta na tela informando que o movimento pode significar um ato criminoso.

Resultados

Para a realização deste trabalho foram estudadas as duas arquiteturas do YOLOv3. A arquitetura original e sua versão reduzida, YOLOv3 *tiny*. Com o intuito de obter a arquitetura com melhor equilíbrio entre desempenho e precisão. Para os testes, foram utilizados dois vídeos, o primeiro contendo uma pessoa na cena com a duração de nove segundos e o segundo vídeo contendo onze pessoas na cena com a duração de treze segundos, em ambos os vídeos a taxa de 30 FPS.

Ao aplicar o filtro de Kalman para realizar o rastreamento dos objetos na cena, a performance do algoritmo caiu significativamente, tornando inviável os testes utilizando a CPU, dessa forma, os testes com as duas arquiteturas foram executados somente com a GPU.

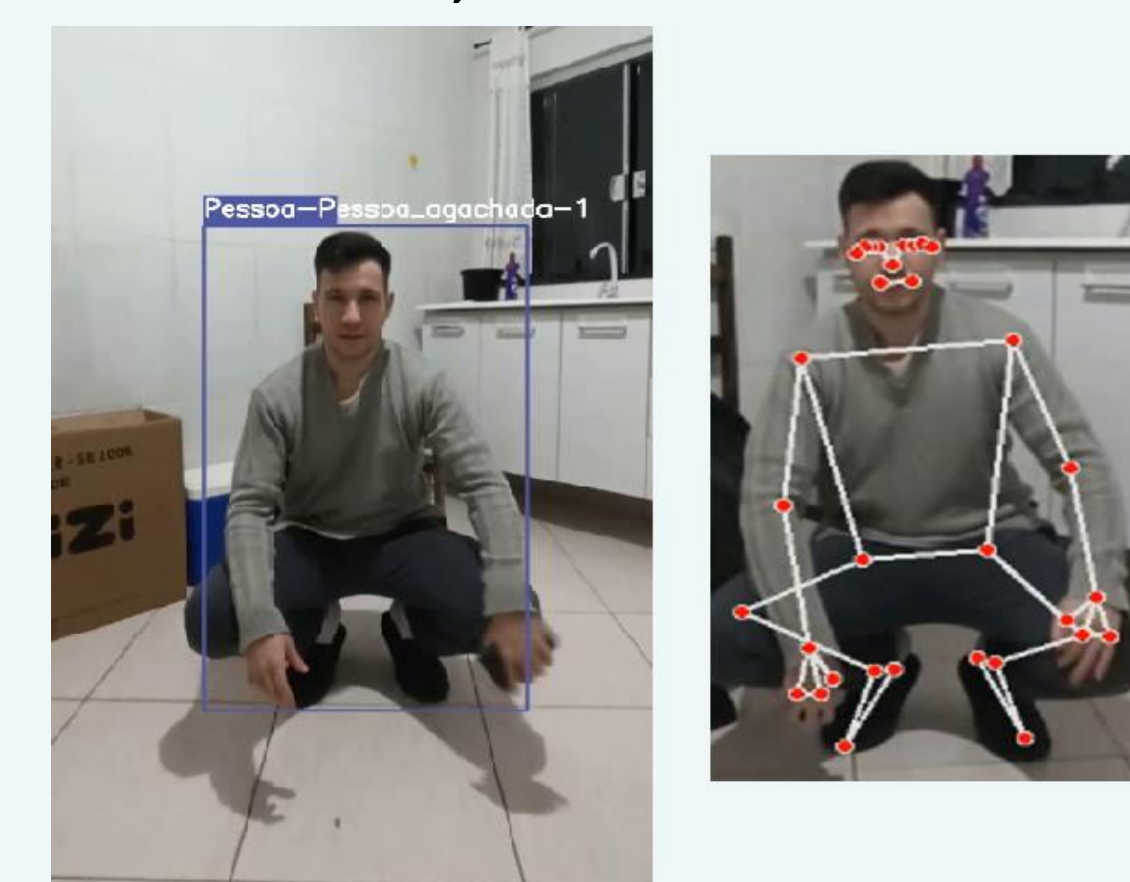
Tabela 1. Testes de performance com o filtro de Kalman.

Arquitetura	Processamento	Resolução	FPS	Precisão
YOLOv3	GPU GTX 1050ti	608x1080	7~12	88% a 95%
YOLOv3 tiny	GPU GTX 1050ti	608x1080	20~24	88% a 95%

Fonte: Realizado pelo autor.

O que se espera para os próximos passos, é unir o algoritmo *Blazepose* de modo a agrega-lo ao trabalho atual mantendo uma boa performance e uma boa precisão nas previsões dos pontos de referência. Ao obter todos os pontos de referência, calcular os ângulos formados por eles e obter o movimento do indivíduo. Na sequência comparar o movimento com a lista de movimentos possíveis, e caso seja um movimento suspeito, emitir um alerta informando que o indivíduo pode estar cometendo um ato criminoso.

Imagem 1. Objeto identificado, rastreado e com os pontos de referência.



Fonte: Realizado pelo autor.

Conclusões

Este trabalho foi motivado pela constante taxa de criminalidade que assola o país, e pelo crescimento de trabalhos utilizando redes neurais convolucionais aplicados nas mais diversas áreas. Acredita-se que o trabalho é relevante, mostrando a possibilidades de utilizar novas combinações, que até então não haviam sido reunidas para essa finalidade. Diferentes algoritmos são utilizados para identificação e rastreamento de pessoa. Os resultados obtidos até o momento nos testes foram considerados satisfatórios, o que justifica ainda mais a relevância do trabalho.

Bibliografia

BAZAREVSKY, V. et al. BlazePose: On-device Real-time Body Pose tracking. Seattle: CVPR Workshop on Computer Vision for Augmented and Virtual Reality, 2020.

IAEXPERT. Detecção de Objetos com YOLO – Uma abordagem moderna. 2020. Disponível em: <https://iaexpert.academy/2020/10/13/deteccao-de-objetos-com-yolo-uma-abordagem-moderna/>. Acesso em: 05 jun. 2021.

WELCH, GREG. AND BISHOP, GARY. An introduction to the kalman filter, Chapel Hill: University of North Carolina at Chapel Hill, 2006.