UNIVERSIDADE REGIONAL DE BLUMENAU

PROCESSAMENTO DE IMAGEM

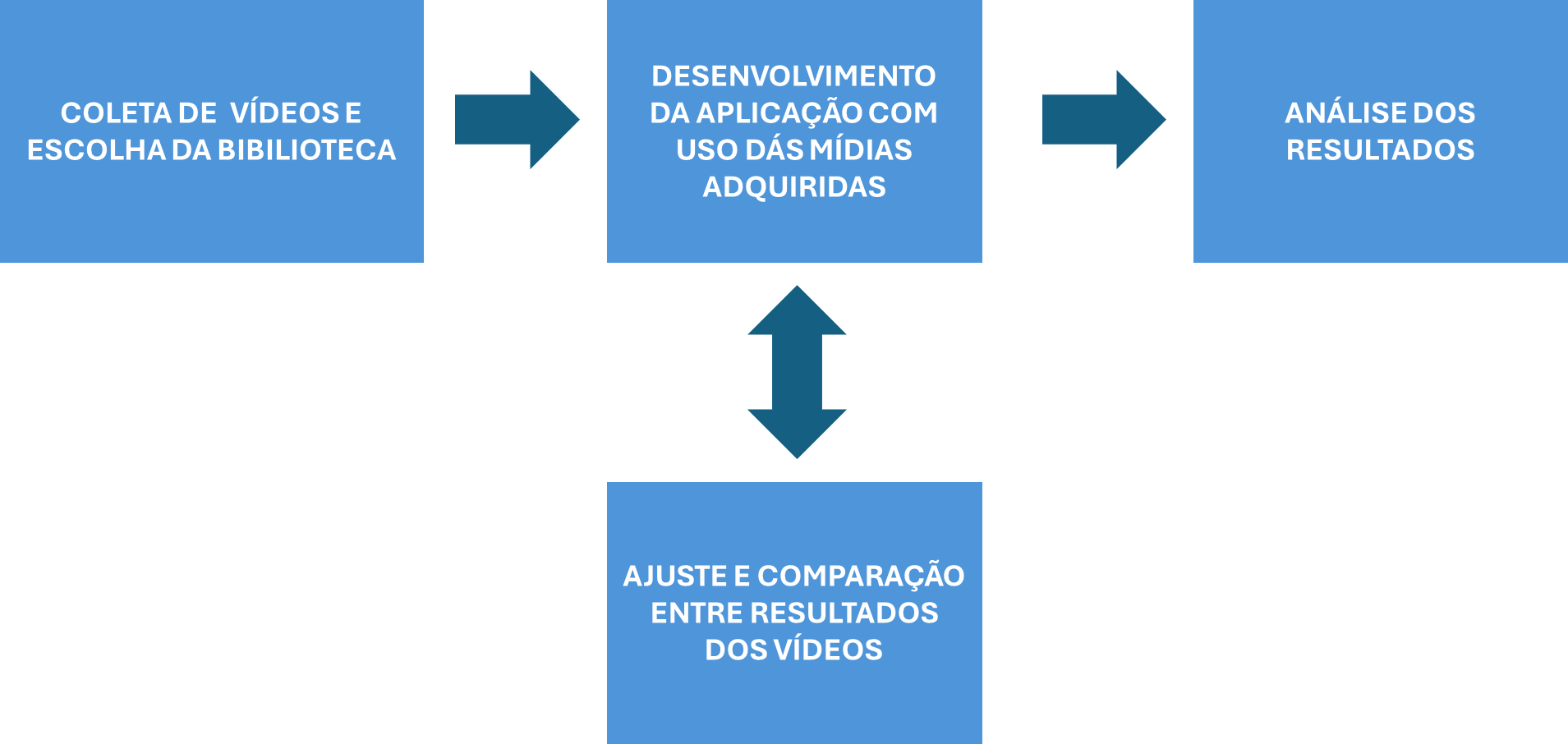
PROFESSOR AURÉLIO HOPPE

ACADÊMICOS: Alani Rigotti de Oliveira, Ari Júnior, Felipe Augusto de Carvalho de Araújo Luigi Marchetti

**TRABALHO 3 – DETECÇÃO DE OBJETOS**

Processamento de imagem é uma área da computação dedicada à extração, interpretação e manipulação de informações presentes em imagens digitais. Essa tecnologia tem sido amplamente aplicada em diversos setores, como no desenvolvimento de sistemas inteligentes de trânsito, na seleção de grãos na agricultura de precisão, entre outros. Dentre essas aplicações, destaca-se o uso em ambientes urbanos, especialmente na análise de semáforos, cruzamentos e faixas de pedestres, elementos essenciais para o monitoramento e controle do tráfego.

Neste contexto, este trabalho tem como objetivo analisar vídeos de mobilidade urbana com ênfase na detecção de semáforos. Para isso, é desenvolvido com a linguagem de programação Python, utilizando a biblioteca **YOLO (You Only Look Once)** em específico as versões 8, 11n e 11m. Essa biblioteca utiliza uma rede neural convolucional (CNN) e, por meio de caixas delimitadoras e probabilidades de classe diretamente a partir de frames do vídeo ou imagem, realiza a detecção de objetos em tempo real. A escolha do Yolo foi motivada pela vantagem da alta precisão e desempenho na identificação de elementos visuais relevantes em ambientes complexos.

Figura 1: Metodologia utilizada para o desenvolvimento do presente trabalho. 

(Fonte: elaborado pelos autores)

Para realizar o trabalho, o professor Aur´leio Hppe orientou a usar um modelo já treinado usando a documentação do. Então, feito isto, foi feito a configuração do Pytoch, ambiente de desenvolvimento, a biblioteca Yolo e, principalmente, o vídeo para aplicar o algoritmo desta biblioteca. Com isso, a partir do MOTChallenge, site que disponibiliza conjuntos de imagens e vídeos, foi escolhido dois vídeos. Um dos vídeos é mais simples, apenas com pessoas e carros para entendermos a biblioteca Yolo. E o outro mais específico, com fluxo de pessoas no período noturno.

Vídeo 1: <https://motchallenge.net/vis/MOT17-13>.

Vídeo 2:

Os testes iniciaram, de fato, com a aplicação da versão 11n da biblioteca Yolo no vídeo 1. Conforme na figura 2, podemos perceber que o resultado consistiu na identificação precisa de pessoas (person), adultos e crianças, e tipos de veículos, como caminhão (truck) e carro (car). Enquanto na versão 8x, além de identificar pessoas e veículos, identificou objetos mais distantes da câmera, conforme figura 3.

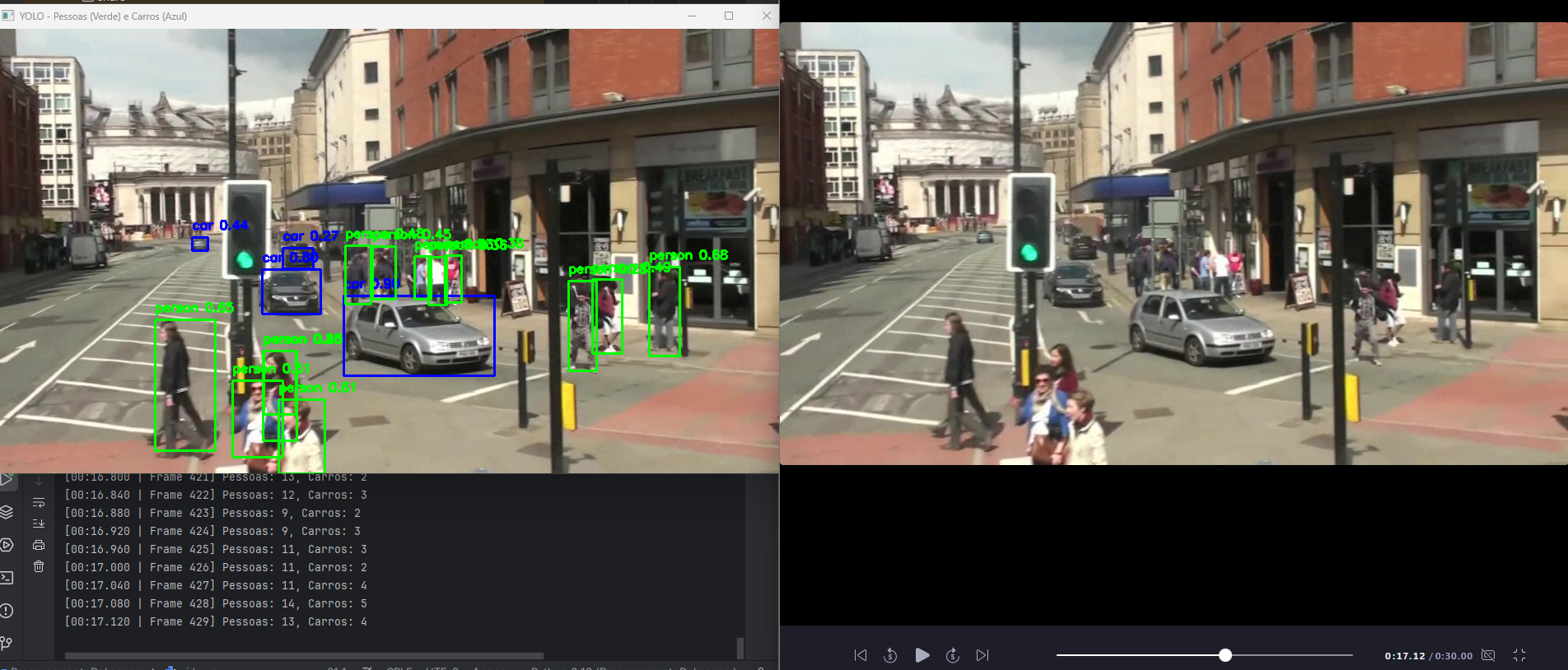
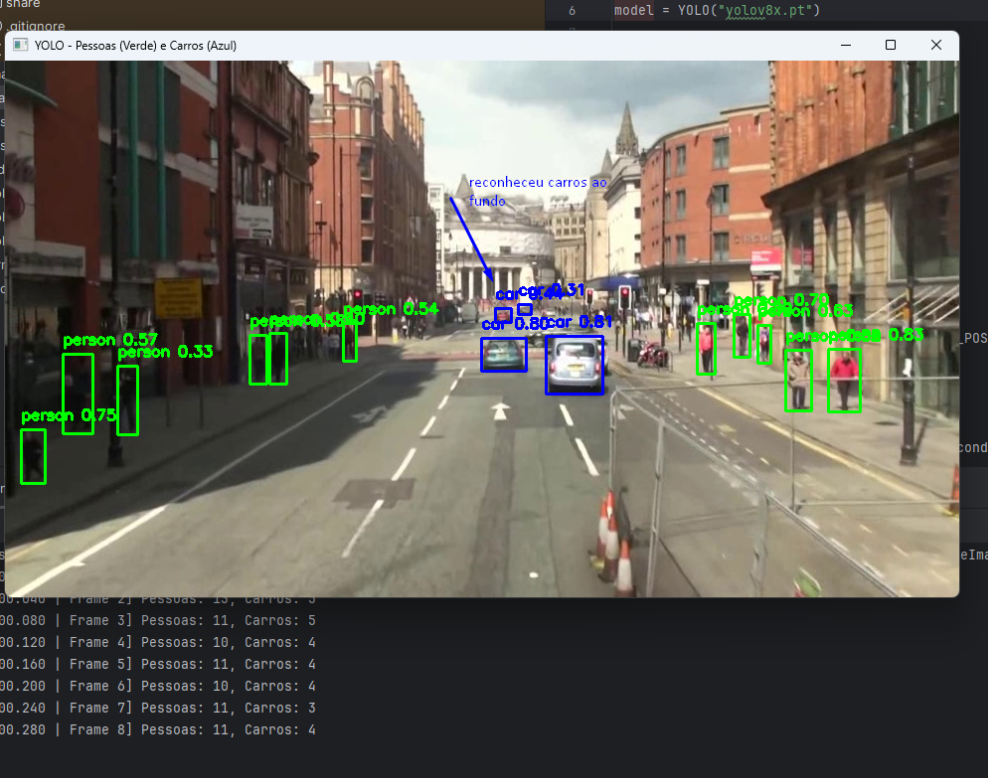
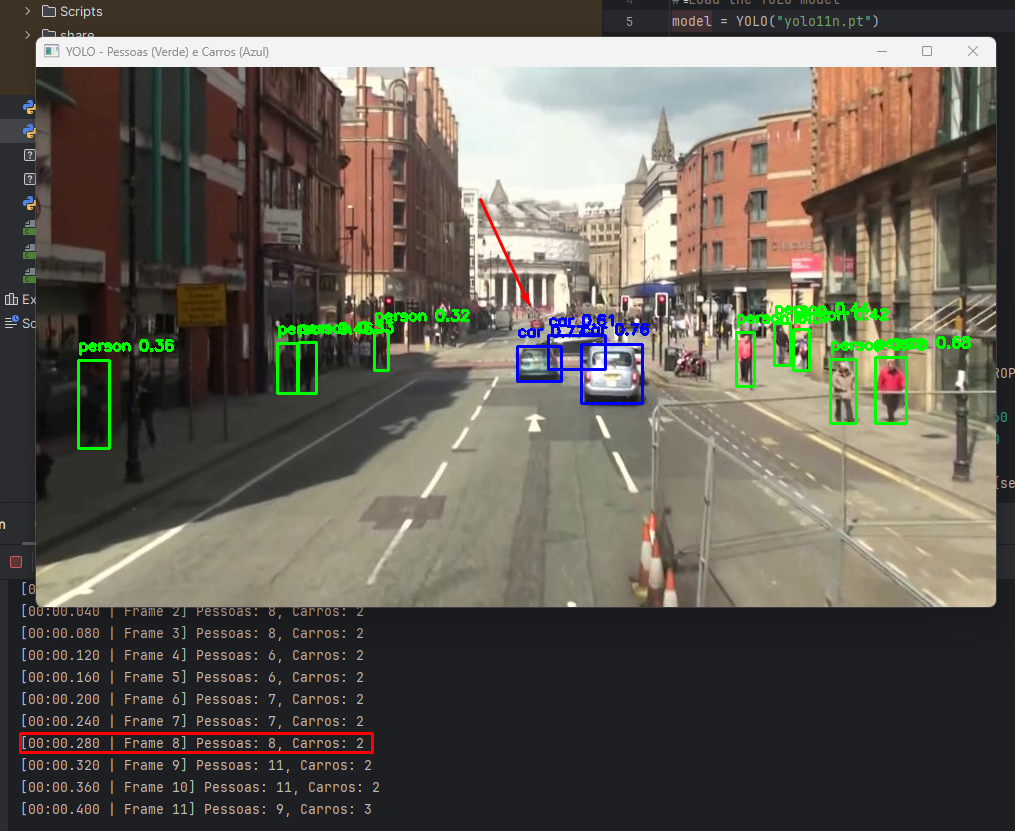
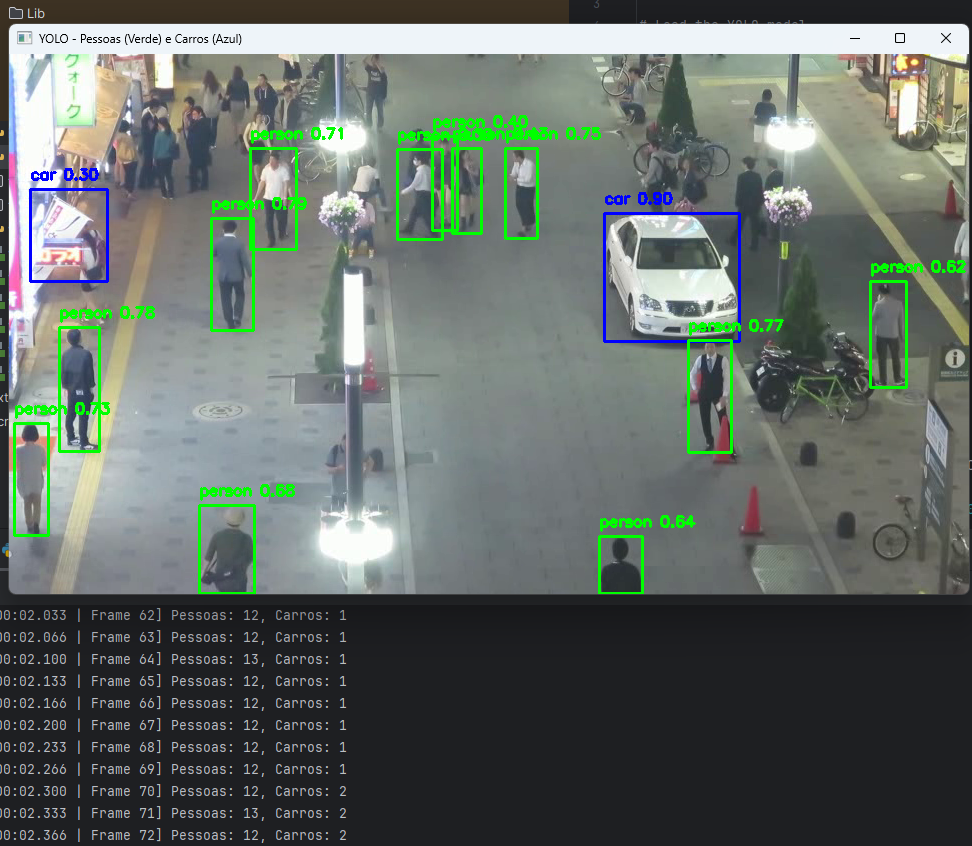
Figura 2: captura de frame do vídeo 1 utilizando a versão 11n da biblioteca Yolo.(Fonte: elaborado pelos autores)

Figura 3: Comparação no frame 8, yolo11n detectou 8 pessoas e 3 carros, yolo8x detectou 4 carros e 11 pessoas.(Fonte: elaborado pelos autores)

Para comparar melhor o desempenho da biblioteca Yolo em diferentes contextos de iluminação e objetos, foi utilizado o vídeo 2. Nesse segundo teste, foram utilizadas as versões 11n, 11x, 11m e 8x. Na comparação das versões, é possível observar que a detecção possui obstáculos como a baixa iluminação e sobreposição de objetos. Na versão 11x performance baixa (ainda melhor em relação a 8x), e a maior precisão de detecção de objetos, poucos falsos positivos, alto grau de confiança, porém ainda com certa dificuldade de detectar pessoas em aglomerações. Já o modelo 11m apresentou uma performance média, com pouca melhora na detecção em relação ao 11n.

Figura 4: captura de frame do vídeo 1 utilizando a versão 11x da biblioteca Yolo.

(Fonte: elaborado pelos autores)

Figura 5: Captura de falso positivo utilizando a versão 11m.

(Fonte: elaborado pelos autores)

Após dos testes, concluímos que o melhor resultado obtido foi com o modelo 11n, devido a performance ser muito boa e a detecção ter uma qualidade relativamente próxima a 11x. Caso se preze principalmente pela precisão seria recomendado o 11x, mas seria necessário uma máquina muito potente. Para referência, em nossos testes utilizamos uma máquina com um processador Ryzen 5 5600 com clock de 3.5 GHz, e placa gráfica RTX 3060 com 3584 CUDA Cores, e rodou a (TENTAR VER QUANTOS FPS RODA).