

## **APRIMORAMENTO DE SOFTWARE DE ANOTAÇÃO: RASTREAMENTO DE OBJETOS EM VÍDEO DE FORMA MANUAL**

MELLO, João  
MACHADO, Matheus

Silva, Luciano  
joao.mellof@furg.br  
Universidade Federal do Rio Grande

**Palavras-chave:** detecção de objetos; marcação de objetos manual; rastreamento manual; rastreamento de objetos.

### **1 INTRODUÇÃO**

A detecção de objetos consiste na identificação e localização de itens que pertençam a uma ou várias classes de objetos de interesse. As aplicações se estendem do setor automobilístico - carros autônomos ao agropecuário, na contagem de animais de um rebanho. A tarefa de detecção de objetos pode ser facilmente realizada por seres humanos, dado que o seu sistema de visão passou por centenas de anos de evolução e é extremamente preciso. Entretanto, a detecção manual realizada por humanos apresenta limitações no que tange à volume e escala das atividades. Nesse cenário, a computação vem consistentemente apresentando soluções eficientes, sendo a mais conhecida delas as redes neurais.

O uso de redes neurais para a detecção de objetos foi sempre muito custoso, principalmente por se utilizar das GPUs para o processamento. Porém, as recentes redes neurais convolucionais (RNC) constituem uma solução mais eficiente, sobretudo performática, como indica (REN, 2015). Para que uma RNC opere com qualquer grau de confiabilidade e precisão, é necessário que esta seja treinada, ou seja, que pessoas, de forma manual, desenvolvam bases de dados de treinamento contendo exemplos de como o problema é resolvido. A partir das bases de dados desenvolvidos por humanos, a rede busca uma configuração que chegue mais próximo ao resultado estipulado pelas pessoas.

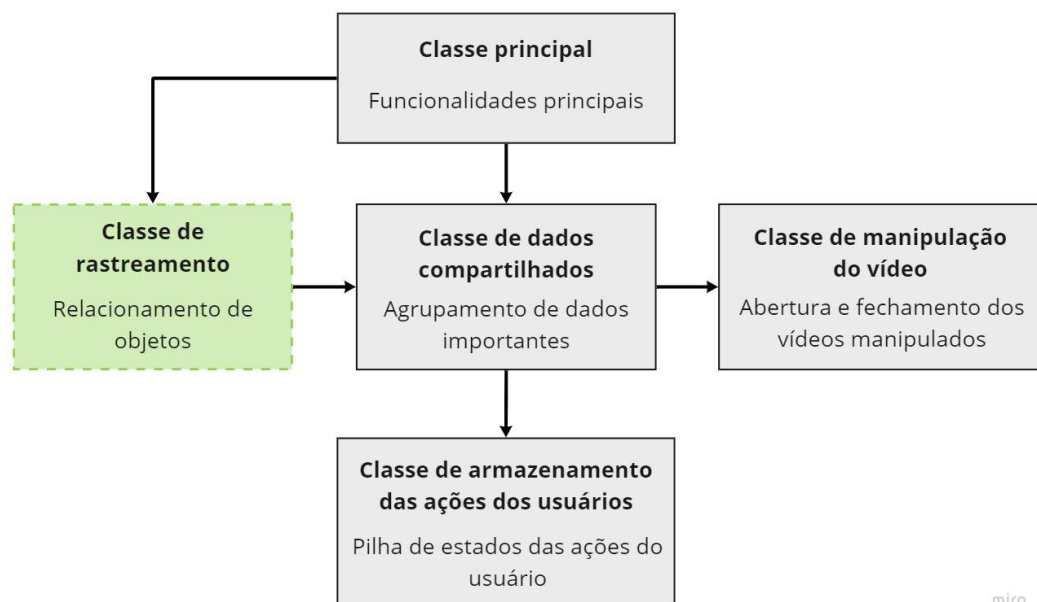
Com intuito de se desenvolver uma base de dados de treinamento, foi elaborada uma ferramenta de marcação que, a partir de um vídeo de entrada, permite realizar anotações - *bounding boxes* - dos objetos de interesse do usuário. Porém, em um contexto onde se precisa contar objetos, por exemplo, é necessário determinar a correspondência entre um mesmo objeto detectado em diferentes quadros do vídeo, para que este não seja contado múltiplas vezes. Essa situação corresponde ao problema de rastreamento, que conforme (BERGMANN; MEINHARDT; LEAL-TAIXE, 2019), apresenta outros empecilhos, a citar casos de oclusão ou alta densidade de objetos.

Nesse sentido, o presente artigo possui como propósito abordar como a funcionalidade de relacionamento de objetos entre quadros distintos foi pensada, desenvolvida e adicionada na ferramenta.

## 2 METODOLOGIA

A partir do estudo da estrutura e modo de implementação da ferramenta atual, optou-se por implementar a nova funcionalidade de rastreamento a partir de uma nova classe, de forma a seguir princípios importantes, no que tange à organização e manutenção do código existente. De maneira simplificada, a estrutura das classes da aplicação já com a nova implementação pode ser vista na Figura 1.

Figura 1 - Diagrama UML da ferramenta de marcação atualizada



Fonte: Os autores.

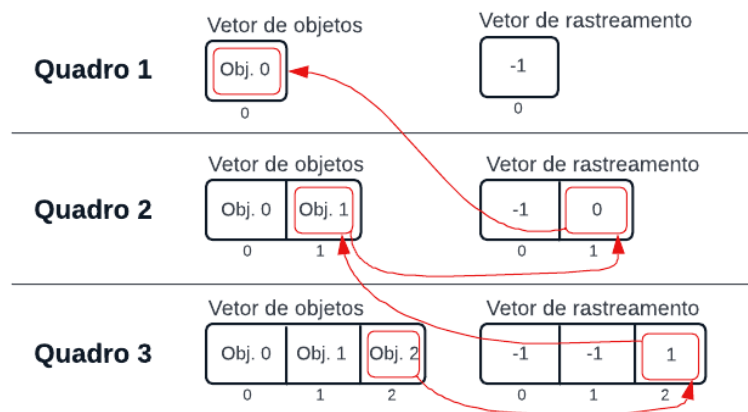
A nova funcionalidade de rastreamento foi implementada por meio de uma classe adicional com intuito de não modificar as demais estruturas da ferramenta. Como pode ser visto, na Figura 1, a aplicação já possui divisão de tarefas e os módulos ficam desvinculados, o que é essencial na organização e escalabilidade do código.

A funcionalidade de rastreamento foi desenvolvida com a linguagem de programação C++, conforme os demais módulos da aplicação. Com intuito de se utilizar estruturas de dados complexas, como por exemplo vetores, a biblioteca *standard* foi requisitada. Outro pacote utilizado foi a *OpenCV*, com a finalidade de se obter os eventos de botões pressionados no teclado e renderizar a interface.

A respeito da estrutura de dados utilizada para o rastreamento, são utilizados dois vetores: o primeiro armazena as marcações de objetos, enquanto o outro guarda os relacionamentos, conforme a Figura 2. Cada quadro do vídeo possui

esse par de vetores, de forma que o de rastreamento indica com qual objeto do quadro anterior um dado objeto do respectivo quadro possui relação. No caso da Figura 2, o objeto 1 do segundo quadro possui relação com os exemplares 0 e 2 dos quadros 1 e 3, respectivamente. Já os objetos 0 e 2 do quadro 3 não possuem correspondentes.

Figura 2 - Estrutura de dados da nova classe



Fonte: Os autores.

Uma vez que o usuário realizou o rastreamento dos objetos na ferramenta, é necessário persistir os dados em disco, para que depois seja possível a utilização dos relacionamentos em outros contextos. Para implementar a persistência dos dados, se modificou a rotina já existente, responsável por salvar as anotações de objetos. As informações são guardadas em um arquivo binário de extensão .nt.

### 3 RESULTADOS E DISCUSSÃO

Após a implementação da nova classe de rastreamento e experimentações, o resultado visual dos relacionamentos representados da Figura 2 pode ser visto na Figura 3.

Figura 3 – Capturas de tela do relacionamento de objetos



Fonte: Os autores.

A Figura 3 mostra o rastreamento de maçãs ao longo de três quadros consecutivos. Os quadrados azuis indicam a marcação de uma maçã, e o índice interno à caixa, o número identificador da mesma.

A partir do momento em que o usuário estabelece uma relação entre objetos, nesse caso maçãs, no canto superior do retângulo em questão aparece um círculo preenchido, na cor magenta, juntamente com um número identificador. Essas informações mostram com qual objeto do quadro adjacente o exemplar em questão se relaciona.

Se há vínculo com o quadro anterior, o círculo e identificador aparecem no canto superior esquerdo, e de maneira similar, a relação com o quadro posterior é mostrada na direita. Essa funcionalidade pode ser vista explicitamente no segundo quadro, cuja maçã de índice 1 possui vínculo com ambos os quadros adjacentes.

Por fim, objetos que não possuem correspondentes não recebem nem círculos nem identificadores nos cantos superiores, o que pode ser visto com algumas maçãs. Como empecilhos da nova funcionalidade, pode-se destacar a possível confusão que pode ser originada com as informações de relacionamento, caso a densidade de objetos seja alta. Além disso, como todos os relacionamentos são da mesma cor - magenta, o usuário pode ser facilmente confundido e levantar dados equivocados.

#### 4 CONSIDERAÇÕES FINAIS

A partir da necessidade de se realizar o rastreamento de objetos em vídeo, implementou-se uma nova classe na ferramenta de marcação de objetos já existente. A integração da funcionalidade foi realizada de forma desacoplada, garantindo a integridade e organização do código da aplicação.

Como sugestão de aprimoramento, destaca-se a busca por uma alternativa que deixe a interface menos confusa, quando a densidade de objetos for maior. Além disso, é conveniente implementar diferentes cores para os relacionamentos, o que ajuda o usuário a realizar os vínculos.

Por fim, vale ressaltar que os dados de rastreamento gerados de forma manual serão utilizados por uma outra aplicação, responsável por comparar a precisão e eficiência do algoritmo de rastreamento automático.

#### 5 REFERÊNCIAS

BERGMANN, Philipp; MEINHARDT, Tim; LEAL-TAIXE, Laura. Tracking without bells and whistles. In: **Proceedings of the IEEE/CVF International Conference on Computer Vision**. 2019. p. 941-951.

REN, Shaoqing et al. Faster r-cnn: Towards real-time object detection with region proposal networks. **Advances in neural information processing systems**, v. 28, 2015.