

REO  
ATC

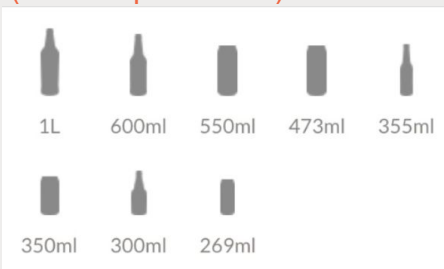
# Visão computacional na devolução pós-consumo

Fase 1: Testes de feasibility da solução

12/12/2020

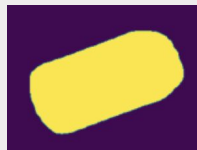
# Motivação

(Gerado pelo GTIN)



Embalagem	Área	Perímetro	...
Garrafa 1L	$190\text{cm}^2$	30cm	...
Lata 350ml	$60\text{cm}^2$	22cm	...
Lata 269ml	$45\text{cm}^2$	16cm	...

RECO  
ATO :



$56\text{cm}^2$ , 25cm, ...



**Lata 350ml**

# Overview

1. Goals and Non-goals
2. *Object Detection*
3. *Instance Segmentation*
4. Combinação de OD + IS
5. Questões em aberto e próximos passos

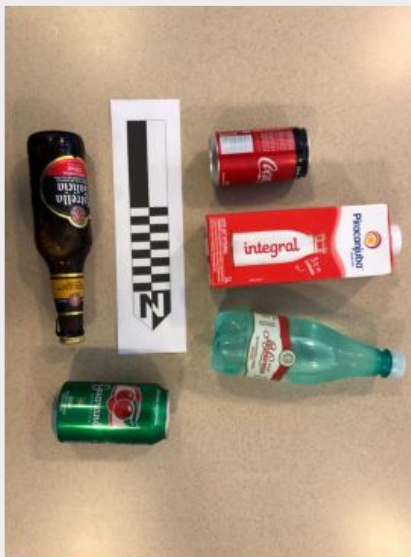
# Goals

- Prototipagem rápida com apenas três imagens parecidas, para possível *fail fast*
- Avaliar a possibilidade de dividir o problema em 1. Determinação do espaço ocupado por cada objeto, 2. classificação do objeto com ajuda de referências.
- Avaliar métodos de Object Detection e de Instance Segmentation para solução da parte 1 do problema.
- Avaliar a feasibility de usar redes pré treinadas no nosso contexto de reciclagem

# Non-goals

- Ter previsões prontas apenas com as redes pré treinadas (i.e. parte 2 do problema em aberto)
- Possuir técnicas que já são imediatamente implementáveis em um app, ou com velocidade satisfatória.
- Garantir funcionalidade para o caso real de uso

# Três imagens para prototipagem



***Object detection***



Usando *Object Detection* nós geramos caixas limite (*bounding boxes*) ao redor de cada objeto detectado. Pode haver superposição entre múltiplas caixas. Cada caixa é etiquetada com o objeto previsto.





É possível fazer a retirada dos objetos individuais das fotos a partir das caixas limite



Embora por vezes acertem, as previsões de classe da rede poder errar grosseiramente, como no exemplo da caixa de leite.

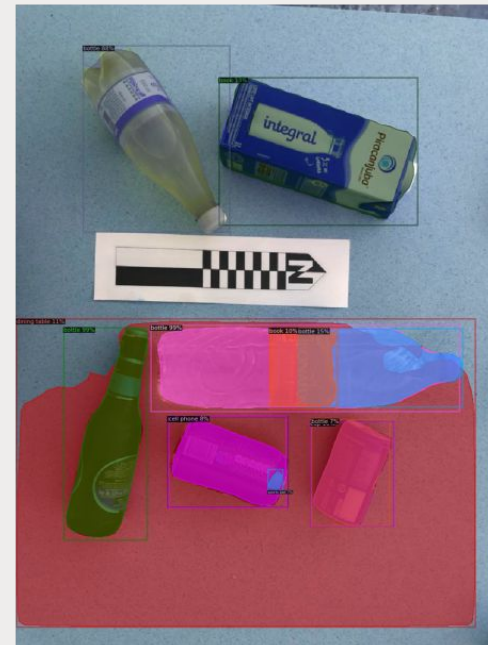
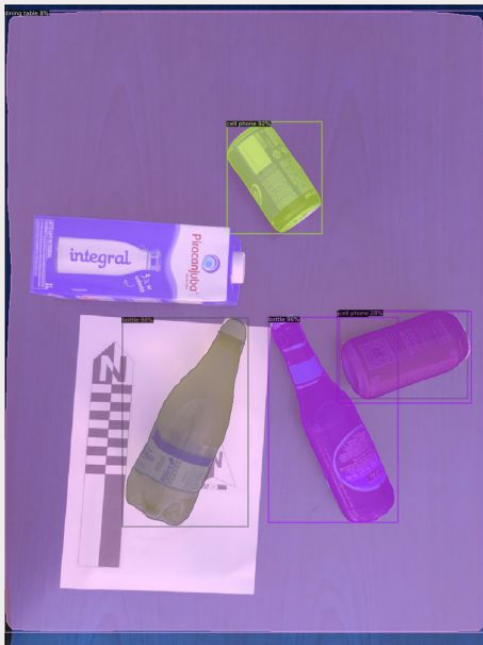


É possível perceber que há utilização de contexto nas previsões da rede, que chamam de “mesa de jantar” um amontoado de garrafas.

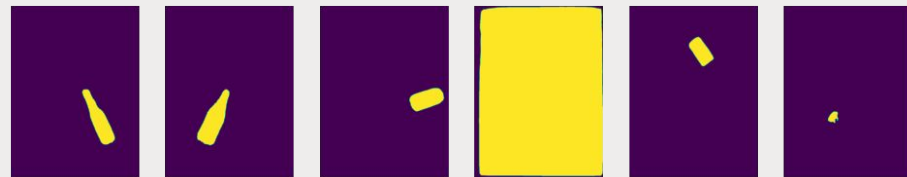
Pode-se ver duas caixas sobrepostas na garrafa bonafont.

***Instance segmentation***

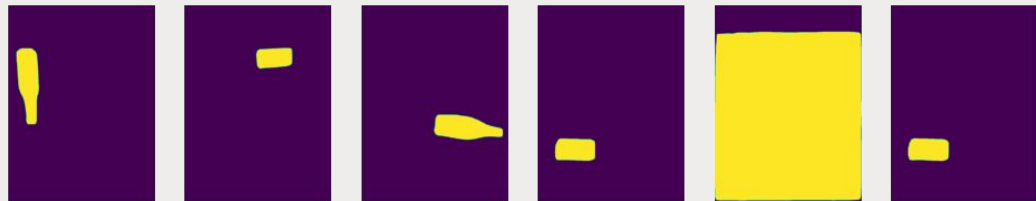




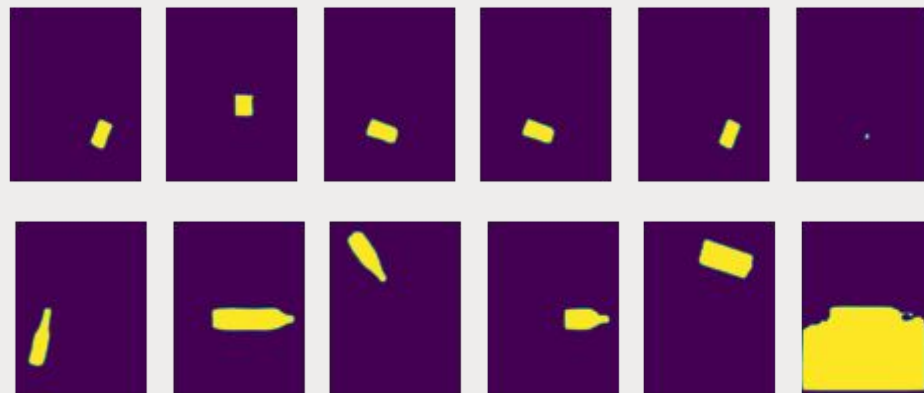
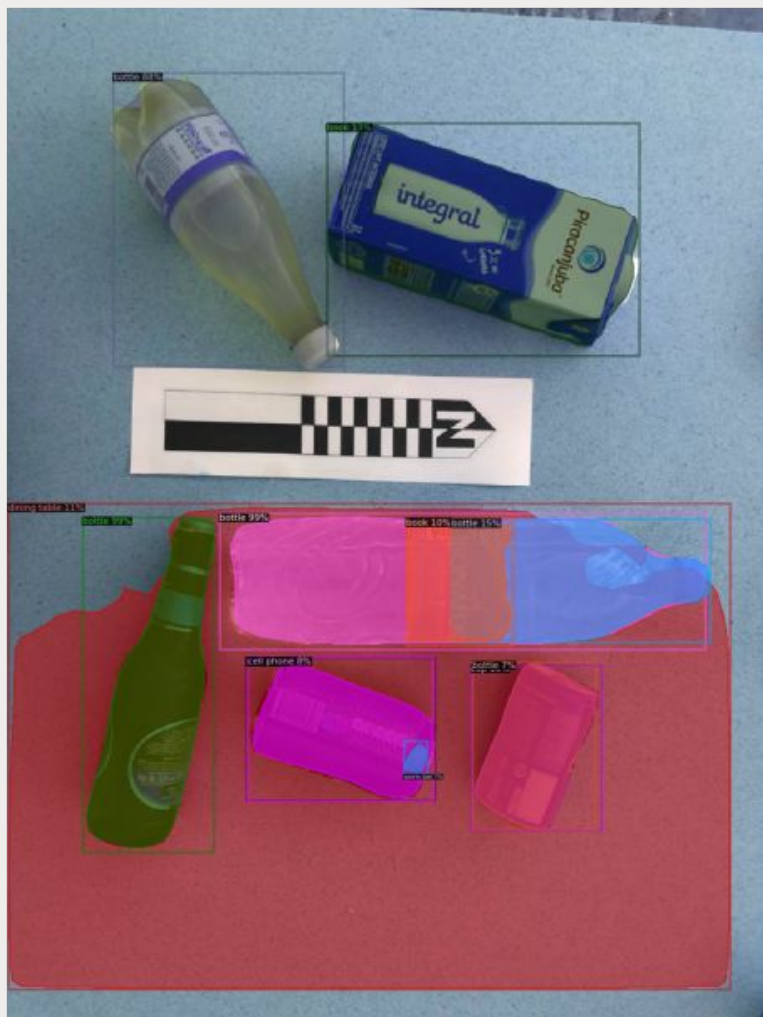
Usando *Instance Segmentation* nós geramos *máscaras* (*masks*) sobre cada objeto. Pode haver superposição entre múltiplas máscaras. Também são geradas caixas-limite e previsões para os objetos.



A segmentação de instâncias é capaz de fazer identificações diretamente da área de objetos. A detecção de objetos parece um pouco pior como na caixa de leite.



As máscaras poderiam ser usadas diretamente para comparação com referências.



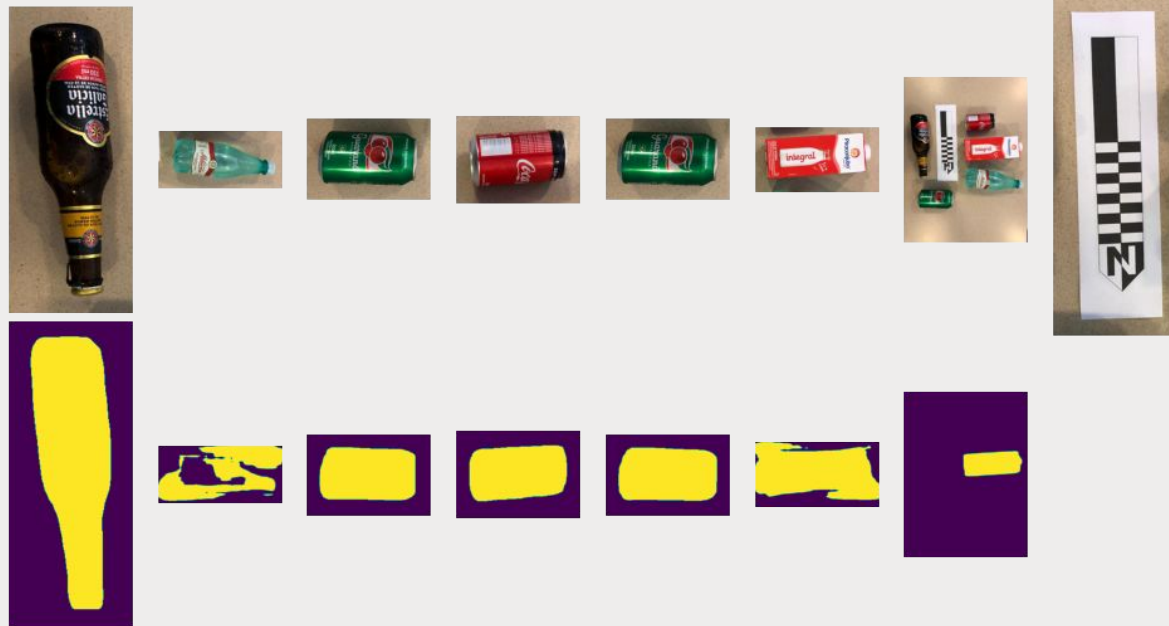
Também pode haver sobreposição, como podemos ver na garrafa PET.



*OD + IS*



Parece haver maior dificuldade de usar a solução neural de segmentação com a imagem já cortada. São geradas *máscaras* com partes adicionados aos objetos



Talvez nesse caso de segmentação posterior à Detecção, outros métodos pudessem ser usados. Com essa formulação no entanto não parece uma solução viável.

# Métricas relevantes

- Approx 6s para Object Detection de uma foto.
- Approx 7s para Instance Segmentation.
- Tempo para IS é o mesmo mesmo para fotos pequenas (e.g. caso OD + IS)
- Provavelmente há formas de melhorar

***Questões em aberto  
e próximos passos***

# Algumas perguntas em aberto

- A qualidade de detecção é suficiente?
  - Não temos exemplos o suficiente para dizer que a segmentação será boa o suficiente para as classificações futuras. Podem ser necessárias otimizações, ou a técnica pode ainda ser descartada dependendo de sua performance.
- Como lidar com a sobreposição de máscaras?
  - Escolhendo a maior / ordenando de acordo com sua prob e testando cada uma
- Como capturar a régua de tamanho automaticamente, permitindo avaliar os tamanhos dos objetos sem depender da altura da câmera?
- Como atacar a parte 2 do problema, de classificar as máscaras?