

## DESCRIÇÃO ESPACIAL QUALITATIVA DE CENAS EM AMBIENTES INTERNOS COM VISÃO COMPUTACIONAL

**Aluno:** Gustavo Fuentes, Jefferson Bizarri e Thomas Neves (fuentesfei99@gmail.com, jefferson.bizarri@gmail.com, tdn.1999@gmail.com)

**Orientador:** Prof. Dr. Danilo Hernani Perico (DPERICO@FEI.EDU.BR)

### Resumo

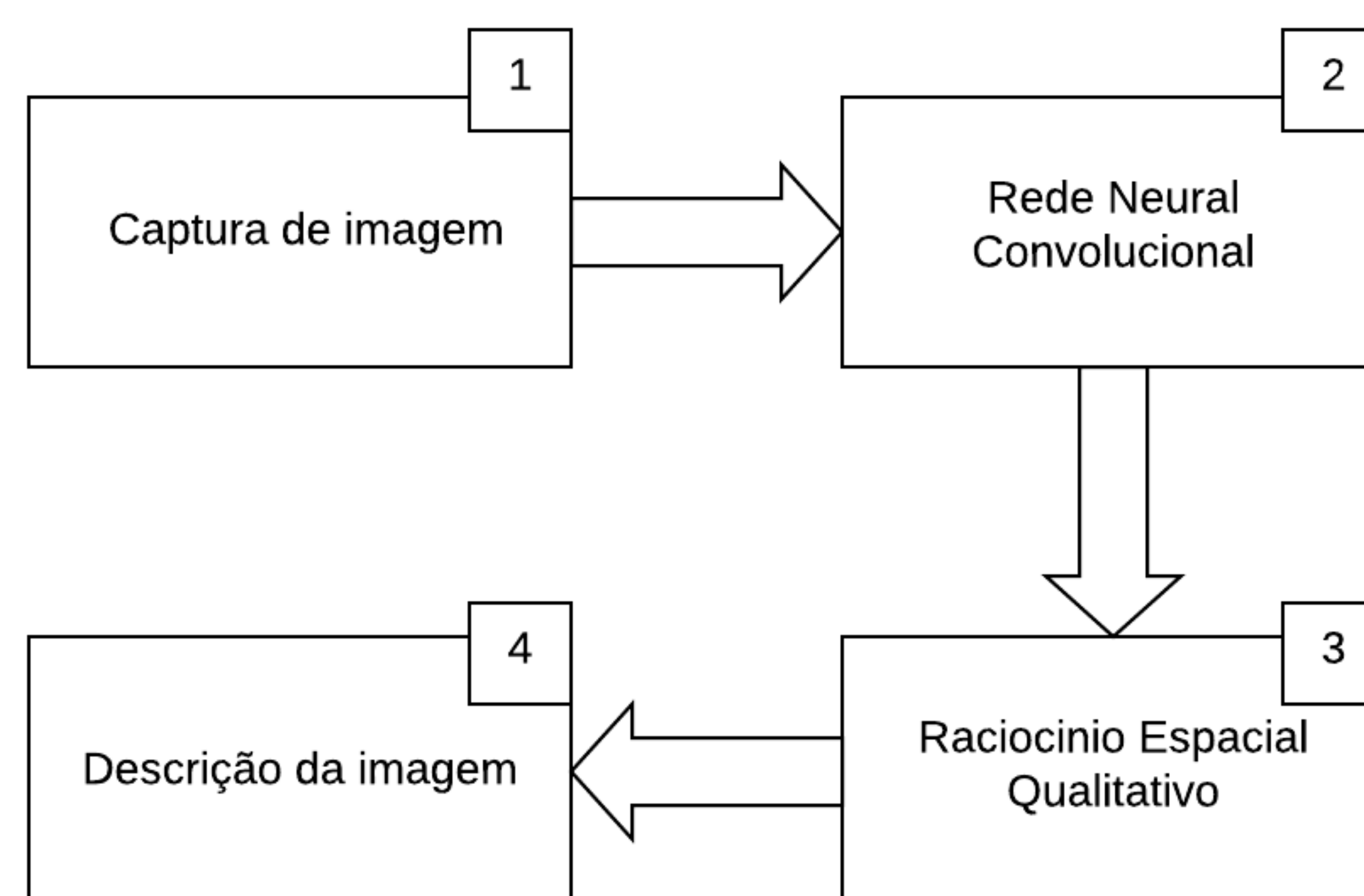
Com avanço da área de visão computacional podemos automatizar processos que envolvem imagem de uma forma que não era possível antigamente. Podemos assim propor formas de reduzir ou resolver problemas como, a dificuldade de um deficiente visual se localizar em um ambientes novos, a catalogação de objetos em uma cena de crime ou até mesmo melhorar a comunicação entre humano e máquina. Podemos utilizar técnicas de localização espacial de objetos em imagens e descrever suas posições de forma textual, assim um deficiente visual teria uma melhor noção espacial de um ambiente novo, um perito criminal poderia gerar documentações de objetos em uma cena de crime ou um humano poderia passar instruções mais flexíveis para um robô. Para a nossa solução técnica utilizamos uma Rede Neural Convolutiva que é responsável por identificar objetos que serão processados por um algoritmo de Reconhecimento Espacial Qualitativo encarregado de classificar a posição relativa do objeto identificado. Utilizando esses métodos alcançamos uma precisão de 52% entre 15 objetos que podem ser identificados em ambientes internos. Com essas aplicações, o dia-a-dia de pessoas que enfrentam as situações citadas anteriormente pode ser facilitado, trazendo mais conforto para tarefas realizadas diariamente.

### Descrição

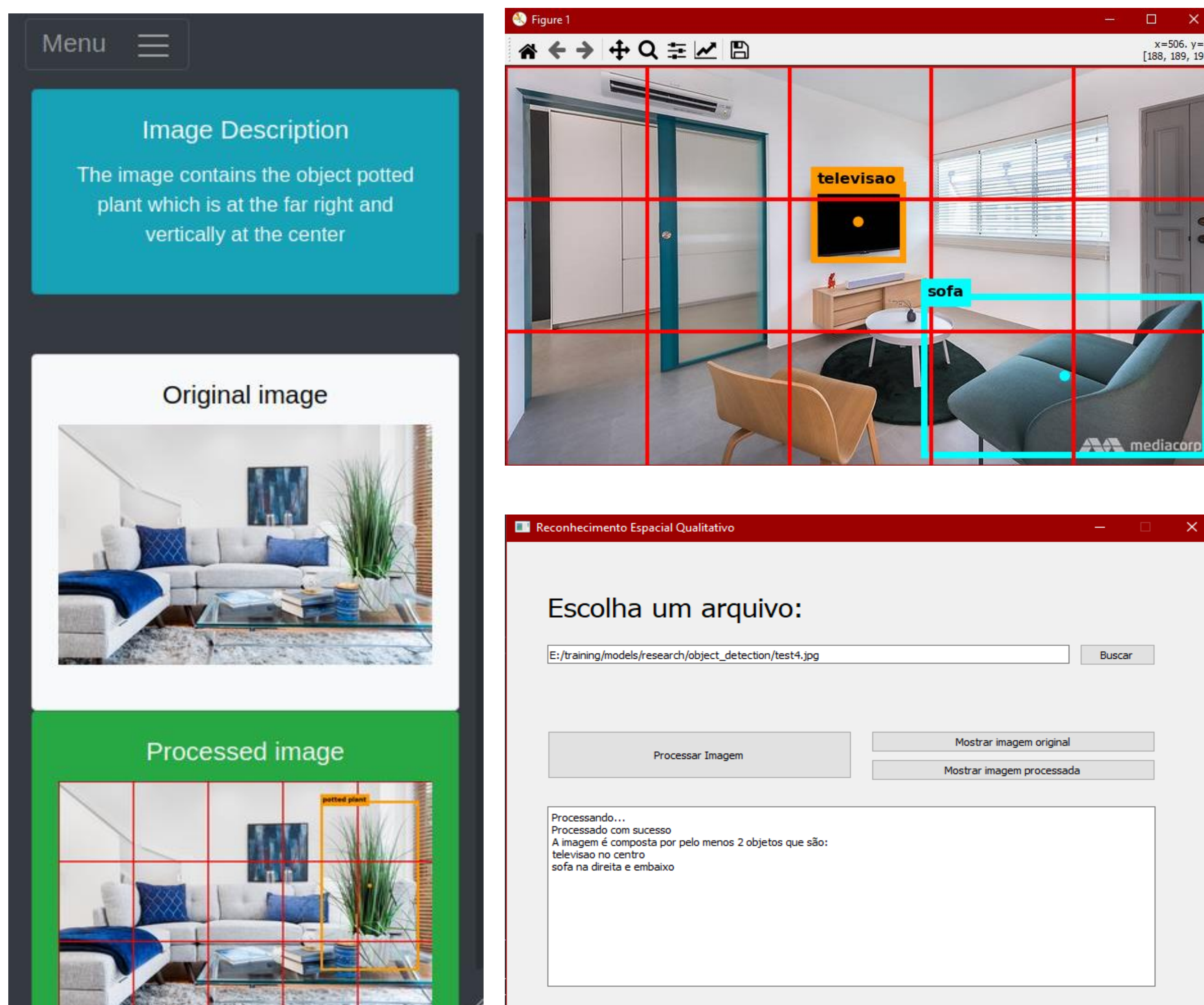
De acordo com a evolução tecnológica em áreas como visão computacional e inteligência artificial, podemos ver a criação de diversas ferramentas afim de agregar valor em muitos campos, como a área da saúde e lazer, por exemplo, com interações surreais de um deficiente visual com o mundo, ou até mesmo de um simples jogo interagindo com a realidade a sua volta. Com isso, propomos uma tecnologia de modo a aumentar o conhecimento desses campos, desenvolvendo uma aplicação que identifica objetos em ambientes internos e descreve a posição relativa desses objetos na imagem processada. Para o desenvolvimento da proposta, foi utilizado conceitos de Redes Neurais Convolucionais (RNC), utilizando a Faster R-CNN para cumprir o papel da detecção dos objetos além de ser criado um algoritmo baseado no fundamento de Raciocínio Espacial Qualitativo, o qual indica as posições de um objeto em relação a câmera (tendo como representação horizontal as seguintes coordenadas: muito à esquerda, esquerda, centro relativo a vertical, direita e muito à esquerda e como coordenadas verticais temos: em cima, centro relativo a vertical e embaixo), tomando como exemplo, podemos supor que temos uma foto de uma sala de estar, o algoritmo irá detectar os objetos presente e processar essa imagem de forma que gere um resultado como, "A imagem é composta por 2 objetos, sejam eles o sofá que está na direita e embaixo e a televisão que está na esquerda em cima". Com a conclusão desse projeto tivemos a concretização de uma API disponível em um servidor do AWS, com uma RCN treinada com 15 classes de objetos sendo elas, Pessoa, Gato, Cachorro, Cadeira, Sofá, Cama, Mesa de Jantar, Fogão, Geladeira, Pia, Vaso sanitário, Micro-ondas, Televisão, Notebook e Vaso de planta, de precisão média entre todos os objetos de 52%. Com esse nível de aplicação podemos trazer uma melhora para o dia-a-dia de deficientes visuais conseguirem criar um mapa mental de um ambiente que ele ainda não conhece bem, mapear e documentar cenas de crime com grande velocidade, além de poder aumentar ainda mais a qualidade com a qual um robô pode interagir com o ser humano.

### Metodologia

Consiste em fazer a captura de uma imagem, processá-la através de uma Rede Neural e aplicar um algoritmo (REQ) capaz de classificar as posições dos objetos, gerando assim a descrição da imagem.



### Resultados



### Conclusão

O descritor espacial qualitativo trouxe resultados capazes de identificar e descrever as posições dos objetos em uma imagem qualquer de um ambiente interno. O descritor está disponível como uma aplicação Desktop e em formato de website. O código-fonte pode ser encontrado no GitHub.