Módulos e tarefas por módulo:

1. Classificador que identifica em que ecrã nos encontramos.
   1. Pesquisa de modelos pré-treinados em GUIs.
   2. Retirar screenshots. Pesquisa de vídeo, script para gravar diversas frames.
   3. Anotação das frames com o label do menu em que nos encontramos
2. Leitor de conta e preço final
   1. Text detection (CRAFT?)
   2. Text recognition (OCR)
   3. Regras de aglomeração
   4. JSON adapter for API output
3. Identificador de botões de mesas
   1. Text detection (CRAFT?)
   2. Text recognition (OCR)
   3. Extrator de coordenadas dos botões
   4. Módulo de navegação inteligente (ações a fazer para chegar onde se quer no menu, de acordo com o menu onde nos encontramos – finite state machine)
   5. Auto clicker
4. Resiliência a erros e outros
   1. Criação de dataset com sobreposição de mensagens de erro a ecrãs
   2. Retreino de classificador de GUI, que inclua categorias/tipos de erro
   3. Adaptação do módulo de navegação inteligente/FSM para saber agir de acordo com o erro