

Para que o OCR (Reconhecimento Óptico de Caracteres) funcione corretamente, é necessário garantir que o arquivo PDF seja processado corretamente. Existem algumas possibilidades que podem afetar o sucesso do reconhecimento de texto:

1. **Formato do PDF:** Se o PDF é baseado em imagem (por exemplo, um escaneamento de um documento), o OCR só funcionará se a imagem contiver texto visível. Se for um PDF com texto incorporado (não uma imagem), a função OCR não será necessária, pois o texto já está embutido no PDF.
2. **Conversão de PDF para Imagem:** PDFs com texto embutido não precisarão de OCR, mas se o seu PDF for uma imagem, você precisará primeiro converter cada página para imagem e depois aplicar o OCR nessas imagens.

