1. Conclusões
   1. Considerações Finais

O objetivo principal deste projeto é a solução de um problema que está despertando a cada dia mais interesse nas empresas, que é a necessidade de se possuir sistemas que agilizem o processo por meio da automação de tarefas, reduzindo não somente o tempo gasto para as mesmas, mas o custo empregado. Atualmente, os sistemas de OCR tradicionais permitem extrair um texto de uma imagem, porém não é atribuído sentido à esta informação extraída. Como discutido a seção XXX, os sistemas atuais possuem a limitação de estarem limitados à uma tecnologia de digitalização de imagens e uma interface gráfica.

Portando, neste projeto foi proposto a criação de uma ferramenta que além de trazer todos os benefícios da leitura automática ainda possui o diferencial de ser um sistema completamente desacoplado da interface gráfica, possibilitando uma rápida atualização das interfaces, ou mesmo do sistema de reconhecimento, sem que o funcionamento do mesmo seja comprometido.

A estrutura de reconhecimento das regiões de interesse do documento foram definidas depois de pesquisas e testes. A arquitetura definida e implementada mostrou-se bastante eficiente, cumprindo com precisão e rapidez seus objetivos.

O resultado do trabalho é um sistema que está muito próximo de um sistema que pode ser comercializado e aplicado em situações reais para o cadastramento de clientes, como hotéis, locadoras de veículos, entre outros.

* 1. Propostas de continuidade

Com o objetivo de transformar o sistema o mais genérico possível algumas sugestões de continuidade do trabalho:

* Otimização no processamento: O método de processamento da imagem até a identificação de linhas é o mais lento do sistema, comprometendo o desempenho do sistema. A maneira que foi desenvolvido atende às especificações, uma vez que ao velocidade de leitura do documento é muito superior ao processo manual, porém este é um ponto em pode ser realizado uma melhoria;
* Acrescentar um identificador de documentos: este sistema foi desenvolvido com o objetivo de digitalizar a carteira de habilitação, porém existem diversos outros documentos, como documento de identidade, passaporte, carteira de trabalho. Pode-se acrescentar um módulo para reconhecer o documento automaticamente, a partir de modelos já criados.
* Acrescentar um modulo para cadastrar novos modelos: além de tornar o sistema mais genérico, permitindo a identificação e leitura de diversos tipos de documentos, seria interessante criar uma interface para que o próprio usuário dos sistema pudesse cadastrar novos documentos.
* Alterar o método de leitura: Apesar do Tesseract OCR apresentar um bom desempenho, foi observado que antes da criação dos modelos de leitura, ocorriam diversos erros, por exemplo, reconhecia uma letra no campo CPF. Para evitar estes erros, criado três leitores delimitando o conjunto de caracteres possíveis em cada campo. Porém, seria interessante avaliar as outras tecnologias de OCR existentes, com o objetivo de melhorar a qualidade e assertividade do sistema.
* Criar um mecanismo para identificação de fraudes: Atualmente, o número de falsificações de documentos, principalmente a carteira de habilitação é crescente. Isto se deve ao fato que este documento pode ser utilizado como documento de identificação, que permite que a pessoa se passe por outra ou mesmo aumente sua idade. Além disto, é um documento necessário para certificar que a pessoa está habilitado para conduzir aquele veiculo e muitas vezes a pessoa não tem dinheiro para pagar o processo para retirar a documentação de maneira legal e acaba optando pela falsificação. Apesar dos crescentes esforços para evitar as falsificações, que resultam em diversas marcas de segurança nos documentos, as falsificações estão cada vez mais fidedignas. Dessa forma, um sistema que auxiliasse na análise de falsificação dos documentos, conferindo fonte, tamanho da escrita, alinhamento dos campos entre outros fatores, seria de grande utilidade, aumentando a segurança contra pessoas mal intencionadas, como estelionatários.

Dessa forma, este trabalho termina com a sugestão de que este sistema consiga ser uma ferramenta completamente autônoma e genérica, que extraia a informação dos mais variados tipos de documentos, permitindo a criação de novos modelos sem que seja necessário intervenção de um desenvolvedor, facilitando assim, a vida do operador do sistema.