

COMPROVAÇÃO DE IDENTIDADE A PARTIR DE DOCUMENTOS OFICIAIS FORNECIDOS POR USUÁRIOS COMUNS

Aluno: Filipi de Luca Valim dos Santos, Gabriel Diniz Gisoldo (lipi-valim@hotmail.com, gabrielgisoldi@gmail.com)

Orientador: Prof. Dr. Plinio Thomaz Aquino Junior (plinio.aquino@fei.edu.br)

Resumo Validar a identidade é uma forma de verificar as informações fornecidas pelo usuário, a partir disso muitas empresas possuem setores apenas para validar a identidade de um usuário e afim de diminuir custos operacionais e poupar muito tempo no processo, propomos um sistema de validação digital validando e finalizando o cadastro.

Em testes foram obtidos bons resultados, com uma media de 80% de validação de documentos.

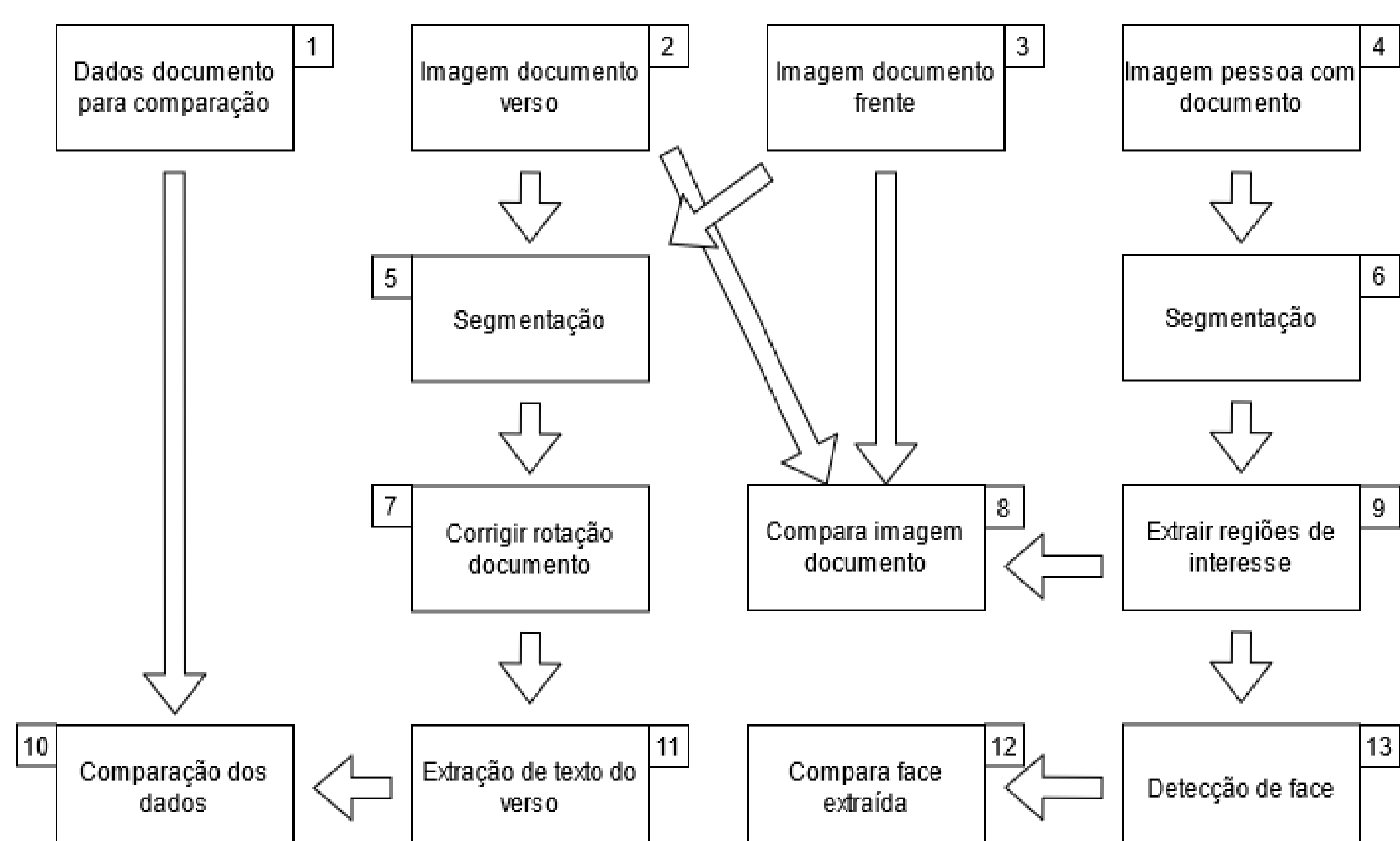
O sistema diminui a necessidade de uma pessoa fazendo o trabalho de verificação de identidade ele diminui o tempo e custos para verificação de identidade.

Descrição Validar a identidade é uma forma de verificar as informações fornecidas pelo usuário a fim de analisar se a pessoa é, de fato, quem se apresenta no ato do cadastro, a partir disso muitas empresas possuem setores com uma infraestrutura e pessoal apenas para receber documentos e validar os dados de forma manual de um cliente e aprovar o cadastro e afim de diminuir custos operacionais e poupar muito tempo no processo, propomos um sistema de validação digital a partir de uma self do cliente e uma foto de seu documento retornando uma análise em segundos validando e finalizando o cadastro.

Em testes foi obtidos bons resultados, com IOU Score de 94% na segmentação de imagens, 62% de acurácia na extração de texto dos documentos e uma acurácia de 40% em semelhança de rostos em idades diferentes e 68,3% em idades mais próximas.

O sistema diminui a necessidade de uma pessoa fazendo o trabalho de verificação de identidade ele diminui o tempo e custos das empresas para verificação de identidade.

Metodologia Utilizada A proposta metodológica deste sistema consiste em treze etapas. A primeira etapa, como na figura de 1 a 4, representa a entrada dos dados usados para verificar e autenticar a pessoa. O bloco 1 representa os dados já conhecidos da pessoa que está se autenticando, o segundo e o terceiro bloco representam a entrada das imagens da frente e verso do documento usado para autenticação e o quarto bloco representa a imagem da pessoa segurando o mesmo documento inserido nos blocos 2 e 3



Em seguida, nos blocos 5 e 6, é feito a segmentação das imagens para poder realizar a extração das regiões de interesse que serão utilizadas nos blocos 7, 8 e 9. No bloco 7 será feita a correção da rotação do documento, usando a máscara gerada no bloco 5. No bloco 8, será feita a comparação simples entre o recorte das imagens do documento com o recorte do documento extraído no bloco 9. No bloco 10 será feita a comparação entre os dados conhecidos e os dados extraídos do documento e, por fim, nos blocos 12 e 13 será feita a identificação do rosto na selfie e a comparação desse rosto com o rosto encontrado na segmentação da imagem recebida no bloco 3.

Resultados Neste projeto, foram alcançados 4 resultados diferentes, referentes à segmentação de imagem com redes neurais, extração de texto em documentos oficiais, comparação entre dois documentos oficiais com o uso de métricas de similaridade e comparação de rostos através da semelhança de vetores de características.

Segmentação de Imagem foi obtido um *IOU Score* de 94%.



Entrada

Esperado

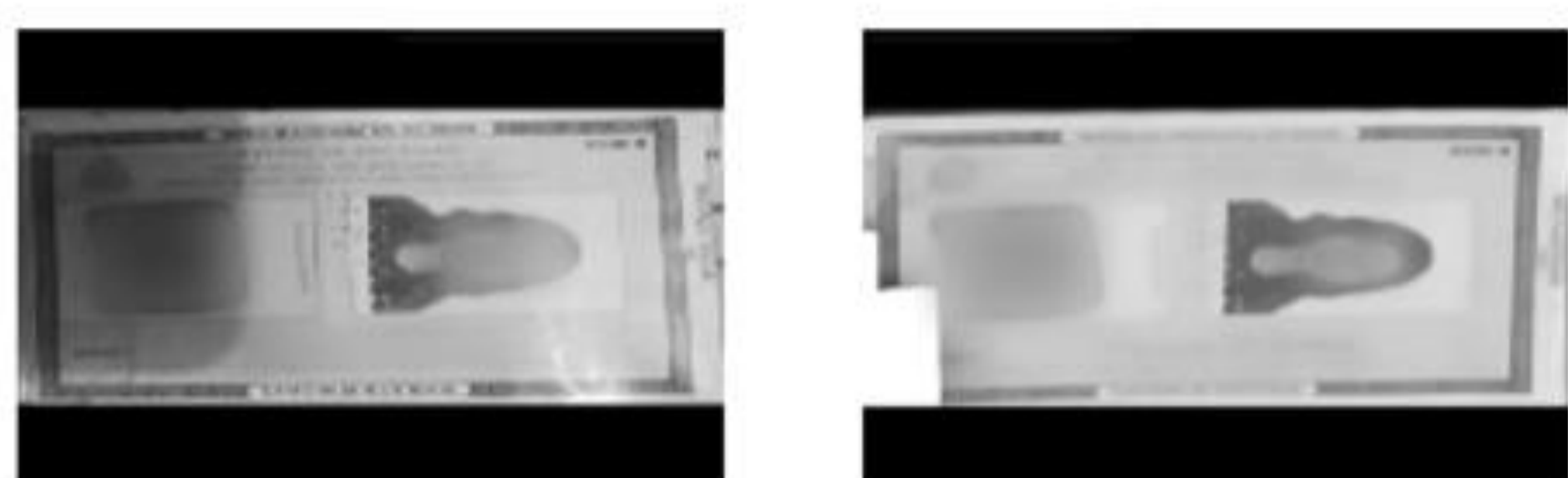
Saída

Considerando apenas resultados com extração perfeita ou extrações nas quais o índice de similaridade ficou acima de 90%, a metodologia obteve um acurácia de 62%

Resultado da extração	Quantidade
Extrações perfeitas	44
Não reconheceu	13
Aproximações	47

Comparação de Documentos, quando os documentos estão em condições relativamente semelhantes, o índice de similaridade estrutural(SSIM) apresenta bons resultados, enquanto, em geral, o erro quadrático médio apresenta valores muito altos mesmo quando o SSIM indica haver pelo menos 50% de semelhança.

MSE: 3279.56, SSIM: 0.78



Comparação de Rostos obteve uma acurácia de 65,5%, após a extração dos vetores de características, foi calculada a similaridade usando o algoritmo de similaridade de cossenos (Seção 3.5.3). Nos testes realizados pode-se observar que a idade da foto do documento em relação a idade atual da pessoa influencia bastante nos resultados, assim como o estado de conservação do documento também faz diferença.

Em documentos mais antigos e/ou danificados, a média no índice de similaridade ficou abaixo de 40%. Em documentos com fotos mais recentes e em melhor estado, a média sobe para 68,3%



Imagem documento



80,9% de semelhança



79,4% de semelhança

Conclusão o sistema é capaz de validar a identidade de uma pessoa a partir de dados e imagens fornecidas por usuários e que uma grande parte das pessoas não renovam seus documentos de identificação, o que dificultou a validação.