

TÉCNICAS DE RESTAURAÇÃO E PRESERVAÇÃO DE FOTOGRAFIAS E DOCUMENTOS DIGITAIS DETERIORADOS

Vitor Hugo Evaldt Silveira¹, Marlon de Matos de Oliveira²

Resumo: O trabalho aborda as técnicas de restauração e preservação de fotografias e documentos digitais deteriorados, com foco em métodos de processamento de imagens que buscam recuperar a legibilidade e integridade de materiais afetados pelo tempo. O objetivo foi comparar diferentes algoritmos de redução de ruído e binarização, analisando sua eficiência na restauração de documentos históricos, utilizando como base o conjunto de dados *Document Image Binarization Contest* (DIBCO). A pesquisa é de caráter descritivo e bibliográfico, com aplicação prática de algoritmos implementados em Python por meio da biblioteca OpenCV, sendo estes: Nível Médio de Cinza, Porcentagem de Preto, Dois Picos, Niblack, Savoula, Otsu e White e Rohrer, contemplando as etapas de pré-processamento, filtragem e análise quantitativa e qualitativa. Os resultados demonstraram que os métodos de binarização global, especialmente o de nível médio de cinza, Otsu e Dois Picos, obtiveram desempenho superiores nas métricas *Peak Signal-to-Noise Ratio* (PSNR) e *Structural Similarity Index Measure* (SSIM), preservando com maior fidelidade a estrutura e o contraste dos textos em relação às imagens originais. Já os métodos adaptativos, como Niblack e White-Rohrer, apresentaram menor precisão em contextos de baixo contraste e ruído acentuado. Conclui-se que, mesmo com abordagens simples, é possível alcançar resultados quantitativamente sólidos na restauração digital, contribuindo para a preservação e o acesso a acervos históricos, embora o estudo aponte limitações quanto à escala e à variedade de documentos analisados.

Palavras-chave: processamento de imagens; preservação digital; binarização; restauração de documentos; visão computacional.

¹Curso de Ciência da Computação, Universidade do Extremo Sul Catarinense (Unesc), vitorhgtt@gmail.com

²Orientador. Curso de Ciência da Computação, Universidade do Extremo Sul Catarinense (Unesc), marlon.oliveira@unesc.net

ABSTRACT: The study addresses techniques for the restoration and preservation of deteriorated digital photographs and documents, focusing on image processing methods aimed at recovering the readability and structural integrity of time-damaged materials. The main objective was to compare different noise reduction and binarization algorithms, analyzing their efficiency in restoring historical documents using the Document Image Binarization Contest (DIBCO) dataset. The research follows a descriptive and bibliographic approach, combined with practical experimentation through the implementation of algorithms in Python using the OpenCV library, these are: Medium Gray Level, Percentage of Black, Two Peaks, Niblack, Savoula, Otsu, White and Rohrer, encompassing preprocessing, filtering, and quantitative and qualitative analysis. The results showed that global binarization methods, especially the mean gray level, Otsu, and Two Peaks, achieved the best performance in Peak Signal-to-Noise Ratio (PSNR) e Structural Similarity Index Measure (SSIM) metrics, preserving text structure and contrast more effectively than adaptive methods such as Niblack and White-Rohrer, which showed lower accuracy in low-contrast and high-noise conditions. The findings indicate that even with simpler approaches, it is possible to achieve quantitatively sound results in digital restoration, contributing to the preservation and accessibility of historical archives, although the study highlights limitations related to the scope and diversity of the analyzed documents.

Keywords: image processing; digital preservation; binarization; document restoration; computer vision.

1 INTRODUÇÃO

A preservação é a aquisição, organização e distribuição de recursos a fim de que venham a impedir posterior deterioração ou renovar a possibilidade de utilização de um seletor grupo de materiais (Conway, 2001). De acordo com Baggio (2013), os documentos são frágeis e não existe segurança suficiente para garantir sua preservação digital a longo prazo. Essa fragilidade pode estar em sua rápida degradação física, na obsolescência tecnológica, na complexidade e nos custos. Com isso posto, compreende-se como uma problemática a degradação em níveis e fontes diversas que poderiam dificultar a compreensão de documentos afetados.

Segundo Gonzales e Woods (2000), a maioria das funções de processamento de imagens pode ser implementada em software. Em outras palavras, chama-se de processamento de imagens a área na qual a

entrada é uma imagem, que será processada geralmente por procedimentos expressos de forma algorítmica, e cujo resultado de saída será uma nova imagem processada.

O processamento de imagens é fundamental tanto para a otimização visual de imagens digitais, facilitando sua interpretação por seres humanos, quanto para a automação do reconhecimento de informações. Além disso, métodos como a binarização adaptativa e técnicas híbridas têm se mostrado eficazes para otimizar a análise de documentos digitais, conforme mencionado por Vaz, Araújo e Varjão (2011).

A visão computacional, de acordo com Milano e Honorato (2010), pode ser entendida como a área da ciência que busca possibilitar que computadores interpretem o ambiente ao seu redor, extraindo informações relevantes de imagens obtidas por câmeras, sensores e outros dispositivos, de modo a reconhecer, manipular e analisar os elementos presentes nelas.

Os processos de visão computacional frequentemente começam com uma etapa de pré-processamento baseada no processamento de imagens. Essa fase pode incluir a conversão das imagens para formatos ou tamanhos específicos e a aplicação de filtros para a remoção de ruídos gerados durante a captura da imagem. Os ruídos podem ser provenientes de diversas fontes, como o tipo de sensor usado, a iluminação no ambiente, as condições climáticas no momento da aquisição ou até mesmo a posição relativa entre o objeto e a câmera. Além disso, destaca-se que os ruídos não se limitam à interferência no sinal da imagem, mas também abrangem elementos que dificultam a interpretação ou o reconhecimento de objetos presentes. Como exemplificado por Marengoni (2009), filtros são ferramentas fundamentais para minimizar esses ruídos, especialmente aqueles surgidos durante o processo de aquisição.

A redução de ruídos é uma técnica criada como uma solução prática para recuperar a imagem original. Algoritmos que aplicam essa tecnologia são denominados *denoising methods*, ou métodos de remoção de ruído. Essas ferramentas são presentes em aplicações de compressão de dados, no reconhecimento de padrões e análise de exames médicos. Esses métodos buscam preservar ao máximo as bordas e outras características essenciais da imagem enquanto eliminam as degradações existentes (Jain; Tyagi, 2014).

O *Non-Locals Means* é um algoritmo de remoção de ruído não-local, que utiliza informação global da imagem para filtrar cada pixel individualmente, foi desenvolvido e apresentado no artigo de Buades, Coll e

Morel (2011). Ele busca explorar padrões ou características que se repetem ao longo da imagem como um todo e desde sua publicação encontra-se em destaque na área de remoção de ruídos e diversos autores estudaram e propuseram melhorias no seu funcionamento. Entre os algoritmos de binarização global estão: o tom de cinza médio, define como ponto de corte o ponto médio dos valores de luminância presentes na imagem, o algoritmo de porcentagem de preto, onde o ponto de corte é calculado para que a imagem final tenha uma porcentagem pré-definida de tons preto apenas, e o algoritmo de dois picos, que calcula o ponto de corte como o ponto entre os dois tons mais presentes na imagem. A Figura 1 demonstra a utilização desses filtros para redução de ruídos.

Figura 1 - Imagem com ruído à esquerda, e filtrada na direita.



Fonte: Adaptado de Marengoni (2009).

Segundo Mello (2013), o processo de binarização trata da conversão de uma imagem para apenas duas cores. Geralmente, ao tratar de documentos, essa conversão resulta em tons de branco, referentes ao papel, ou pretos, referentes ao conteúdo. Isso é feito definindo um ponto de corte nas cores onde tons menores do que esse valor se tornam preto e tons maiores em branco. Existem classes diferentes de algoritmos de binarização que podem ser usados de acordo com o problema. De modo geral, os algoritmos podem ser classificados como globais, ponto de corte único para toda a imagem, ou locais, divisão em regiões e um ponto de corte para cada região. O processo de binarização de imagens transforma uma imagem colorida em uma nova imagem com apenas duas cores, representando o fundo e o objeto de interesse, sendo uma das etapas de aplicações como Reconhecimento Óptico de Caracteres (OCR) e reconhecimento de texto manuscrito.

Segundo Gonzales e Woods (2000), a binarização é considerada uma técnica essencial no processamento de imagens digitais, sendo um passo importante em várias aplicações de visão computacional, especialmente em tarefas de segmentação de imagens, detecção de objetos e reconhecimento de padrões.

Figura 2 - Documento escurecido pela ação do tempo e aplicado o processamento de imagem com binarização.



Fonte: Adaptado de Roe (2015).

Na Figura 2, pode-se entender como um processamento da imagem é capaz de preservar as informações de um documento deteriorado. Diferentes algoritmos de binarização têm sido propostos nas últimas décadas, contudo, ainda não há uma solução consolidada para o problema, principalmente para imagens que possuem um fundo complexo. Isso deixa o campo de pesquisa aberto a inovações e melhorias.

O objetivo deste trabalho é comparar técnicas de restauração e preservação de fotografias e documentos digitais. Busca-se entender o processamento e restauração de fotografias e documentos digitais, as técnicas existentes e suas limitações, os algoritmos de redução de ruído e de binarização e, por fim, comparar a eficiência das técnicas utilizadas. Para isso, propõe-se uma análise comparativa dos algoritmos de redução de ruído e binarização estudados, para avaliar seus desempenhos no processamento de documentos deteriorados e verificar a eficiência na restauração.

Este artigo está ordenado em cinco seções, na primeira seção é apresentado o escopo do projeto, contendo a introdução que contextualiza o trabalho a ser realizado mostrando os objetivos, a justificativa da realização desse projeto e a fundamentação teórica que esta pesquisa usa de base. Na segunda seção, são descritos os trabalhos correlatas à esta pesquisa. Na terceira seção, são apresentados os materiais e métodos utilizados para a concretização e replicação deste trabalho. Na quarta seção, tem-se os resultados e discussões dos dados obtidos através da aplicação da metodologia proposta. Por fim, a quinta seção contém a conclusão e sugestões para trabalhos futuros.

2 TRABALHOS CORRELATOS

A dissertação de mestrado de Pimenta (2009) na Pontifícia Universidade Católica do Paraná (PUCPR), tem origem na área de Computação Forense. A pesquisa propôs uma abordagem para a reconstrução de documentos fragmentados e realizou um estudo comparativo com o método de referência de Solana (2005) para validar sua eficácia em termos de precisão e robustez e uma versão modificada do algoritmo de Prim. Os aspectos computacionais centrais se concentram no uso de Visão Computacional para extrair características geométricas do contorno dos fragmentos, que são então utilizados em um algoritmo de Programação Dinâmica para encontrar a sequência ótima de junção dos pares de fragmentos. Como resultado, a metodologia proposta alcançou uma precisão de 81% na combinação de fragmentos, demonstrando um ganho mensurável na robustez do sistema ao reduzir em pelo menos 4% a taxa de erros e o número de falsos candidatos a parceiros em comparação com o método referencial, confirmando sua capacidade de reconstruir documentos completos, mesmo que com algumas limitações em casos de fragmentação extrema.

O trabalho de Nogueira e Amorim (2019) do Laboratório de Estudos Avançados em Cidades, Arquitetura e Tecnologias Digitais (LCAD) da Universidade Federal da Bahia, aborda a reconstrução digital como meio de preservação da herança cultural. Com base em métodos científicos e rigor histórico, os autores exploram como modelos digitais podem representar e transmitir valores materiais e imateriais de bens culturais. A pesquisa destaca que, apesar dos desafios técnicos e éticos, essas reconstruções digitais podem ir além da recuperação visual, constituindo-se como ferramentas para conservação e investigação histórica, ampliando o alcance de objetos culturais e permitindo múltiplas interpretações de dados históricos.

A tese de doutorado de Roe (2015) publicado na Universidade Federal de Pernambuco, tem origem na área de ciência da computação e processamento de imagens. A metodologia desta tese de doutorado envolveu a proposição e o desenvolvimento de três novos algoritmos de processamento de imagens, dois de binarização e um de restauração de cores, seguidos por uma avaliação comparativa de seu desempenho contra os métodos de ponta existentes na literatura. Os aspectos computacionais estão ancorados no Processamento Digital de Imagens (PDI), com destaque para algoritmos de binarização aprimorados, incluindo o uso de equalização local e XDoG, e um método de restauração de cores, além da imple-

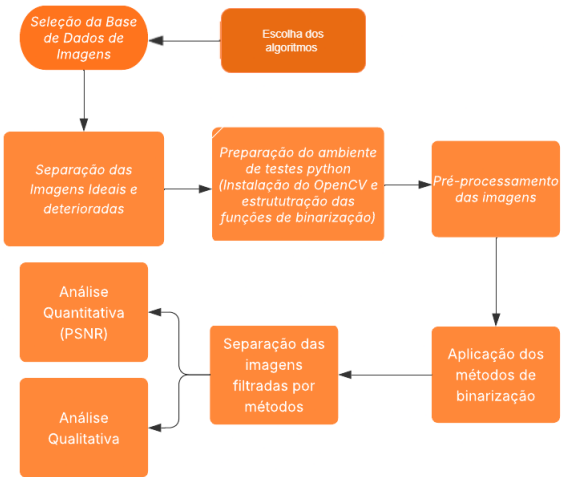
mentação de um novo filtro bilateral para binarização de documentos com elementos gráficos. Como resultado, o algoritmo de binarização proposto demonstrou desempenho superior em comparação com os métodos atuais, exibindo grande adaptabilidade a diferentes tipos de documentos e fontes, sem a necessidade de reajustes de parâmetros, cumprindo o objetivo de preservação digital.

3 MATERIAIS E MÉTODOS

O desenvolvimento do trabalho sobre técnicas de restauração e preservação de fotografias e documentos digitais deteriorados consiste em uma pesquisa descritiva, bibliográfica e com base tecnológica, que possui o objetivo de utilizar algoritmos de filtragem e redução de ruído e binarização, juntamente com o pré-processamento das imagens da base escolhida. Juntamente, a biblioteca OpenCV do Python foi utilizada para a implementação do trabalho, com uso dela foram aplicadas as técnicas e obtidos os dados que serão relevantes para a análise final dos resultados. Com base nesses dados obtidos, é discutida uma conclusão onde são comparadas a eficiência das técnicas estudadas e em quais situações cada uma delas possui um desempenho ou não que as de mais.

A metodologia desse trabalho consiste no levantamento bibliográfico, busca de trabalhos correlatos, escolha de uma base de imagens para testes, escolha dos algoritmos utilizados, aplicação de métodos de pré-processamento das imagens, utilização das técnicas estudadas, testes, análise e validação dos dados obtidos. O fluxograma que demonstra a metodologia dessa pesquisa pode ser visto na Figura 3.

Figura 3 - Fluxograma da metodologia aplicada.



Fonte: O autor (2025).

3.1 BASE DE DADOS

Para a avaliação quantitativa e qualitativa dos algoritmos de binarização, foi utilizado o conjunto de dados da competição *Document Image Binarization Contest* (DIBCO) do ano de 2017. A base de dados escolhida consiste em 20 imagens de documentos, dos quais 10 são manuscritos e 10 são impressos, com diferentes graus de preservação digital que serão submetidas ao processo que se visa com essa pesquisa.

Das características da base de dados, nota-se que as imagens de entrada originais são coloridas, a imagem de menor resolução possui 351 X 292 pixels, enquanto a maior imagem possui 2397 X 1350 pixels. Na totalidade, o tamanho da base é de 27.910.150 pixels tendo cada imagem em média 1.395.507,5 pixels. A base de imagens *open source* está disponível no DUTH-Visual Computing Group³.

O pré-processamento tratou dos dados brutos das imagens onde iniciou-se a calibração e correção de distorções nas imagens para a preparação dos algoritmos que foram aplicados sequencialmente. Nessa pesquisa, utilizou-se a escala de cinza (*grayscale*) para eliminar alguns ruídos iniciais. Para converter uma imagem em escala de cinza foi extraído a média dos três canais da imagem, ou seja, dividiu-se por 3 a soma dos valores RGB da imagem.

3.2 UTILIZAÇÃO DOS ALGORITMOS

Após o pré-processamento, foram aplicados os algoritmos, começando pela redução de ruído sendo utilizado o *Non-Local Means* em que o algoritmo usa da média ponderada de todos os pixels da imagem para calcular o valor do pixel na imagem filtrada, e seguido pela binarização onde são explorados os diferentes métodos vistos escolhidos, nível médio de cinza, porcentagem de preto, dois picos, Niblack, Savoula, White e Rohrer, Otsu. Cada algoritmo funciona de um modo diferente, justifica-se a utilização destes da seguinte forma, Nível Médio de Cinza é um método simples, rápido e fornece um ponto de referência básico. Útil para comparar métodos mais sofisticados. Aplica-se bem quando a iluminação é uniforme e o contraste entre fundo e texto é moderado. O algoritmo da porcentagem de preto permite controle direto sobre a densidade de pixels classificados como preto, sendo útil em imagens com grande variação de intensidade ou quando se busca manter a coerência visual em conjuntos de imagens. É frequentemente aplicado em digitalização de documentos com degradação

³<https://vc.ee.duth.gr/dibco2017/benchmark/>

não uniforme. O método dos Dois Picos é utilizado quando o histograma é claramente bimodal, esse método obtém limiares robustos e precisos. É indicado para documentos com bom contraste entre tinta e papel, mesmo em condições moderadas de ruído. Em Niblack, busca-se imagens com iluminação não uniforme, sombras e manchas típicas de documentos deteriorados. É sensível ao ruído, mas preserva bem detalhes de regiões escuras ou com variações locais significativas. O algoritmo de Savoula é um métodos utilizado para documentos antigos, com manchas, fungos, rasuras ou degradação severa. Mantém detalhes e reduz falsos positivos, especialmente em áreas claras ou degradadas. White e Rohrer proporciona a separação entre texto e fundo em documentos com estrutura fina, ruído de textura e papel envelhecido. Por último, o método de Otsu funciona quando existe separação nítida entre fundo e objeto. É rápido, robusto e serve como referência padrão em experimentos acadêmicos e industriais. É frequentemente utilizado em trabalhos de binarização.

Por fim, tem-se a análise quantitativa da Relação Pico Sinal-Ruído (PSNR) e qualitativa do *Structural Similarity Index Measure* (SSIM) dos resultados, onde são comparadas as imagens originais com as suas versões filtradas, que é a base para a validação da eficiência dos algoritmos em cada situação por meio dos testes propostos.

No âmbito da análise quantitativa, O PSNR mede a similaridade entre duas imagens. Na binarização, ele quantifica o erro ou a perda de qualidade entre a imagem ideal (*Ground Truth*) e a imagem gerada pelo seu algoritmo (*Output Binarizado*).

O PSNR é expresso em decibéis (dB) e definido pela seguinte equação:

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX^2}{MSE} \right)$$

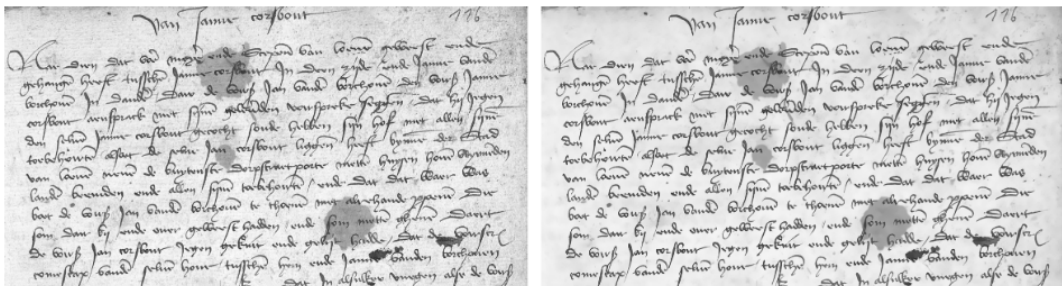
Onde MAX é o valor máximo de pixel possível, para imagens binarizadas em 8 bits, tons de cinza, onde 0 é preto e 255 é branco, MAX = 255. E MSE (*Mean Squared Error*) é a média dos erros quadráticos entre a imagem de referência e a imagem binarizada.

A análise SSIM é uma métrica de avaliação de qualidade de imagem que mede o grau de similaridade estrutural entre uma imagem processada e uma imagem de referência. Diferente de métricas puramente baseadas em erro, como o MSE, o SSIM leva em consideração aspectos perceptuais importantes, avaliando três componentes principais: luminosidade, contraste e estrutura local. Em imagens resultantes de métodos

de binarização, o SSIM é utilizado para quantificar o quanto a versão binarizada preserva a estrutura e o padrão do texto original em relação à imagem de referência, geralmente a máscara *ground truth*. Um valor de SSIM próximo de 1 indica alta similaridade estrutural, ou seja, o método de binarização conseguiu manter os traços e contornos do texto sem distorções significativas. Valores mais baixos, por outro lado, indicam perda de detalhes, ruído residual ou degradação da estrutura textual.

Os métodos de binarização estudados possuem características próprias que os tornam eficazes para diferentes situações de aplicabilidade. Antes de submeter as imagens aos métodos, todas elas coloridas, depois convertidas para escala de cinza passam pela função `cv2.fastNlMeansDenoising`. Nessa etapa, as imagens da base escolhida possuem ruído característico do papel antigo, como manchas e granulação. O algoritmo de *Non-Local Means* funciona comparando *patches*, pequenas janelas, em toda a imagem. Ao invés de somente suavizar, ele remove o ruído do fundo, o que poderia ser confundido com texto nos métodos globais, sem desfocar as bordas finas dos caracteres. O resultado é uma imagem em escala de cinza mais nítida, onde a transição entre o traço do texto e o fundo do papel é mais clara, preparando as imagens para uma binarização mais precisa. Pode-se visualizar o funcionamento da redução de ruído na Figura 4

Figura 4 - À esquerda imagem original, à direita aplicação da função.



Fonte: O autor (2025).

3.2.1 Nível médio de cinza

O algoritmo de nível médio de cinza proposto por (Parker, 1997) é utilizado para definir o limiar I , ele usa apenas os valores limite do histograma de uma imagem em escala de cinza. O cálculo é feito por:

$$I = \frac{\max(H) - \min(H)}{2}$$

onde $\max(H)$ é o maior e $\min(H)$ é o menor nível de cinza presente no histograma imagem. Apesar da sua simplicidade, este método não é eficiente

para vários tipos de imagens, como, por exemplo, imagens de baixo contraste. No algoritmo, ele foi implementado definindo limiar (l) como a média simples entre o valor máximo ($\max(H) = 255$) e o mínimo ($\min(H) = 0$) dos níveis de cinza, resultando em $l = 127.5$.

3.2.2 Porcentagem de preto

De acordo com Srinivasan et al. (2024), o algoritmo de porcentagem de preto assume que a imagem binarizada final possuirá uma certa quantidade de pixels pretos. Assim, é necessário definir essa quantidade *a priori*. O histograma da imagem é calculado e o limiar é encontrado pelo valor que produz a quantidade de pixels pretos definida inicialmente.

No código foi definido o limiar (l) ao percorrer o histograma e somar as contagens de pixels, de 0 a 255, até atingir um percentual pré-definido, por exemplo, 0.4, ou 40%, do total de pixels da imagem. A premissa foi de garantir que o volume de texto esperado seja preservado.

3.2.3 Dois Picos

O código implementado buscou isolar dois picos no histograma, texto e fundo, e calcular o limiar no ponto mais baixo entre eles. A estratégia para encontrar o segundo pico usa a lógica da maximização da distância quadrática. Quando uma imagem possui duas classes de tons de cinza separáveis, escuras e claras, o limiar ocorre no ponto mais baixo entre os dois picos do seu histograma. O algoritmo dois picos age nesse e seleciona automaticamente o seu limiar em dois passos: procura os dois picos do histograma, e encontra o ponto mais baixo entre eles. De acordo com Sezgin e Sankur (2004), o primeiro pico é encontrado facilmente, pois ele é o ponto do histograma com maior valor. O problema é que o segundo maior ponto do histograma geralmente fica ao lado do primeiro maior. Então, localizar o segundo pico torna-se mais difícil do que o primeiro. A estratégia usada define o segundo pico como sendo:

$$\max \left\{ (k - j)^2 \cdot H(k) \mid 0 \leq k \leq 255 \right\}$$

onde H é o histograma que tem 256 tons de cinza [0-255], e o primeiro pico está situado no nível j do histograma.

3.2.4 Niblack

Niblack (1986) propôs um algoritmo de limiarização dinâmico que divide a imagem de entrada em sub-imagens de tamanho M por N .

Para cada sub-imagem, o algoritmo define o limiar l como sendo:

$$l(x, y) = \mu + k \cdot \sigma$$

onde μ e σ são a média local e o desvio padrão da sub-imagem centrada no $pixel(x, y)$. As dimensões de cada sub-imagem, s e k são parâmetros pré-definidos e variáveis em cada caso. Esse método pode separar o texto do fundo em áreas ao redor do texto, contudo o seu principal problema é que, para sub-imagens em que não há nenhum texto, algumas partes do fundo são classificadas como texto, gerando um ruído considerável no fundo.

Na implementação se aplicou o limiar adaptativo $l(x, y) = \mu + k \cdot \sigma$, onde μ e σ são calculados em uma janela, `window_size=25`. Isso é notável pois se ajusta às mudanças de brilho do papel. No entanto, ele ainda pode gerar ruído no fundo porque em áreas sem texto, o desvio padrão (σ) é baixo, fazendo com que o limiar $\mu + k \cdot \sigma$ fique muito próximo da média do fundo, classificando pixels ligeiramente mais escuros como texto.

3.2.5 Savoula

O algoritmo proposto por (Sauvola; Pietikainen, 2000) inseriu um ponderador para o desvio padrão. Assim, o limiar foi definido como :

$$l(x, y) = \mu + k \cdot \frac{\sigma}{R}$$

onde μ e σ são como em Niblack e R é uma constante responsável por mudar o valor do desvio padrão. Sauvola ainda sugeriu os valores de $R = 128$ e $k = 0.5$. Com a mudança, se produz uma imagem binária em caso de contraste alto entre fundo e texto. Contudo, os valores ótimos de R e k são proporcionais ao contraste do texto, de forma que, para imagens com baixo contraste, se os valores não forem definidos corretamente, acarreta em uma perda das regiões de texto.

Foi utilizada a melhoria $l(x, y) = \mu + k \cdot (\sigma/R)$, com $R = 128$ e $k = 0.5$. A introdução do termo de normalização R mitiga o problema de ruído do Niblack. Nas imagens da selecionadas, o limiar é forçado a permanecer próximo de μ em regiões de fundo limpo (σ pequeno), inibindo o ruído. Simultaneamente, em áreas de alto contraste, ele se comporta como Niblack, preservando as bordas.

3.2.6 White e Rohrer

White e Rohrer (1983) propuseram um método local que classifica o pixel de acordo com a média dos valores dos pixels da sua vizinhança. Assim, dada uma janela de tamanho centrada no pixel $p(x,y)$, a classificação para cada pixel é feita da seguinte forma:

$$p_{(x,y)} = \text{objeto}, \text{ se } m_j > p_{(x,y)} \cdot k$$

$$p_{(x,y)} = \text{fundo}, \text{ caso contrário}$$

onde m_j é a média dos pixels da janela e k é um peso auxiliar definido a priori. O pixel classificado como objeto recebe a cor preta, e o classificado como fundo, a cor branca.

Durante o processo, classificou-se um pixel como texto objeto se $p_{(x,y)} \cdot k < m_j$, onde m_j é a média local e k é um peso auxiliar $k=1.0$. Nas imagens escolhidas, este método fornece um limiar adaptativo rápido e razoavelmente eficaz. Ele assume que o texto é significativamente mais escuro que a média local. É menos sensível a flutuações de ruído que o Niblack, mas pode ser menos preciso que o Sauvola na preservação de traços finos em áreas de baixo contraste.

3.2.7 Otsu

O método de Otsu (1979) é um dos mais referenciados na literatura. Nele, considera-se que o histograma de uma imagem possui duas classes C_0 e C_1 . Dada uma imagem com T níveis de cinza, o número de pixels no nível i é denotado por n_i e o número total de pixels por $N = n_1 + n_2 + \dots + n_i$. Assim é calculada a distribuição de probabilidade através de uma normalização do histograma da imagem em tons de cinza:

$$P_i = \frac{n_i}{N}, \quad P_i > 0, \quad \sum P_i = 1$$

Então, é necessário encontrar o limiar no nível que divida os pixels nas duas classes C_0 e C_1 , dado que os níveis dos pixels de C_0 estão entre $[1, \dots, l]$ e os de C_1 entre $[l + 1, \dots, T]$. Logo, a probabilidade da ocorrência da classe ω_c e a média dos seus níveis de cinza μ_c são dados por:

$$\omega_0 = \sum_{k=1}^l p_k = \omega(l)$$

$$\omega_1 = \sum_{i=l+1}^T p_i = 1 - \omega(l)$$

e

$$\mu_0 = \sum_{k=1}^l k \cdot p_k / \omega_0 = \mu(l) / \omega(l)$$

$$\mu_1 = \sum_{i=l+1}^T i \cdot p_i / \omega_1 = \frac{\mu_T - \mu(l)}{1 - \omega(l)}$$

onde

$$\omega(l) = \sum_{l=1}^l p_l$$

e

$$\mu(l) = \sum_{l=1}^l i \cdot p_l$$

O nível médio total da imagem original é dado por:

$$\mu_T = \sum_{l=1}^T l \cdot p_l$$

Desse modo, para qualquer l escolhido, é possível verificar que:

$$\omega_0 \cdot \mu_0 + \omega_1 \cdot \mu_1 = \mu_T, \quad \omega_0 + \omega_1 = 1$$

As variâncias das classes são dadas por:

$$\sigma_0^2 = \sum_{l=1}^l (i - \mu_0)^2 \cdot \frac{p_l}{\omega_0}$$

$$\sigma_1^2 = \sum_{l=l+1}^T (i - \mu_1)^2 \cdot \frac{p_l}{\omega_1}$$

Para avaliar o limiar ótimo, Otsu analisou as seguintes medidas de separabilidade de classes:

$$\lambda = \frac{\sigma_B^2}{\sigma_W^2}, \quad K = \frac{\sigma_T^2}{\sigma_W^2}, \quad \eta = \frac{\sigma_B^2}{\sigma_T^2},$$

onde

$$\sigma_W^2 = \omega_0 \cdot \sigma_0^2 + \omega_1 \cdot \sigma_1^2$$

$$\sigma_B^2 = \omega_0 \cdot (\mu_0 - \mu_T)^2 + \omega_1 \cdot (\mu_1 - \mu_T)^2 = \omega_0 \cdot \omega_1 \cdot (\mu_1 - \mu_0)^2$$

$$\sigma_T^2 = \sum_{i=l+1}^T (i - \mu_T)^2 \cdot p_i$$

são a variância intraclasse, variância entre classes e a variância total dos níveis de cinza da imagem original, respectivamente. Então, a otimização do limiar l é feita através da maximização de uma das medidas de separabilidade de classes acima. Devido η ser a medida mais simples entre as três, é adotado o seu uso. Portanto, o valor que maximiza η , e consequentemente maximiza σ_B^2 , é tido como o limiar ótimo l^* para o algoritmo:

$$\sigma_B^2(l^*) = \max(\sigma_B^2(l)).$$

Uma das principais características desse método é a sua autonomia, visto que ele não faz uso de nenhum tipo de ajuste de parâmetros, nem de um observador para guia-lo ao melhor resultado. Além disso, ele pode ser aplicado para resolver qualquer caso em que há duas classes. Contudo, o algoritmo não é adequado para imagens com muita diferença de variâncias das classes.

Durante a aplicação do código, utilizou-se a implementação otimizada do OpenCV `cv2.THRESH_OTSU`, onde ele itera por todos os possíveis limiares para maximizar a variância entre classes σ_B^2 , o que teoricamente separa melhor os grupos de pixels escuros, texto, dos claros, fundo. No entanto, nas imagens da base de dados com variação de iluminação, o histograma é complexo e o limiar de Otsu acaba sendo influenciado pelas características do ruído de fundo.

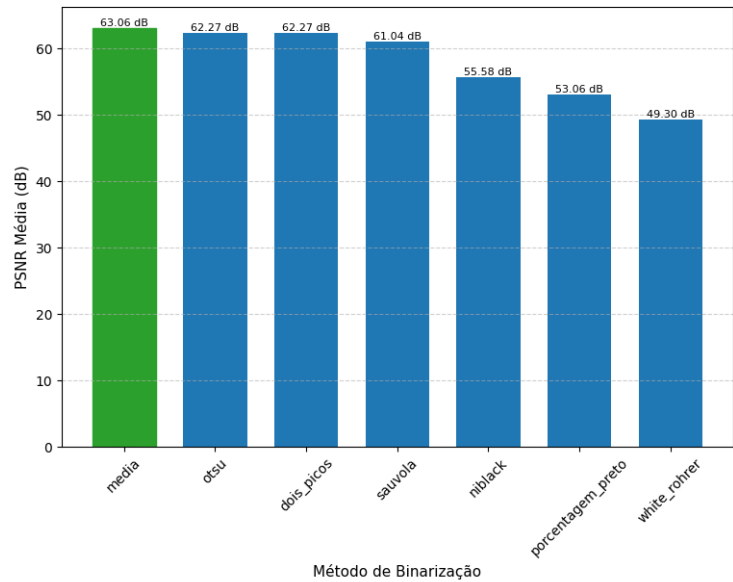
Todas as etapas descritas foram aplicadas localmente e as imagens foram submetidas individualmente aos métodos através de um algoritmo feito pelo autor implementando cada um dos modelos estudados.

4 RESULTADOS E DISCUSSÃO

As métricas de análise dos dados obtidos foi fundamental para expor de forma objetiva e perceptível aquilo que a pesquisa se propôs. Uma vez que os resultados se tratam de imagens, a mera abstração de dados quantitativos, como no PSNR, podem ir ao encontro da percepção daquele

que visualiza as imagens, como pode-se visualizar nas imagens obtidas pela aplicação dos métodos presentes no Apêndice A. Por isso, a comparação com o SSIM se mostrou muito mais correlacionável com a subjetividade das interpretações humanas. Pode-se verificar as médias do desempenho quantitativo no gráfico da Figura 5.

Figura 5 - Gráfico do desempenho quantitativo médio (A)



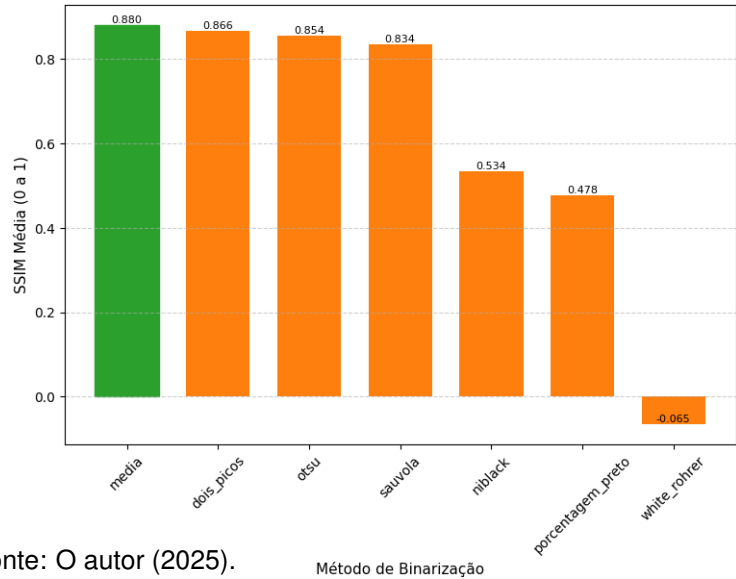
Fonte: O autor (2025).

Os resultados da análise de PSNR indicam que o método de binarização pelo nível médio de cinza apresentou desempenho quantitativo superior, alcançando 63,06 dB, seguido pelos métodos Otsu e Dois Picos, ambos com 62,27 dB. Esses valores evidenciam que tais métodos preservaram a intensidade e a fidelidade do sinal original após a binarização, com menor degradação visual. Em contrapartida, os métodos White-Rohrer e Porcentagem de Preto obtiveram os menores índices, 49,30 dB e 53,06 dB, respectivamente, sugerindo maior perda de informação e aumento do ruído introduzido pelo processo de limiarização. De modo geral, métodos globais mostraram-se mais consistentes em ter relação sinal-ruído do que os adaptativos, como Niblack e Sauvola.

Na análise de SSIM, o método de binarização pela média também se destacou, alcançando o maior índice de similaridade (0,880), indicando alta correspondência estrutural com a imagem de referência. Os métodos Dois Picos, Otsu e Sauvola apresentaram desempenhos próximos, com valores entre 0,834 e 0,866, o que demonstra manutenção das estruturas textuais e bordas. Em contrapartida, os métodos Niblack e Por-

centagem de Preto tiveram desempenho inferior (0,534 e 0,478), e, o método White-Rohrer apresentou resultado negativo (-0,065), refletindo forte distorção estrutural. Esses resultados reforçam que métodos globais tendem a preservar a estrutura visual do texto em imagens de documentos, especialmente em contextos com baixo ruído residual. Pode-se verificar esses resultados no gráfico da Figura 6.

Figura 6 - Gráfico da semelhança estrutural (B)



Fonte: O autor (2025).

Na tese de Roe (2015), o PSNR foi uma das métricas principais de avaliação de qualidade das imagens binarizadas em relação às versões de referência (*ground truth*). O método proposto pelo autor apresentou um PSNR médio de 65,90 dB, superando os algoritmos clássicos de Otsu (62,56 dB) e Niblack (53,99 dB), bem como outros métodos recentes, como Howe (65,97 dB) e Mesquita et al. (66,01 dB). Em comparação, os resultados obtidos no presente trabalho mostram um PSNR máximo de 63,06 dB (binarização pela média global), valor inferior aos métodos de Roe, mas ainda dentro da faixa considerada de alta fidelidade visual, especialmente levando em conta que as imagens DIBCO utilizadas possuem degradações mais intensas e não passaram por restauração colorimétrica. Os valores de PSNR obtidos reforçam a preservação de contraste e integridade estrutural, ainda que com menor refinamento que os métodos baseados em equalização local e filtragem XDoG descritos por Roe.

Pimenta (2009) demonstrou que a metodologia proposta para a reconstrução de documentos degradados, que se baseia em características geométricas do contorno e Programação Dinâmica, possui um desempe-

nho promissor, alcançando uma precisão de 81% na combinação de fragmentos, comparável ao método de referência (Solana, 2005). No entanto, o método proposto obteve um ganho considerável em termos de robustez, conseguindo uma redução de pelo menos 4% nos erros e uma diminuição significativa no número de falsos candidatos a parceiros. Os resultados também mostraram a capacidade de reconstruir documentos completos, embora em casos de alta fragmentação, alguns pedaços não tenham sido adicionados por falta de parceiros compatíveis, o que levou a uma análise da estatística de erro.

Nogueira e Amorim (2019) discutem o papel da fidelidade estrutural e perceptiva na reconstrução digital de documentos históricos, o que se relaciona diretamente com a métrica SSIM. No presente trabalho, o SSIM máximo alcançado foi 0,880, indicando uma correspondência estrutural entre a binarização e o *ground truth*. Comparativamente, as reconstruções digitais mencionadas por Nogueira e Amorim destacam a importância de preservar a textura, contraste e forma dos objetos originais em modelos digitais, objetivos análogos aos da binarização documental. Assim, o SSIM obtido demonstra que os métodos aplicados mantêm a integridade estrutural das imagens de texto, o que está em consonância com a abordagem de preservação digital proposta por esses autores.

Em síntese, os valores obtidos neste estudo demonstram consistência estrutural, especialmente considerando os métodos utilizados e a natureza degradada das imagens DIBCO. Isso sugere que, mesmo com abordagens menos complexas, é possível alcançar resultados competitivos, preservando adequadamente a estrutura textual essencial para tarefas de reconhecimento óptico e arquivamento digital.

5 CONCLUSÃO

A presente pesquisa teve como propósito comparar diferentes técnicas de binarização aplicadas à restauração e preservação de documentos históricos digitalizados, avaliando sua eficiência na recuperação da legibilidade e na manutenção da estrutura textual. A partir dos experimentos realizados com a base DIBCO 2017, observou-se que os métodos globais de binarização, especialmente o Nível Médio de Cinza, Dois Picos e Otsu, apresentaram desempenho superior nas métricas quantitativas (PSNR) e qualitativas (SSIM), destacando-se pela capacidade de preservar contraste, contornos e elementos estruturais mesmo em imagens com degradação acentuada.

Em contrapartida, métodos adaptativos como Niblack, Porcentagem de Preto e White e Rohrer mostraram-se menos eficazes para o conjunto de dados analisado, apresentando resultados inconsistentes em cenários de baixo contraste ou com presença de ruídos extensos no papel. Tais limitações reforçam que a adequação de um algoritmo depende diretamente das características do documento original e do tipo de degradação presente. Apesar de sua simplicidade, alguns métodos globais demonstraram elevada eficiência, confirmando que abordagens menos sofisticadas podem ser competitivas quando bem contextualizadas.

É importante salientar que a análise realizada neste estudo foi restrita a um conjunto específico de documentos e a uma seleção limitada de algoritmos. Dessa forma, os resultados, embora relevantes, não esgotam as possibilidades de investigação. Diferentes tipos de documentos, graus de deterioração, algoritmos de pré-processamento e combinações híbridas podem alterar significativamente o desempenho alcançado. Assim, recomenda-se cautela na generalização dos resultados e reconhece-se a necessidade de ampliar a escala e a diversidade das imagens estudadas.

Em síntese, os resultados obtidos demonstram que a restauração digital de documentos deteriorados pode ser realizada de maneira eficiente mesmo com algoritmos relativamente simples, desde que aplicados de forma adequada ao contexto. O estudo reafirma o papel do processamento de imagens como ferramenta essencial para a preservação da memória documental e abre caminho para investigações mais amplas e aprofundadas no campo da visão computacional aplicada ao patrimônio histórico.

Por fim, baseado nos resultados obtidos e nos conhecimentos adquiridos durante a pesquisa, propõe-se como trabalhos futuros: uma análise mais ampla com outros tipos de documentos e métodos; realização de uma abordagem de pré-processamento e filtragem diferente; implementação de uma API que acolha os métodos e processos aqui estudados.

REFERÊNCIAS

BAGGIO, C. **Documentos Digitais: Preservação e Estratégias**. Rio Grande, Brasil: Revista do Instituto de Ciências Humanas e da Informação, 2013. v. 27. 11-24 p.

BUADES, A.; COLL, B.; MOREL, J. **Meios não locais de redução de ruído, processamento de imagem on-line**. Paris, França: IPOL - Image Processing on Line, 2011. 208 p.

CONWAY, P. **Preservação no universo digital**. Rio de Janeiro, Brasil: Projeto Conservação preventiva em bibliotecas e arquivos, 2001.

GONZALES, R.; WOODS, R. **Processamento de Imagens Digitais**. São Paulo, Brasil: Edgard Blücher, 2000.

JAIN, P.; TYAGI, V. **Uma pesquisa sobre métodos de redução de ruído de imagem com preservação de bordas**. Nova Iorque, EUA: Information System Frontiers, 2014.

MARENGONI, M. **Introdução à Visão Computacional usando OpenCV**. 1. ed. Porto Alegre, Brasil: Revista de Informática Teórica e Aplicada, 2009. XVI.

MELLO, C. et al. **Avanços Recentes no Processamento Digital de Imagens de Documentos Antigos**. Recife, Brasil: Universidade Federal de Pernambuco, 2013.

MILANO, D. de; HONORATO, L. B. **Visão computacional**. Faculdade de Tecnologia - UNICAMP, Limeira, São Paulo, Brasil. 2010.

NIBLACK, W. **An Introduction to Digital Image Processing**. Englewood Cliffs, NJ, USA: Prentice Hall, 1986.

NOGUEIRA, F.; AMORIM, A. **Reconstrução digital: A preservação em meio digital da herança cultural**. Universidade Federal da Bahia - Laboratório de Estudos Avançados em Ciências, Arquitetura e Tecnologias Digitais (LCAD). 2019.

OTSU, N. **Um método de seleção de limiar a partir de histogramas de nível de cinza**. 1. ed. [S.l.]: IEEE Transactions on Systems Man and Cybernetics, 1979. v. 20. 62-66 p.

PARKER, J. R. **Algorithms for Image Processing and Computer Vision**. [S.l.]: John Wiley and Sons, 1997.

PIMENTA, A. M. **Reconstrução Digital de Documentos Mutilados Usando Programação Dinâmica**. Curitiba, Brasil: [s.n.], 2009.

ROE, E. Tese de Doutorado. **Restauração Digital de Imagens Coloridas de Documentos Históricos**. Recife, Brasil: [s.n.], 2015.

SAUVOLA, J.; PIETIKAINEN, M. **Adaptive document image binarization**. [S.l.: s.n.], 2000. v. 33. 225–236 p.

SEZGIN, M.; SANKUR, B. **Survey over image thresholding techniques and quantitative performance evaluation**. Journal of Electronic Imaging, SPIE, v. 13, n. 1, p. 146–165. 2004.

SRINIVASAN, R. et al. Performance of binarization algorithms on tamizhi inscriptions. In: **Proceedings of the International Conference on Computational Intelligence**. [S.l.]: ACM, 2024.

VAZ, F.; ARAÚJO, C.; VARJÃO, F. **Restauração De Documentos Eletrônicos Degradados Com Binarização Adaptativa.** [S.l.]: Rios Eletrônica, 2011.

WHITE, J. M.; ROHRER, G. D. **Image thresholding for optical character recognition and other applications requiring character image extraction.** [S.l.: s.n.], 1983. v. 27. 400–411 p.

A Apêndice A – Resultados Visuais

Nesta seção são apresentadas algumas das imagens resultantes dos métodos aplicados.

Figura 7 - Imagem processada pelo método de Otsu

opdragen, om het volken dajoon bijz. te kennen. —
waaronder in gewijst is, zind den geïncorporeerde
in Raden van India tot nu toe bij de de beo-
zindt niet beschreven van gaten, pamaetkan, gaffan
in Cessing geijken stel, het wt songte geijfend
daar verwoordt, in het volke. Het in andere men-
tege et vade willeen. Zoude houfsten —
gechreven in Cartel Batavia, op 6
Eylend groot Javaans Comdijck Jacaraden
101. maet 17702: (overstent) den geïncorporeerde
gemaakt van Nederlanden India, en gewaant, Willem
van Rithoor, der Ziden hand te my. Zegge in
raden acht donker, en daer men geet, den overnauten
van long gewaant van de ^{te recht} (overstent) Willem de
Roo Gewaant. —

Fonte: O autor (2025).

Figura 8 - Imagem processada pelo método de Dois Picos

16

Adan. Was soll ich antworten? Ich verstehe dein reden nicht.

Liber. Ich will sagen: dem blind-geflügleten Kinde.

Adan. Ich verstehe auch dieses nicht.

Liber. Der Liebes Gottheit.

Adan. Diese ist mir unbekant.

Liber. O alberes Herge / kennst du nicht den Liebes-Scherge?

Fonte: O autor (2025).

Figura 9 - Imagem processada pelo método do nível médio de cinza

Inhalt.	
Vom Gesichte.	Seite 342
Vom Geruche.	344
Erklärungen seltsamer Gelüste; Anwendungen	349
wunderlicher Begierden der Schwangeren und	
hysterischen.	350
Von den Leidenschaften, und der Nothwendigkeit	
des Studiums der Menschenkenntniß etc.	364
Von psychologischen Geheimnissen oder den Wissen-	
schaften der Sybilen.	368
Von sonderheitlichen Gefühlen und Empfindungen.	375
Theorie angenehmer Empfindungen.	378

Fonte: O autor (2025).

Figura 10 - Imagem processada pelo método de Savoula

(28)

Epilogue.

Roger left behind:

IF any thing obscure, not understood
 Be here, the State *expounds* the dark st word:
 And makes the thickest thickets plain and clear,
 As the back of your hand, as *Shot-over:*
 For *Nob* expounds, the *officers* expound,
 The *Souldiers* too expound, *All in a Round:*
 You'd think *Jerusalem* were building new,
 With *Swords in hand* and *Bible*, nay by th' *Jew:*
 The *Jews* buy all, the *Independents* sell,
 The *mares* the *King's* and's *friends*; the *Fair* is hell:
 Dfference there is twixt *that*, and *Bartholomew*,
 That brings *Brimstone* and *Fire*, this the cold *Dem.*

F I N I S.

Fonte: O autor (2025).

Figura 11 - Imagem processada pelo método de Dois picos

Honoured Sir -
I have the honor to acknowledge the receipt of your letter of the 30th inst. and in reply to inform you that the same has been forwarded to the proper authorities for their consideration. I am, Sir, very respectfully,
Yours obedient servant,
James M. D.
John M. D.
Charles M. D.

Fonte: O autor (2025).

Figura 12 - Imagem processada pelo método de Dois Picos

London Jan 6th 1755
I have the honor to acknowledge the receipt of your letter of the 30th inst. and in reply to inform you that the same has been forwarded to the proper authorities for their consideration. I am, Sir, very respectfully,
Yours obedient servant,
James M. D.
John M. D.
Charles M. D.

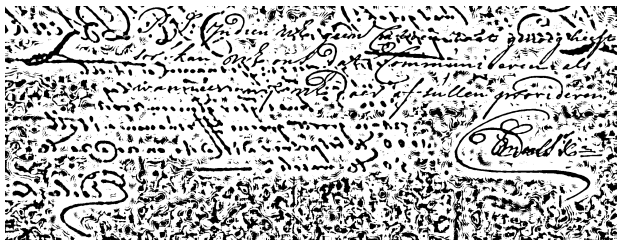
Fonte: O autor (2025).

Figura 13 - Imagem processada pelo método de Otsu

London Jan 6th 1755
I have the honor to acknowledge the receipt of your letter of the 30th inst. and in reply to inform you that the same has been forwarded to the proper authorities for their consideration. I am, Sir, very respectfully,
Yours obedient servant,
James M. D.
John M. D.
Charles M. D.

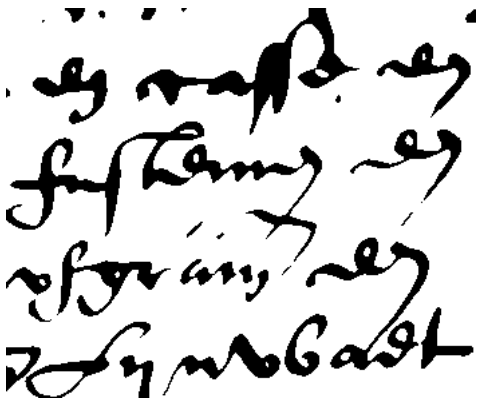
Fonte: O autor (2025).

Figura 14 - Imagem processada pelo método de Niblack



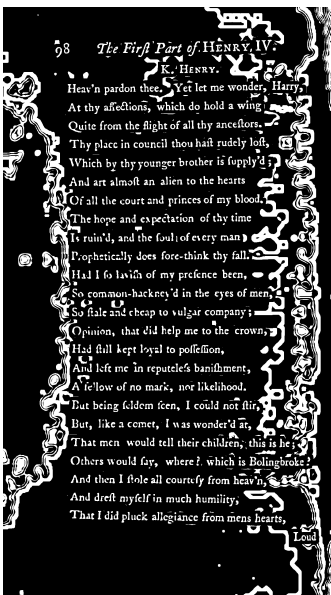
Fonte: O autor (2025).

Figura 15 - Imagem processada pelo método de Dois Picos



Fonte: O autor (2025).

Figura 16 - Imagem processada pelo método de White e Rohrer



Fonte: O autor (2025).