

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Um Novo Algoritmo para a Remoção de Bordas Ruidosas em Documentos Monocromáticos

Elaborado por:

Ednardo Jerônimo Cavalcante Mariano

Orientador: Prof. Dr. Rafael Dueire Lins

Recife, Janeiro de 2012.

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**Um Novo Algoritmo para a Remoção de Bordas
Ruidosas em Documentos Monocromáticos**

por

Ednardo Jerônimo Cavalcante Mariano

Dissertação submetida ao Programa de Pós-Graduação em Engenharia Elétrica da
Universidade Federal de Pernambuco como parte dos requisitos para a obtenção do grau de
Mestre em Engenharia Elétrica.

ORIENTADOR: Prof. Dr. Rafael Dueire Lins

Recife, janeiro de 2012.

© Ednardo Jerônimo Cavalcante Mariano, 2012

Catálogo na fonte
Bibliotecária Margareth Malta, CRB-4 / 1198

M333u Mariano, Ednardo Jerônimo Cavalcante.
Um novo algoritmo para a remoção de bordas ruidosas em documentos monocromáticos / Ednardo Jerônimo Cavalcante Mariano. - Recife: O Autor, 2012.
126 folhas, il., gráfs., tabs.

Orientador: Prof. Dr. Rafael Dueire Lins.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-Graduação em Engenharia Elétrica, 2012.
Inclui Referências Bibliográficas e Apêndices.

1. Engenharia Elétrica. 2. Documentos digitalizados monocromáticos. 3. Remoção de ruído de borda. 4. Processamento de documentos. 5. Processamento de imagens. 6. Imagens monocromáticas. I. Lins, Rafael Dueire. (Orientador). II. Título.

UFPE

621.3 CDD (22. ed.)

BCTG/2012-096



Universidade Federal de Pernambuco

Pós-Graduação em Engenharia Elétrica

PARECER DA COMISSÃO EXAMINADORA DE DEFESA DE
DISSERTAÇÃO DO MESTRADO ACADÊMICO DE

EDNARDO JERÔNIMO CAVALCANTE MARIANO

TÍTULO

**“UM NOVO ALGORITMO PARA A REMOÇÃO DE BORDAS RUIDOSAS
EM DOCUMENTOS MONOCROMÁTICOS”**

A comissão examinadora composta pelos professores: RAFAEL DUEIRE LINS, CIN/UFPE, VALDEMAR CARDOSO DA ROCHA JÚNIOR, DES/UFPE e MARCELO RESENDE THIELO, HP/BRASIL sob a presidência do primeiro, consideram o candidato **EDNARDO JERÔNIMO CAVALCANTE MARIANO**
APROVADO.

Recife, 31 de janeiro de 2012.

CECÍLIO JOSÉ LINS PIMENTEL
Vice-Coordenador do PPGEE

RAFAEL DUEIRE LINS
Orientador e Membro Titular Interno

MARCELO RESENDE THIELO
Membro Titular Externo

VALDEMAR CARDOSO DA ROCHA JÚNIOR
Membro Titular Interno

Resumo da Dissertação apresentada à UFPE como parte dos requisitos necessários para a obtenção do grau de Mestre em Engenharia Elétrica.

UM NOVO ALGORITMO PARA A REMOÇÃO DE BORDAS RUIDOSAS EM DOCUMENTOS MONOCROMÁTICOS

Ednardo Jerônimo Cavalcante Mariano

Janeiro/2012

Orientador: Prof. Dr. Rafael Dueire Lins.

Área de Concentração: Telecomunicações.

Palavras-chave: documentos digitalizados monocromáticos, remoção de ruído de borda, processamento de documentos, processamento de imagens, imagens monocromáticas.

Número de Páginas: 126.

RESUMO: A digitalização de grandes quantidades de documentos através de *scanners* de linha de produção geralmente incorpora na imagem defeitos que acarretam dificuldades desde a sua leitura até sua transcrição automática. Tais defeitos devem ser removidos, obtendo imagens de documentos com melhor qualidade, através de técnicas de processamento de imagens digitais. Devido à grande quantidade de documentos, essas técnicas devem ser rápidas e capazes de abranger a maior parte das falhas nessas imagens. Esta dissertação apresenta um novo algoritmo para remoção de bordas ruidosas cujo tempo de processamento chegou a ser reduzido em até 62% em relação ao algoritmo de Ávila-Lins, com qualidade semelhante, quando testado num conjunto de cinco mil imagens.

Abstract of Dissertation presented to UFPE as a partial fulfillment of the requirements for
the degree of Master in Electrical Engineering

A NEW AND EFFICIENT ALGORITHM FOR NOISY BORDER REMOVAL IN MONOCHROMATIC DOCUMENTS

Ednardo Jerônimo Cavalcante Mariano

January/2012

Supervisor: Prof. Dr. Rafael Dueire Lins, PhD.

Area of Concentration: Telecommunications.

Keywords: monochrome digitized documents, noise removal border, document processing, image processing, monochrome images.

Number of pages: 126.

ABSTRACT: The scanning large quantities of documents by production line scanners usually incorporates image defects that cause difficulties since his reading until his automatic transcription. Such defects should be removed, obtaining document images with better quality, through techniques of digital image processing. Due to the large amount of documents, these techniques must be quick and able to cover most of the flaws in these pictures. This paper presents a novel algorithm for removing noisy edges whose processing time has come to be reduced by 62% compared to Avila Lins algorithm, similar in quality when tested in a group of five thousand images

SUMÁRIO

1 – Introdução	1
1.1 - Digitalização de Documentos.....	3
1.2 – Objetivos, limitações e critérios	9
1.3 – Visão Geral desta dissertação	10
2 – Processamento Digital de Imagens de documentos monocromáticos	11
2.1 – Remoção de bordas ruidosas	12
2.2 – Estado da Arte	17
2.2.1 – O algoritmo Ávila-Lins para remoção do ruído de borda ruidosa.....	17
2.2.2 – O algoritmo de pré-processamento para o Ávila-Lins	20
3 – Uma nova proposta de pré-processamento	23
3.1 - Implementação.....	33
4 - Resultados.....	34
4.1 – Comparação dos resultados	34
4.2 – Análise qualitativa das imagens filtradas.....	87
4.2.1 – Comparação com ferramentas comerciais	118
5 – Conclusões e Trabalhos Futuros.....	120
5.1 – Trabalhos Futuros.....	121
6 – Referências.....	122
APÊNDICE A - Código fonte dos algoritmos e programas.....	125
APÊNDICE B - Publicações	126

1 – Introdução

A linguagem escrita é uma tecnologia de comunicação, que marca o início da história na sociedade humana, e consiste em registrar marcas em um meio para exprimir ideias e guardar informação. A escrita é uma das maiores invenções da humanidade, ampliando a capacidade de transmissão do conhecimento e garantindo a integridade da informação passada ao longo das gerações, ao contrário da linguagem oral, que sofre distorções à medida que é passada de geração em geração, ou mesmo entre pessoas da mesma geração.

Enquanto a linguagem escrita evolui para melhor expressão e compreensão a evolução do meio onde são feitos os registros escritos facilita o trabalho de registro, o transporte, a divulgação e a durabilidade da informação contida neles. Diversos meios foram utilizados ao longo da história para o registro da escrita: pedra, cerâmica, madeira, fibras vegetais e finalmente o papel como conhecemos.

O papel é utilizado pelo homem há vários séculos e tornou-se o principal meio de armazenamento e publicação de conhecimento e informações dispostas em forma de ideias, relatos, estórias, mapas, desenhos e obras de artes que são transmitidos de geração em geração através de livros, jornais, revistas, artigos, teses, documentos, etc. Apesar da importância do papel para o homem, ele apresenta uma série de desvantagens na sua utilização. O papel se desgasta com o tempo tornando-se frágil. Caso não seja conservado em um ambiente adequado, pode sofrer a ação de fungos e insetos. Além disso, ele pode ser rasgado, amassado, dobrado, molhado, perdido, queimado, falsificado, etc. Outra desvantagem é o custo relacionado ao papel. Mesmo que o preço unitário do papel seja extremamente pequeno, o gasto com materiais relacionados à sua utilização e manuseio, como por exemplo, cartucho de impressoras, lápis, pastas, cliques e pranchetas, cópias, etc., agregam um alto valor ao seu custo. A enorme quantidade de papéis, hoje existente nos mais diversos ambientes, ocupa um grande espaço. Em empresas públicas e privadas, tornou-se comum o uso de estantes e salas de arquivos para armazenar e organizar os documentos mais importantes, como por exemplo, contratos, notas fiscais, relatórios, manuais, etc. Essa necessidade também agrega um valor ainda maior ao custo do papel. As empresas maiores, com grandes quantidades de papéis, terceirizam o serviço de armazenamento e organização dos documentos que são guardados em galpões.

Segundo dados da Environmental Paper Network [30][31] houve uma queda no consumo de papel em algumas regiões do mundo, porém houve um maior aumento no consumo de papel em outras regiões.

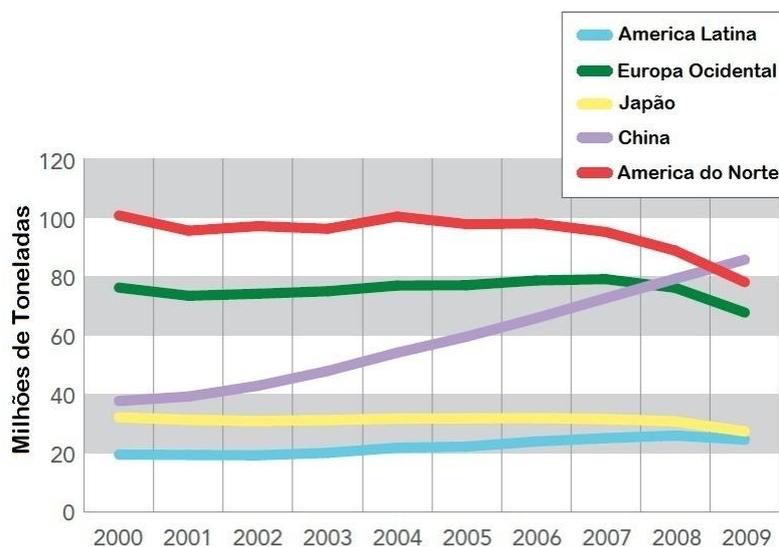


Gráfico 1 - Evolução do consumo de papel no mundo [30][31]

No Brasil, segundo dados da Bracelpa [24], a produção e o consumo de papel aumentaram nos últimos anos (Tabela 1).

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
PAPEL - TOTAL										
PRODUÇÃO	7.200	7.438	7.774	7.916	8.452	8.597	8.725	9.008	9.409	9.428
IMPORTAÇÃO	839	632	560	578	734	770	967	1.097	1.328	1.085
EXPORTAÇÃO	1.225	1.368	1.455	1.778	1.853	2.039	1.990	2.006	1.982	2.008
CONSUMO APARENTE	6.814	6.702	6.879	6.716	7.333	7.328	7.702	8.099	8.755	8.505
CONSUMO PER CAPITA (kg/hab.)	40,2	38,9	39,5	37,7	40,0	39,5	41,2	44,0	46,2	44,2

Tabela 1 - Evolução do consumo e produção de papel no Brasil [24]

A aparente estagnação tanto no consumo quanto na produção é o reflexo do aumento na taxa de recuperação de papéis recicláveis, que no Brasil alcança 46% [24]. Em outros países esta taxa supera os 60% (Tabela 2).

Países Selecionados	Taxa de recuperação*
Coreia do Sul	91,6%
Alemanha	84,8%
Japão	79,3%
Reino Unido	78,7%
Espanha	73,8%
Estados Unidos	63,6%
Itália	62,8%
Indonésia	53,4%
Finlândia	48,9%
México	48,8%
Brasil	46,0%
Argentina	45,8%
China	40,0%
Rússia	36,4%
Índia	25,9%

Tabela 2 - Taxa de recuperação de papéis recicláveis até 2010 [24][31]

*Volume de aparas recuperadas no país dividido pelo consumo aparente de papel.

Com o advento dos computadores, havia-se suposto que o papel se tornaria obsoleto. Então, surgiu o conceito de *Paperless Office*, segundo o qual as empresas não utilizariam papel e tudo seria feito no computador. Contudo, este conceito não se concretizou e a tendência atual é que o uso do papel seja feito de forma mais racional.

Atualmente, a facilidade e a velocidade para transmissão documentos digitais, grande capacidade e segurança de armazenamento em dispositivos cada vez menores e a redução de preço dos equipamentos para armazenamento e digitalização de documentos, tornam o armazenamento em meio digital a melhor e mais segura forma de armazenamento.

1.1 - Digitalização de Documentos

Em geral, os documentos físicos são dados desestruturados [02], ou seja, não possuem dados que os caracterize ou identifique. A digitalização de forma estrita é um método que toma como entrada documentos físicos e gera imagens de cada um deles. Para identificar cada imagem e associá-la ao documento físico, um conjunto mínimo de dados é criado, tornando, desta forma, o documento físico em um dado semiestruturado. Em outras palavras, para que a digitalização seja efetiva, esta tem que transcender a fase de simples conversão em imagem do documento original (captura) e deve compreender o processo de transformação dos dados desestruturados em dados semiestruturados. As fases do processo de digitalização (Figura 1) são basicamente:

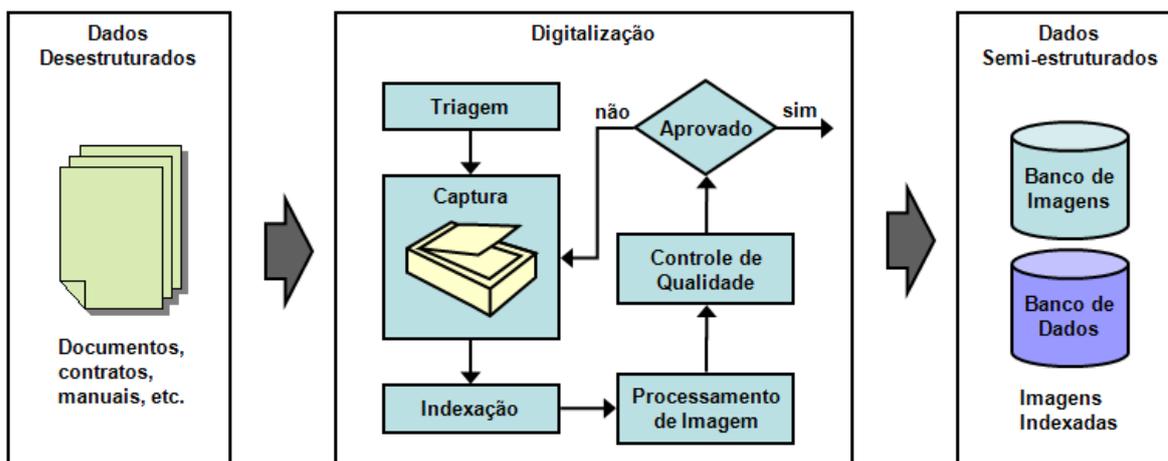


Figura 1 - Processo básico de digitalização: transformação de dados desestruturados em semiestruturados [01]

Triagem: processo manual de separação e preparação dos documentos físicos para a captura. A triagem é uma fase caracterizada por ser, em sua maioria, manual e é responsável em preparar e organizar os documentos físicos para a fase de captura. Em projetos de digitalização de alto volume de documentos, esta fase é assistida pelo computador para a geração de códigos de barra com dados referentes à sua identificação. Os códigos de barra impressos serão incluídos entre os papéis de forma a organizá-los e serão, posteriormente, detectados e seus dados reconhecidos e utilizados na fase de indexação.

Captura: o documento físico (papel) é capturado através de um dispositivo eletrônico, como por exemplo, *scanners* e câmeras digitais. Uma vez que há diversos equipamentos que podem ser utilizados para converter o documento físico em uma imagem digital (digitalizadores manuais, semimanuais ou automáticos, câmeras digitais, canetas digitalizadoras), a escolha do equipamento depende do tipo de documento (livros, documentos históricos, papéis individuais), qualidade da imagem desejada (resolução, imagem colorida, em escala de cinza ou preto e branco) e da quantidade de documentos a serem digitalizados. Para trabalhos de digitalização de grandes quantidades de documentos sem valor iconográfico, tais como documentos burocráticos, são utilizados digitalizadores automáticos, em geral, configurados para aquisição de imagens em escala de cinza ou preto e branco com resolução de 200 dpi [01][18].

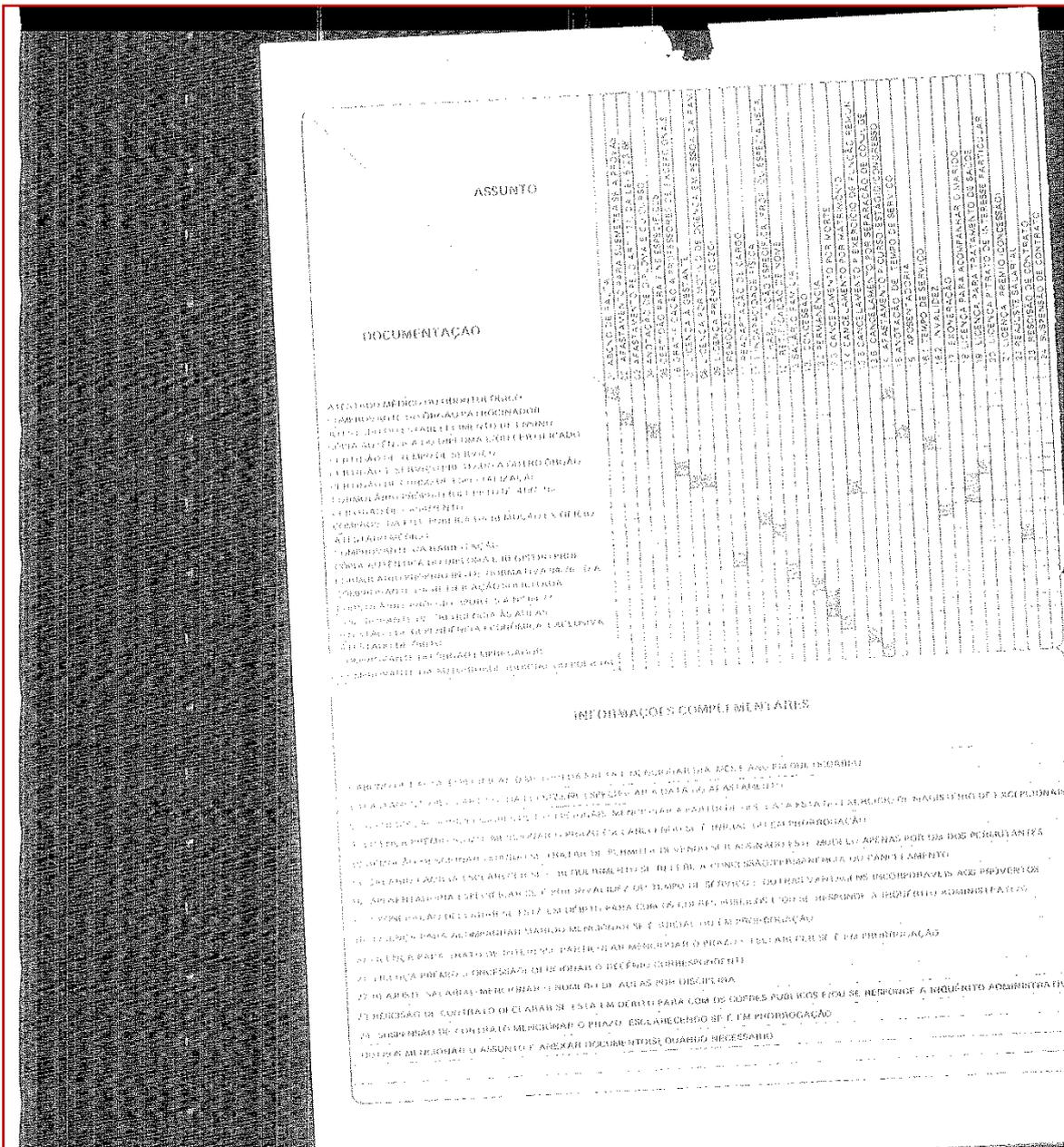
Indexação: criação de um conjunto mínimo de dados de forma que possibilite a identificação da imagem gerada e a sua associação com o documento físico. A indexação é iniciada na fase de triagem, onde os documentos a serem digitalizados são marcados para posterior identificação da imagem obtida.

Para trabalhos com grandes quantidades de documentos são utilizados códigos de barra para identificação e organização dos grupos de imagens gerados. Após o processamento da imagem é possível melhorar a indexação a partir de palavras obtidas do conteúdo dos documentos, através de técnicas de transcrição automática do documento (OCR - *Optical Character Recognition*).

Processamento da imagem: frequentemente o documento é capturado com falhas como, por exemplo, bordas pretas, ruídos, rotação ou distorção (Figuras 2, 3 e 4). Estas falhas podem ocorrer por diversos fatores como: má conservação do documento físico, apresentando perfurações, ranhuras e/ou fissuras; poeira na mesa de digitalização; diferenças entre o tamanho do papel digitalizado e o tamanho da mesa de digitalização, em geral a área do papel é menor que a área da mesa de digitalização, ocasionando o aparecimento de bordas pretas, rotação do documento e/ou distorções. Vários filtros de tratamento da imagem capturada são aplicados com o objetivo de melhorar a sua qualidade visual, diminuir o tamanho de armazenamento e opcionalmente executar a transcrição automática do documento. Em um projeto de digitalização de uma grande quantidade de documentos é recomendado que toda a fase de processamento de imagem deva ser automatizada.

Controle de qualidade: processo manual assistido pelo computador para avaliar cada uma das imagens em relação à qualidade visual e aos resultados da indexação. Caso a imagem seja rejeitada, volta para a fase de captura. A fase de controle de qualidade caracteriza-se por ser a mais lenta do processo, em consequência da necessidade de inspecionar cuidadosamente os resultados das indexações e de processamentos de todas as imagens manualmente, auxiliadas pelo computador. As tarefas desta fase incluem: verificar e corrigir a ordem das páginas e os resultados do OCR; avaliar a qualidade visual da imagem final e aplicar filtros específicos para tratamento da imagem, caso detectado algum erro. As imagens com falhas que não puderem ser corrigidas manualmente serão rejeitadas e reenviadas à fase de captura.

A avaliação da qualidade do documento digitalizado, assim como, uma avaliação automática, ainda são problemas em aberto na área de engenharia de documentos e que devem ser motivo de estudos futuros. Caso o projeto de digitalização tenha que transcrever o conteúdo do documento digitalizado, a medição da taxa de acerto do OCR pode ser utilizada apenas para aprovar ou rejeitar parte dos documentos, podendo ser adequado como uma medida quantitativa da qualidade do documento digitalizado.



**Figura 2 – Exemplo de documentos digitalizados com falhas:
Rotação e forte ruído originado por falha do digitalizador.**

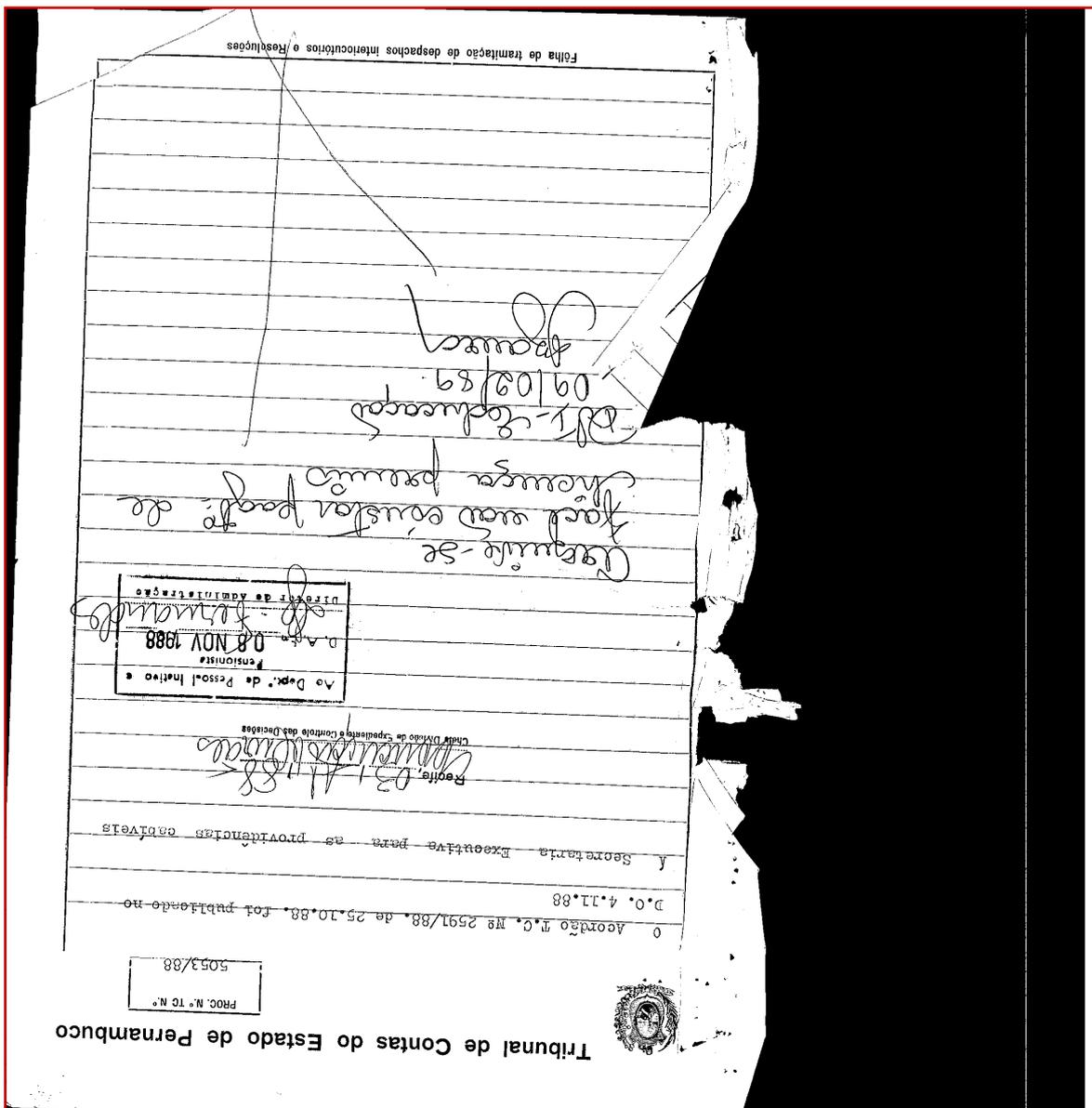


Figura 3 - Exemplo de documento digitalizado com falhas:

Erro de orientação (180°) e rotação, borda “preta” irregular e rasgada na margem esquerda do documento.

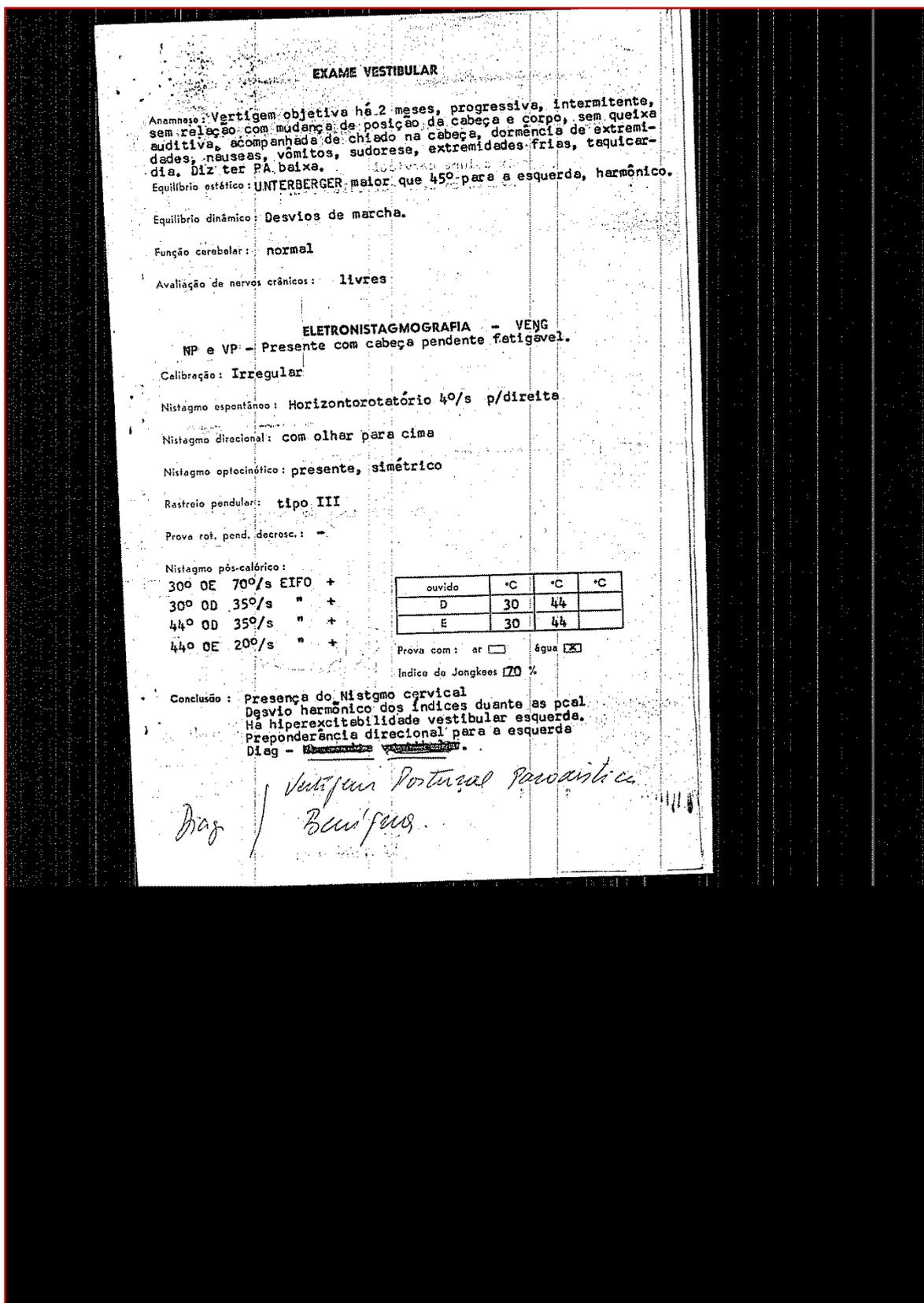


Figura 4 - Exemplo de documento digitalizado com falhas:
 Rotação e borda "preta" por falha no ajuste do tamanho do papel a ser digitalizado.

1.2 – Objetivos, limitações e critérios

O armazenamento em meio digital de documentos demanda um grande trabalho e organização para completa aquisição das informações contidas nos documentos com o mínimo de perdas ou falhas. As etapas de captura, indexação e processamento da imagem são, em sua maioria, efetuados de forma completamente automatizada, para reduzir o esforço humano e acelerar o trabalho. Na fase de captura, *scanners* de alta produção [02][21][22] conseguem digitalizar centenas de páginas por hora. Técnicas de indexação automatizadas que utilizam OCR permitem classificação e a busca de documentos através até mesmo do seu conteúdo. Porém o foco desta dissertação é a etapa de processamento da imagem, onde a necessidade de automação e melhoria no desempenho e confiabilidade deste processo é fundamental para o bom desempenho das etapas subsequentes e evitar a perda de tempo em muitos casos de falhas do processo de captura, evitando que o documento capturado com falhas seja novamente submetido ao processo de captura.

O algoritmo Ávila-Lins [02], estado da arte na remoção do ruído de borda [04] em documentos preto e branco, combina alto desempenho, com baixo tempo de processamento por imagem, e confiabilidade, remoção do ruído de borda sem perda ou comprometimento da informação contida no documento. No entanto, por ser um algoritmo do tipo “preenchimento” (*flood fill*) [19] o tempo de processamento por imagem é diretamente proporcional à área da imagem processada [01][02], além de possuir limitações quanto à remoção do ruído de borda quando este é adicionado com o ruído tipo “*sal e pimenta severo*” [02][16][17].

Em [01] há a descrição de um algoritmo de pré-processamento que remove significativamente a área do ruído antes da aplicação do algoritmo Ávila-Lins [02] [03] e [04], melhorando consideravelmente seu desempenho em relação ao tempo de processamento por imagem. O algoritmo descrito em [01], será aqui referenciado como algoritmo de Formiga-Lins e utiliza como base o próprio algoritmo Ávila-Lins [02].

Esta dissertação apresenta nova proposta de algoritmo de pré-processamento para o algoritmo Ávila-Lins que melhora seu desempenho em relação ao tempo de processamento por imagem e reduza suas restrições quanto ao ruído do tipo “*sal e pimenta severo*”, aumentando o alcance e confiabilidade do Algoritmo Ávila-Lins.

1.3 – Visão Geral desta dissertação

Esta dissertação está organizada em quatro capítulos, além desta introdução e das referências bibliográficas. O capítulo 2 descreve os principais algoritmos em processamento de imagens da literatura que foram utilizados neste estudo. O capítulo 3 aborda o problema da remoção de bordas ruidosas com a descrição de uma nova proposta de algoritmo para remoção dessas. No capítulo 4 são apresentados e comparados os resultados dos algoritmos descritos nesta dissertação para remoção de bordas ruidosas. Finalmente, o capítulo 5 conclui este estudo com considerações finais e propostas de trabalhos futuros.

Além do tratamento específico do problema de remoção de borda, foco desta dissertação, encontram-se no apêndice B um trabalho em área correlata desenvolvidos pelo autor durante o mestrado. Em anexo a esta dissertação está um DVD com todas as imagens de teste, todas as imagens processadas e o código fonte dos algoritmos propostos, Formiga-Lins e Ávila-Lins e o Novo Algoritmo. O DVD está organizado segundo a descrição do apêndice A desta dissertação.

2 – Processamento Digital de Imagens de documentos monocromáticos

Uma imagem pode ser definida como uma função bidimensional, $f(x,y)$, em que x e y são coordenadas *espaciais* (plano), e a amplitude de f em qualquer par de coordenadas (x,y) é chamada de *intensidade* ou *nível de cinza* da imagem nesse ponto [16]. Quando x , y e os valores de intensidade de f são quantidades finitas e discretas, é denominada de *imagem digital* [16]. O campo do *processamento digital de imagens* se refere ao processamento de imagens digitais por um computador digital [16].

A imagem digital é representada por uma matriz finita onde cada um de seus elementos são os valores de intensidade da imagem. Esses elementos são chamados de *elementos pictóricos* ou simplesmente *pixels* [16]. *Pixel* é o termo mais popular utilizado para representar os elementos de uma imagem digital [16].

É intuitivo que quanto maior for a quantidade de pixels que definam uma imagem, melhor será sua apresentação, com maiores detalhes e maior nitidez. Este é o conceito da *resolução espacial* [16], a medida do menor detalhe discernível em uma imagem. Quantitativamente, a resolução espacial pode ser expressa de várias formas, sendo a mais comum a unidade de pontos (pixels) por unidade de distância. A unidade de resolução espacial utilizada nesta dissertação será a unidade de “*pontos por polegada*” (*dots per inch, dpi*), por ser a mais utilizada nas especificações de *scanners* de produção. É interessante perceber que quanto maior for a resolução espacial, ou seja, o *dpi* utilizado, maior será a quantidade pixels na imagem, conseqüentemente maior o número de bytes utilizados para armazenar a imagem.

Como definido anteriormente, o valor de cada pixel é finito e discreto para que possa ser representado por bits. O número de bits que representa um pixel depende de quantos níveis de intensidade são desejados para representar a imagem. A *resolução de intensidade* [16] refere-se à menor variação discernível de nível de intensidade na imagem.

A menor representação possível é a de um único bit (0 ou 1) por pixel, havendo apenas um único nível de intensidade máxima e mínima. Em geral, “0” para menor intensidade, ou seja, o preto, e “1” para a maior intensidade, o branco.

A digitalização de documentos burocráticos e sem valor iconográfico, se efetuada com resolução de 200 dpi e um único bit por pixel, guarda todos os elementos essenciais do documento oferecendo um bom fator qualidade/espço de armazenamento [04] e transmissão via rede de computadores [05].

Em caso de transcrição automática via OCR, estudos [08] mostram que a resolução de 200 dpi oferece uma boa taxa de acertos, podendo sofrer degradação com o uso de resoluções mais altas. A imagem resultante é monocromática sem escalas de tons de cinza, apenas preto e branco.

Os algoritmos para processamento digital de imagens podem ser classificados [02] em algoritmos de pré-processamento e pós-processamento, de acordo com a ordem ao qual será utilizado. A proposta desta dissertação é um algoritmo de pré-processamento em relação ao algoritmo para remoção do ruído de borda preta [01] [02] [07] [08].

2.1 – Remoção de bordas ruidosas

Scanners de linha de produção são comumente utilizados para capturar documentos burocráticos de diversas dimensões em um projeto de digitalização. As imagens obtidas pela digitalização de documentos de dimensões menores que a mesa de digitalização do scanner, enviesados ou rasgados, apresentam uma borda preta circundando-a (Figuras 5, 6, 7 e 8). Devido a outros fatores como brilho, contraste, poeira na mesa digitalizadora, interferência do ambiente, etc., pode surgir na borda preta uma grande quantidade de padrões de ruído [02].

A borda ruidosa traz diversos problemas:

- Degrada a qualidade da visualização em tela do documento digitalizado produzindo poluição visual, pois parte significativa da área apresentada não possui informação;
- Aumenta significativamente o tamanho (em bytes) da imagem comprimida;
- Aumento da quantidade de tinta utilizado na impressão do documento.

A remoção desse elemento do documento traz benefícios, contudo, a sua remoção manual é inviável para grandes quantidades de documentos a serem processados. Dessa forma, torna-se importante o desenvolvimento de um filtro automático para remover bordas “pretas” ruidosas.

A borda ruidosa apresenta as seguintes características:

- Circunda total ou parcialmente o documento digitalizado;
- A textura pode ser sólida e/ou apresentar ruídos com faixas brancas verticais ou horizontais;
- Pode apresentar um formato irregular devido a uma parte do documento físico (papel) que pode estar rasgada.

Uma característica importante foi identificada e deve-se dar destaque: parte da informação do documento pode estar conectada à borda preta. Desta forma, o filtro de remoção da borda ruidosa deve, fundamentalmente, evitar que parte da informação do documento seja perdida, podendo tolerar a manutenção de borda vestigial.

DOCUMENTAÇÃO	ASSUNTO	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
ATESTADO MÉDICO OU ODONTOLÓGICO																				
COMPROVANTE DO ÓRGÃO PATROCINADOR																				
ATESTADO DO ESTABELECIMENTO DE ENSINO																				
CÓPIA AUTÊNTICA DO DIPLOMA E/OU CERTIFICADO																				
CERTIDÃO DE TEMPO DE SERVIÇO																				
CERTIDÃO T. SERVIÇO PRESTADO A OUTRO ÓRGÃO																				
CERTIDÃO DE CURSO DE ESPECIALIZAÇÃO																				
FORMULÁRIO PRÓPRIO (DECRETO Nº 419/78)																				
CERTIDÃO DE CASAMENTO																				
COMPROV. DA ENT. PÚBLICA DA REMOÇÃO EX OFÍCIO																				
ATESTADO MÉDICO																				
COMPROVANTE DA HABILITAÇÃO																				
CÓPIA AUTÊNTICA DO DIPLOMA E REGISTRO PROF.																				
FORMULÁRIO PRÓPRIO INSTR. NORMATIVA 04/76 - D A																				
COMPROVANTE DA RETIFICAÇÃO SOLICITADA																				
FORMULÁRIO PRÓPRIO (PORT. S A Nº 64/77)																				
COMPROVANTE DE FREQUÊNCIA ÀS AULAS																				
ATESTADO DE DEPENDÊNCIA ECONÔMICA EXCLUSIVA																				
ATESTADO DE ÓBITO																				
COMPROVANTE DO ÓRGÃO EMPREGADOR																				
COMPROVANTE DA AUTORIDADE JUDICIAL OU POLICIAL																				

INFORMAÇÕES COMPLEMENTARES	
1	ABONO DE FALTA - ESPECIFICAR O MOTIVO DA FALTA E MENCIONAR DIA, MÊS E ANO EM QUE OCORREU
3	AFASTAMENTO PELO ART. 177, DA LEI 8123/88: ESPECIFICAR A DATA DO AFASTAMENTO
6	GRATIFICAÇÃO A PROFESSORES DE EXCEPCIONAIS: MENCIONAR A PARTIR DE QUE DATA ESTÁ NO EXERCÍCIO DE MAGISTÉRIO DE EXCEPCIONAIS
9	LICENÇA PRÊMIO (GOZO): MENCIONAR O PRAZO, ESCLARECENDO SE É INICIAL OU EM PRORROGAÇÃO
10	REMOÇÃO: MENCIONAR QUANDO SE TRATAR DE PERMUTA DEVENDO SER ASSINADO ESTE MODELO APENAS POR UM DOS PERMUTANTES
13	SALÁRIO FAMÍLIA: ESCLARECER SE O REQUERIMENTO SE REFERE A CONCESSÃO/PERMANÊNCIA OU CANCELAMENTO
16	APOSENTADORIA - ESPECIFICAR SE É POR INVALIDEZ OU TEMPO DE SERVIÇO E OUTRAS VANTAGENS INCORPORÁVEIS AOS PROVENTOS
17	EXONERAÇÃO: DECLARAR SE ESTÁ EM DÉBITO PARA COM OS COFRES PÚBLICOS E/OU SE RESPONDE A INQUÉRITO ADMINISTRATIVO
18	LICENÇA PARA ACOMPANHAR MARIDO: MENCIONAR SE É INICIAL OU EM PRORROGAÇÃO
20	LICENÇA PARA TRATO DE INTERESSE PARTICULAR: MENCIONAR O PRAZO E ESCLARECER SE É EM PRORROGAÇÃO
21	LICENÇA PRÊMIO (CONCESSÃO): MENCIONAR O DECÊNIO CORRESPONDENTE
22	REAJUSTE SALARIAL - MENCIONAR O NÚMERO DE AULAS POR DISCIPLINA
23	RESCISÃO DE CONTRATO: DECLARAR SE ESTÁ EM DÉBITO PARA COM OS COFRES PÚBLICOS E/OU SE RESPONDE A INQUÉRITO ADMINISTRATIVO
24	SUSPENSÃO DE CONTRATO: MENCIONAR O PRAZO, ESCLARECENDO SE É EM PRORROGAÇÃO
OUTROS: MENCIONAR O ASSUNTO E ANEXAR DOCUMENTO(S), QUANDO NECESSÁRIO	

Figura 5 - Exemplo de documento digitalizado por um scanner de produção. Borda do documento irregular a direita devido a danos no documento físico.

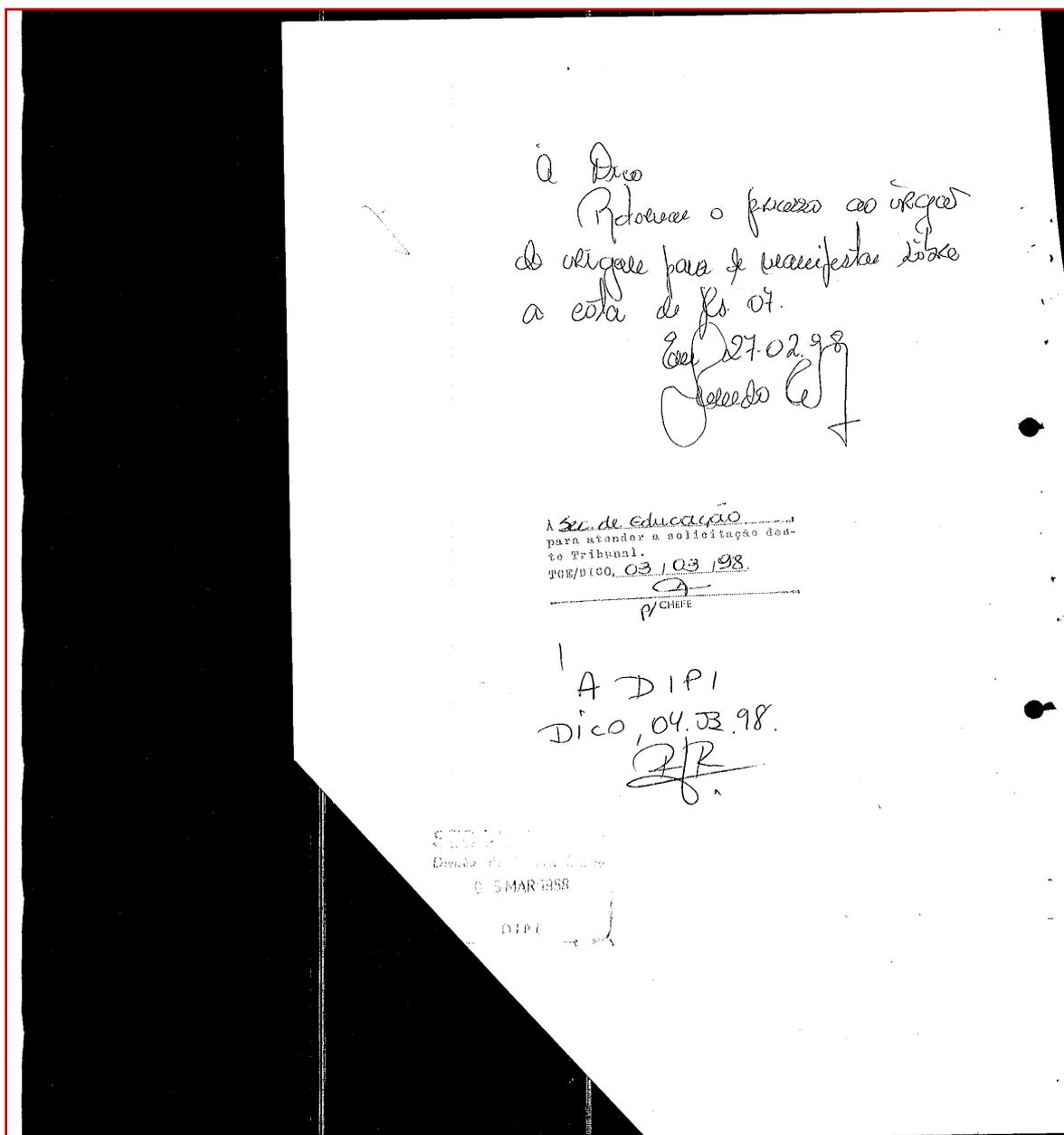
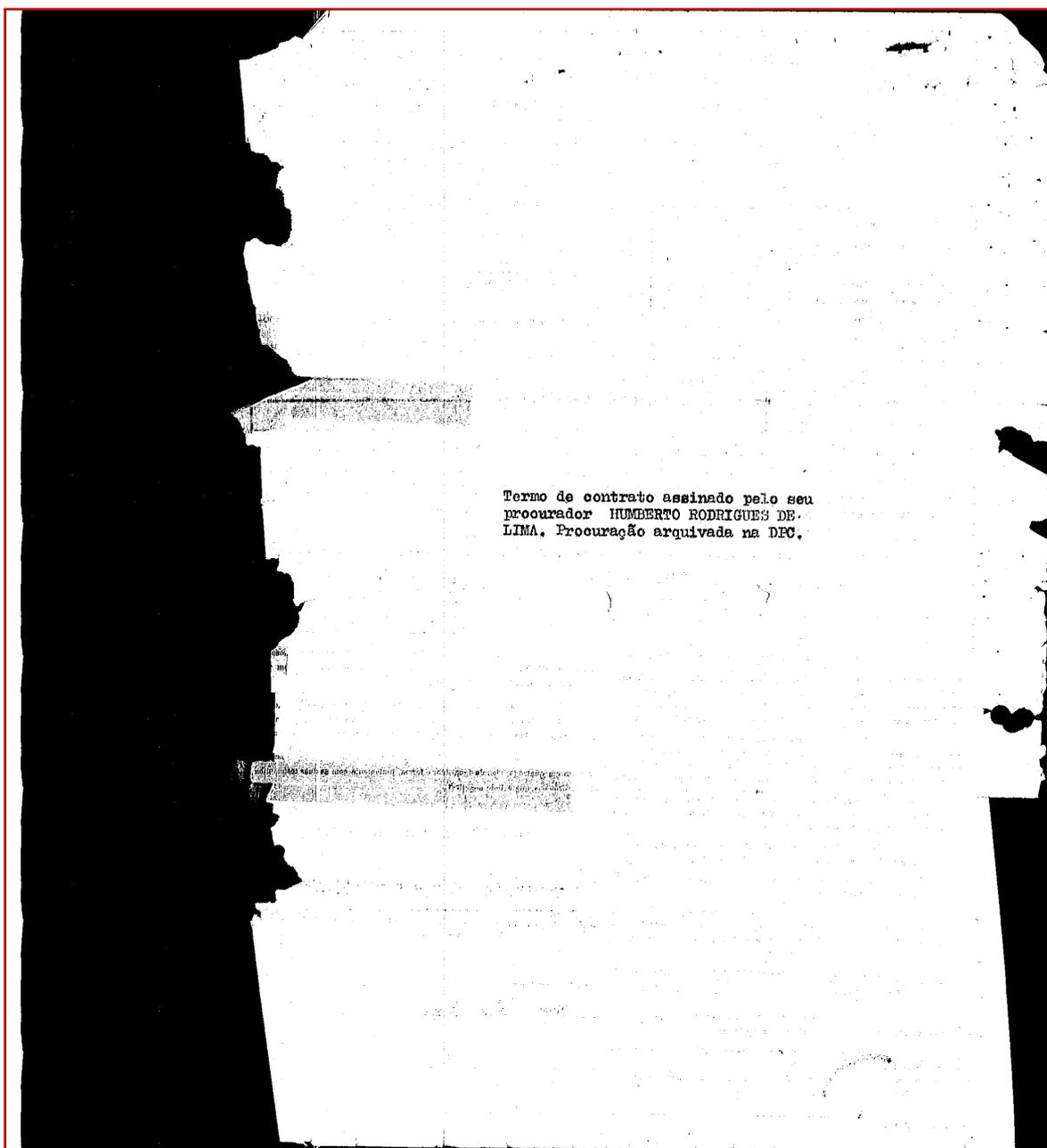


Figura 6 - Exemplo de documento digitalizado por um scanner de produção. Borda irregular à esquerda na parte inferior devido a dobra no documento durante a digitalização, com linhas brancas verticais.



**Figura 7 - Exemplo de documento digitalizado por um scanner de *produção*.
Bordas irregulares à direita e à esquerda da imagem devido a danos no documento físico.**

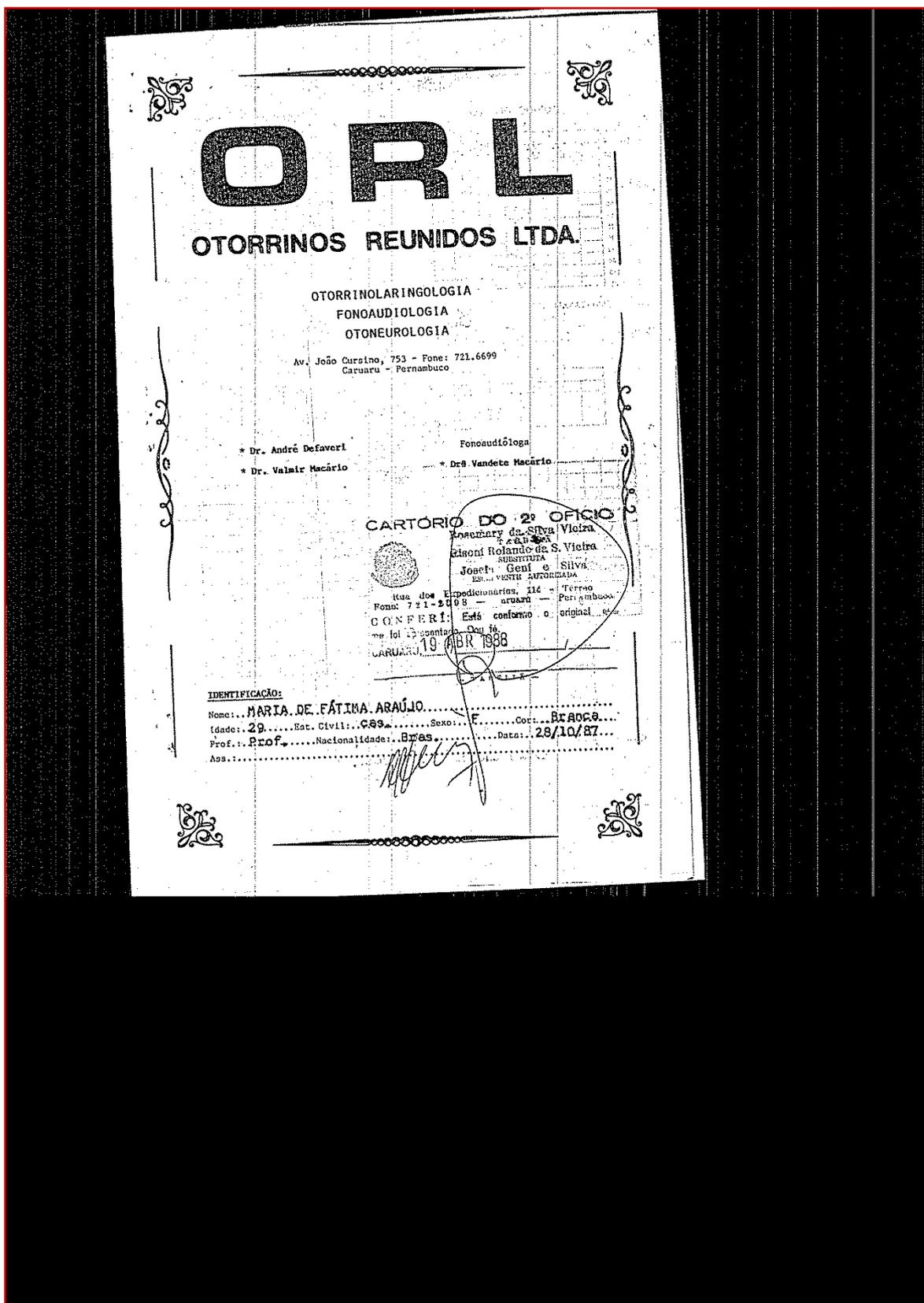


Figura 8 - Exemplo de documento digitalizado por um scanner de produção.
Borda ruidosa circunda quase completamente a imagem do documento,
apresenta linhas brancas verticais na margem direita.

2.2 – Estado da Arte

Existem muitos algoritmos para a melhoria de imagens monocromáticas de documentos [01] [02] [07], dentre eles se destaca o algoritmo *Ávila-Lins* [01] [08] por sua confiabilidade e alto desempenho, tanto pelos resultados qualitativos quanto pela velocidade de processamento por imagem. Recentemente, foi apresentada, por Formiga-Lins [01], uma proposta de pré-processamento para melhoria do desempenho do processamento por imagem de *Ávila-Lins* [03].

2.2.1 – O algoritmo *Ávila-Lins* para remoção do ruído de borda ruidosa

O algoritmo *Ávila-Lins* para remoção do ruído de borda “preta” é baseado no algoritmo de preenchimento e adaptado para segmentar a informação da borda preta [02] [03] [04]. Considera-se que as informações de um documento monocromático apresentam-se principalmente nas letras ou caracteres do texto. Outros elementos não-textuais, como figuras, tabelas, gráficos, podem apresentar informação, contudo, são menos frequentes. Os principais passos do algoritmo *Ávila-Lins* são explicados a seguir.

```

procedure remove_black_border(image, SEGMENT, LINE, CONNECT)
begin
  segment_list := create_horizontal_runs(image);
  mark_horizontal_runs(image, segment_list, SEGMENT);
  mark_vertical_runs(image, segment_list, SEGMENT);
  block_list := component_labelling_4x4(image, segment_list);
  adj_graph := region_adjacency_graph(block_list);
  classify_list := create_list_of_blocks_at_limit(block_list);
  while there exists block b in classify_list do
    begin
      if b.mark is HORIZONTAL_MARK then
        begin
          prj := calc_vertical_projection(block_list, adj_graph, b);
          if prj < LINE then
            b.mark := BORDER;
          end
        else
          if b.mark is VERTICAL_MARK then
            begin
              prj := calc_horizontal_projection(block_list, adj_graph, b);
              if prj < LINE then
                b.mark := BORDER;
              end;
            b.visited := true;
            if b.mark is BORDER then
              add_marked_neighbour_blocks(block_list,
                adj_graph, classify_list, b);
            end;
          process_special_pattern(image, blocks, adj_graph, segment_list, CONNECT);
          substitute_marks(image, blocks);
        end;

```

Pseudocódigo 1 – Algoritmo de remoção de bordas pretas *Ávila-Lins* [2].

Tendo em vista que a borda preta caracteriza-se por estar nos limites da imagem, o algoritmo inicia-se pelo preenchimento dos pixels pretos conectados entre si em pixels brancos a partir dos limites da imagem (Figura 9a). Observa-se que todos os elementos do documento, exceto as figuras, são formados por sequências de pixels pretos que, em média, têm um comprimento relativamente pequeno e constante para uma dada resolução. Portanto, para segmentar a informação da borda preta, o algoritmo verifica estatisticamente se, para cada sequência vertical ou horizontal da borda preta de tamanho igual ou inferior ao parâmetro SEGMENT, o componente segmentado é informação (Figuras 9b e 9c): (a) caso seja, o algoritmo de preenchimento não ultrapassa esta carreira; (b) caso contrário, o algoritmo de preenchimento prossegue com a sequência.

A análise estatística realizada para classificar o componente segmentado como informação deve levar em conta outros elementos não-textuais que aparecem na borda “preta”, como por exemplo, pequenos rasgões e excessivos ruídos. Nestes casos, tais elementos devem ser removidos. De fato, uma vez que o componente é segmentado, é possível extrair diversas características, como área, número de sequências, tamanho médio das sequências, etc. e aplicá-las em classificadores mais complexos, como redes neurais por exemplo. No entanto, o filtro proposto utiliza apenas uma característica: o tamanho da projeção horizontal ou vertical contado da posição da carreira ao final do componente segmentado (Figura 10). Se o tamanho da projeção for maior que o parâmetro LINE, então o componente segmentado é classificado como informação, caso contrário, é classificado como borda preta.

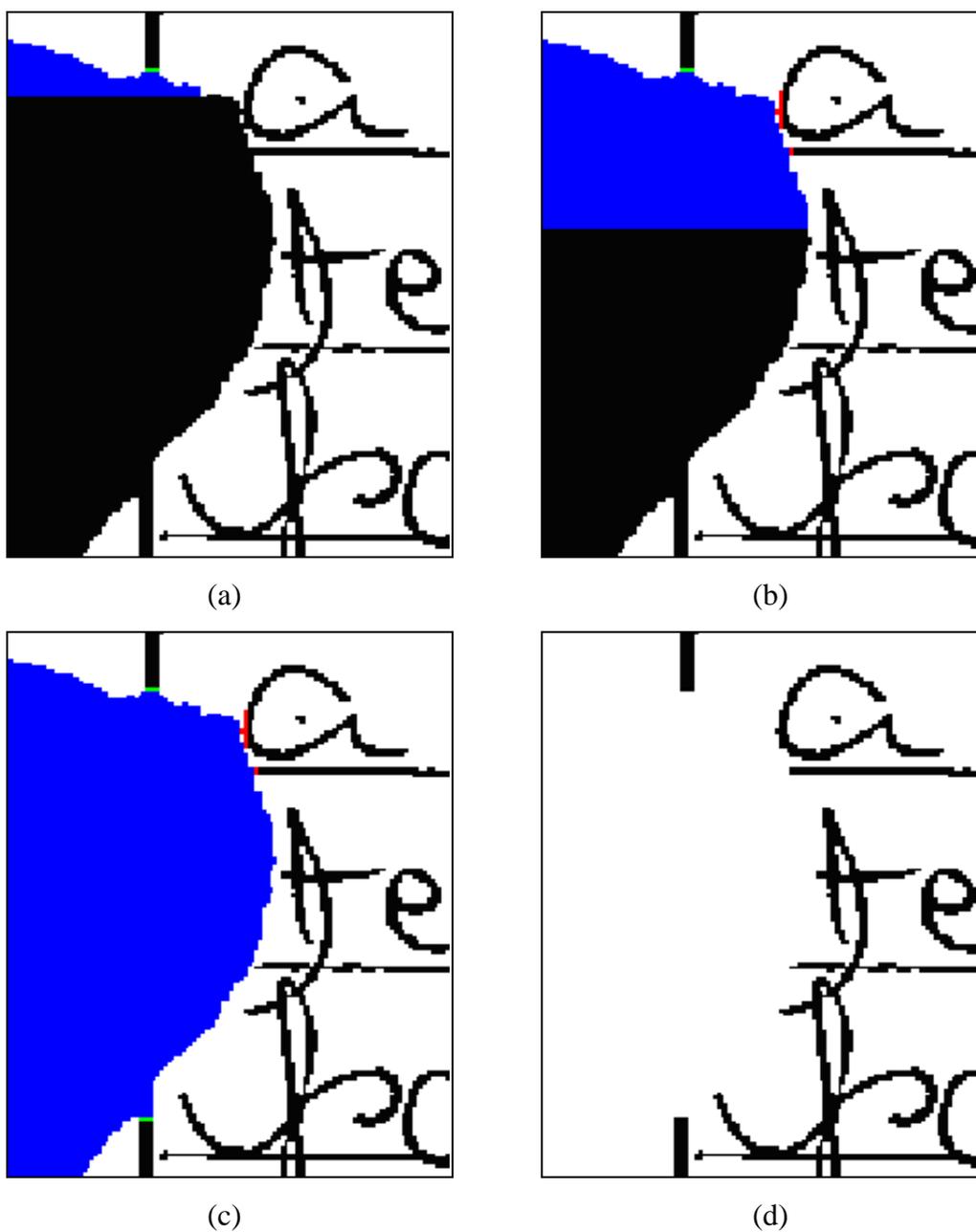


Figura 9 - Passos do Algoritmo Ávila-Lins para remoção de borda preta: (a) pixels em azul classificados como borda preta; (b, c) carreiras em vermelho e verde segmentam a informação da borda preta; (d) borda preta removida.

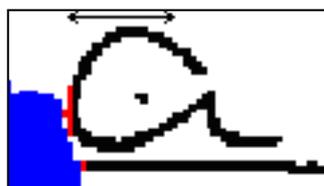


Figura 10 - Tamanho da projeção horizontal a partir da sequência marcada é utilizado para classificar o componente segmentado.

Os pixels da borda preta são *conectados 4x4 entre si* [16]. No entanto, alguns ruídos podem aparecer na borda de modo a se isolar (Figura 11). Este tipo de ruído pode ser caracterizado como borda preta se os pixels em sua volta também forem classificados como tal. Para garantir que apenas os componentes menores sejam classificados como borda preta, o filtro classifica apenas os ruídos que estão entre dois componentes classificados como borda preta a uma distância menor que o parâmetro CONNECT.



Figura 11 - Borda preta com ruídos pretos isolados (ilhados)

Apesar da alta precisão e bons resultados do algoritmo Ávila-Lins, seu desempenho é fortemente prejudicado por ruídos na borda preta do tipo *sal e pimenta* [16], que impedem a detecção do ruído de borda [01] [02], e pelo tamanho da imagem do documento, que devido ao padrão de preenchimento (*flood fill*) [01] [02] [19], utilizado pelo Ávila-Lins, necessita que o documento seja completamente analisado, aumentando o tempo de processamento por arquivo. Estas são as maiores limitações do algoritmo Ávila-Lins.

O algoritmo Ávila-Lins foi implementado em linguagem C, segundo descrição do código em [02], para o sistema operacional LINUX.

2.2.2 – Pré-processamento para o algoritmo Ávila-Lins

Uma forma sugerida para melhoria do desempenho do algoritmo Ávila-Lins é a remoção da maior parte do ruído de borda preta [01] [02], tendo como consequência a redução da área da imagem, antes da aplicação do algoritmo Ávila-Lins. Uma proposta de algoritmo de pré-processamento foi concebida e implementada por Formiga e Lins [01].

A ideia é a aplicação da etapa de detecção e remoção do ruído de borda do algoritmo Ávila-Lins, porém em uma versão reduzida da imagem original, para definição dos limites entre a imagem do documento e o ruído de borda que o circunda. As coordenadas dos limites encontradas na imagem reduzida são projetadas na imagem original, onde a área externa a estas coordenadas é então removida (*crop*) [01].

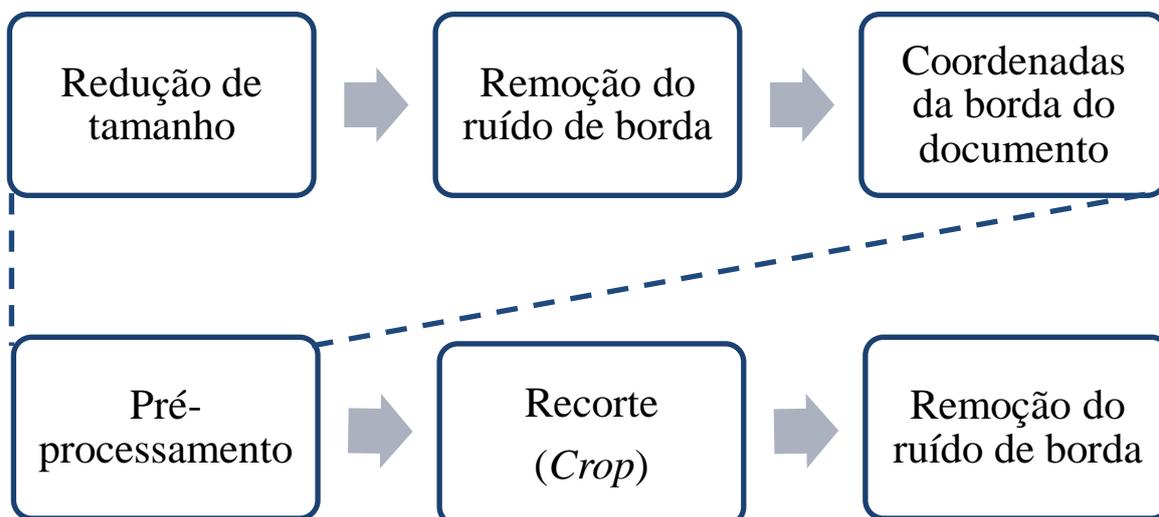


Figura 12 - Diagrama de blocos da remoção do ruído de borda com pré-processamento.

O objetivo do pré-processamento é a redução da área do ruído de borda em torno da região de interesse na imagem que contém as informações úteis do documento.

A Figura 12 define os passos do pré-processamento:

- **Redução de tamanho** – a partir da imagem original é criada outra imagem de menor tamanho a partir de subamostragem, onde algumas linhas e colunas de pixels são descartadas. Este é um processo simples e rápido, onde a imagem original é subamostrada a uma taxa r , isto significa que somente uma a cada r linhas e somente uma a cada r colunas são preservadas, obtendo uma imagem de menor tamanho, com uma área reduzida por um fator r^2 . Segundo [01], r deve ser no máximo igual a dois ($r=2$). Valores maiores de r irão tornar o pré-processamento mais rápido, porém aumentarão o risco de perda de informação relevante da imagem.
- **Remoção do ruído de borda** – esta etapa é realizada pelo algoritmo Ávila-Lins, mas não tem o objetivo de remover a borda da imagem de tamanho reduzido. Visa apenas determinar as coordenadas dos limites entre o ruído de borda e área de interesse que contém a imagem do documento.
- **Coordenadas da borda do documento** – a etapa anterior retorna quatro coordenadas definindo um retângulo que engloba completamente a área de interesse. Estas coordenadas são relativas à imagem de tamanho reduzido, sendo necessária uma transformação de coordenadas para que possam ser utilizadas para definir a área que será removida no documento original.

A imagem reduzida é obtida de uma simples transformação escalar com um fator escalar de $1/r$. Então um ponto qualquer (x,y) na imagem reduzida, corresponde a um ponto $(r.x,r.y)$ na imagem original.

Após o pré-processamento, é efetuado o recorte (*crop*) na imagem original, onde a área externa ao retângulo formado pelas coordenadas definidas no pré-processamento é removida. Em seguida é utilizado o algoritmo Ávila-Lins, agora com o objetivo de remover completamente o restante do ruído de borda (Figura 13).

O processo de criação de uma imagem de menor tamanho por subamostragem seguindo da aplicação do algoritmo Ávila-Lins para determinação das coordenadas de uma área de interesse mostra-se em média duas vezes mais rápido, conforme será mostrado quantitativamente mais adiante. A eliminação da área externa ao retângulo encontrado (*crop*) é uma operação bastante simples, com duração irrelevante para o tempo de execução do processo total. A eliminação parcial do ruído borda antes da aplicação do algoritmo Ávila-Lins mostrou-se bastante efetiva para a redução do tempo de execução do processo [01] [02].

No entanto, a utilização do algoritmo Ávila-Lins para determinação das coordenadas da área de interesse mantém as mesmas limitações deste algoritmo quanto à presença de ruídos intensos do tipo *sal e pimenta* na borda externa a área de interesse.

O algoritmo Formiga-Lins também foi implementado em linguagem C, e o código foi gentilmente cedido pelo Dr. Andrei Formiga [01], para o ambiente LINUX.

3 – Uma nova proposta de pré-processamento

O algoritmo Ávila-Lins possui um ótimo desempenho e poucas limitações, o que justifica o esforço para a criação de algoritmos que o utilizem como base.

A proposta definida aqui descreve um algoritmo de pré-processamento que além de possuir um excelente desempenho quanto ao tempo de pré-processamento, não é limitado pelo ruído sal e pimenta. Há técnicas para remoção deste tipo de ruído, como o filtro de vizinhança. Para cada conjunto de pixels pretos conectados entre si de dimensões $k \times k$ na imagem binária, verificam-se os 4 $(k+1)$ pixels vizinhos. Se não houver nenhum pixel vizinho preto, o conjunto de pixels é considerado ruído e então, marcado de branco. Pixels brancos em um fundo preto também podem ser considerados ruídos e vice-versa [02].

Outra técnica chamada de *k-fill* é sugerida por O’Gorman [11]. O método é uma extensão do filtro de vizinhança. Para cada conjunto de pixels pretos conectados entre si de dimensões $(k-2) \times (k-2)$ na imagem binária, calcula-se: n , o número de pixels pretos na vizinhança; c , o número de carreiras de pixels pretos conectados na vizinhança e; r , o número de pixels pretos nos cantos. A condição para preencher os pontos interiores de branco é:

$$(c = 1) \text{ and } (((n > 3k - 4) \text{ or } (n = 3k - 4)) \text{ and } (r = 2))$$

O filtro *k-fill* foi desenvolvido especificamente para textos e gráficos com a finalidade de remover ruído enquanto mantém a legibilidade [02].

As técnicas citadas anteriormente para a remoção do ruído sal e pimenta são lentas, devido à necessidade de realizar verificações em todos os pixels da imagem [02].

A ideia do novo algoritmo aqui apresentado surgiu para eliminar as limitações com o ruído de borda combinado ao ruído sal e pimenta, onde a alternância entre pixels brancos e pretos é completamente aleatória, assemelhando-se a “uma estrada com muitos buracos”, se observada do ponto de vista de uma área preta com textura sólida e muitas falhas.

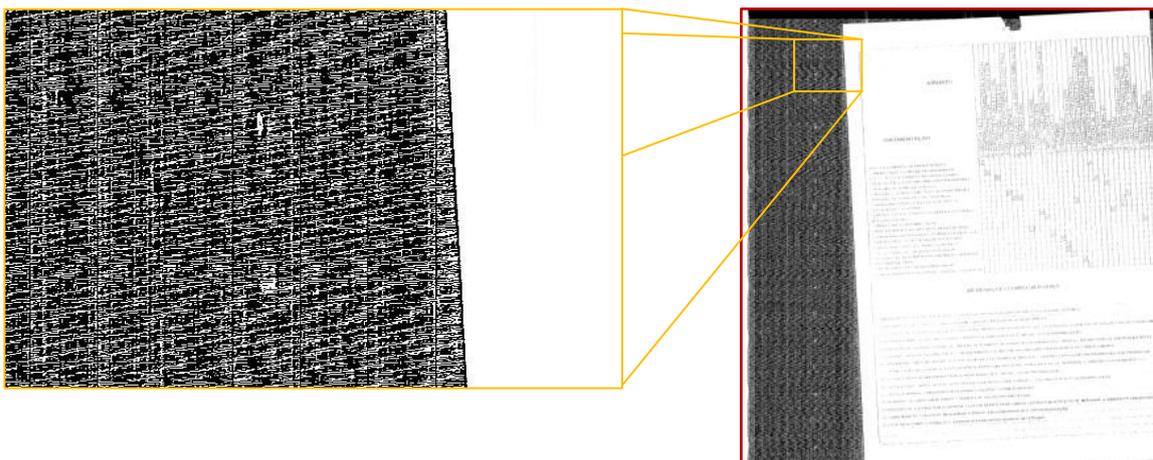


Figura 13 - Detalhe da área de uma imagem com ruído de borda combinado com o ruído sal e pimenta.

A partir dessa analogia, a solução para trafegar melhor em uma estrada com muitos buracos é utilizar um veículo com pneus largos ou mesmo que use esteiras. Então, é aplicando uma técnica morfológica [16] simples para detecção da borda, a partir da extração de fronteiras, da região de interesse, porém ao invés de deslocar-se pixel a pixel, é utilizado um deslocamento de grupos de pixels por vez (uma esteira).

Analisando as características descritas no capítulo 2, seção 2.1, desta dissertação para o ruído de borda, foi possível elaborar uma estratégia para a detecção da borda da área de interesse, baseada na extração de fronteiras:

- A partir do topo do documento, efetuar uma varredura no sentido da esquerda para a direita, até localizar grupos de pixels pertencentes ao documento, ou seja, um grupo contíguo de pixels brancos.
- Observar que continuando a varredura na mesma linha de pixels, mesmo que a imagem do documento esteja rotacionada, a fronteira do lado esquerdo será detectada (Figura 14), não havendo a necessidade de uma varredura na mesma linha no sentido inverso, da direita para a esquerda.
- A varredura na linha é efetuada utilizando um bloco de pixels, de um pixel por linha por 8 pixels por coluna. Reduzindo a largura do documento em um fator de 8 ($largura/8$). Este grupo de pixels é visto como sendo um byte padrão (8 bits), organizados em uma única variável de 8 bits. Para verificar se cada bloco é um conjunto contíguo de pixels pretos ou brancos, basta verificar se o valor do bloco em teste é 0 ou 255, dependendo de como a imagem está configurada para representar pixels pretos e brancos [25].

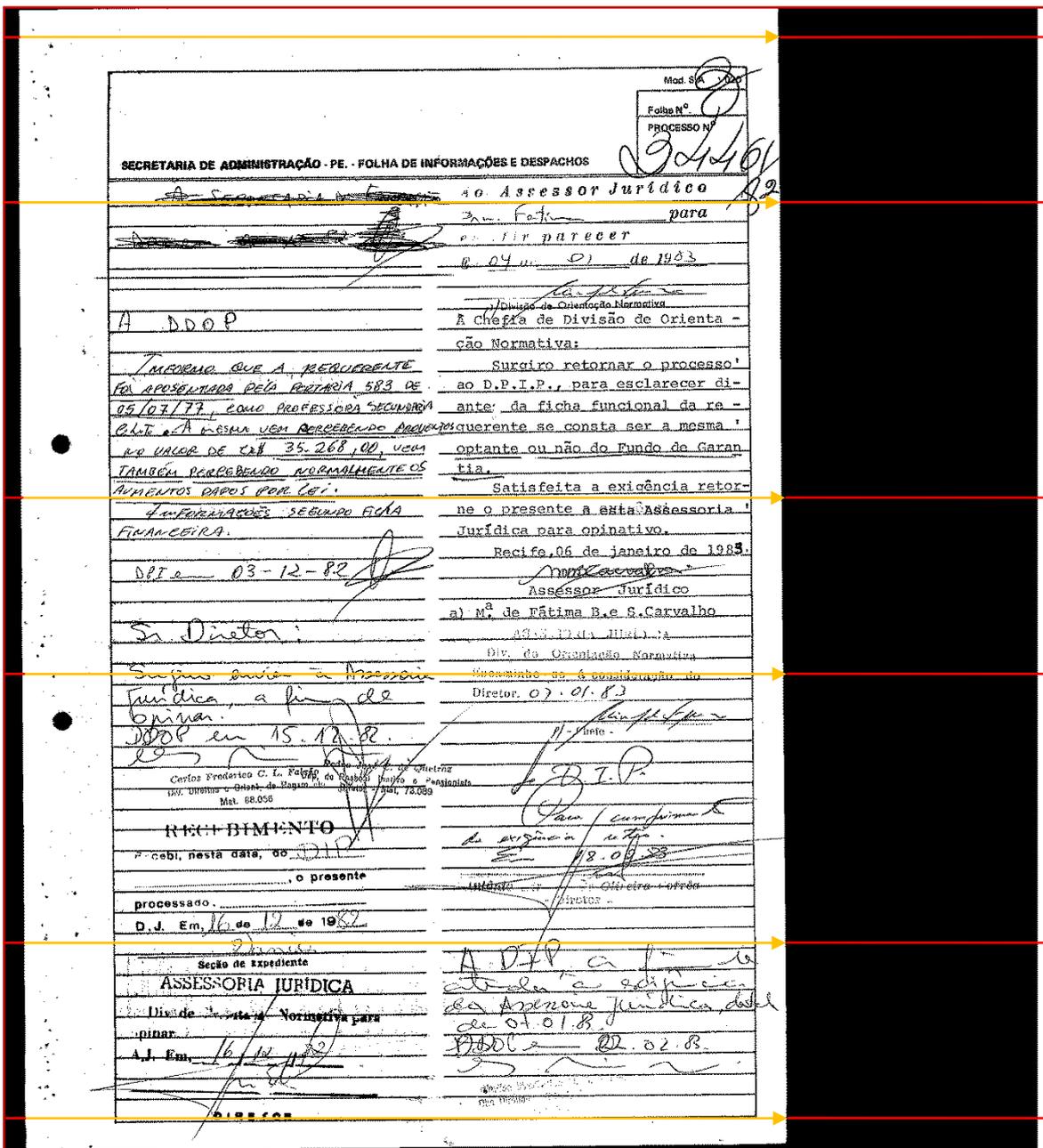


Figura 14 - Sentido de varredura (setas amarelas) para extração da fronteira da região de interesse que contém a imagem do documento.

Foi observado que não há a necessidade de efetuar a varredura ao longo de toda a altura da imagem, apenas em três partes, tentando cobrir a maior parte da região de interesse.

Os gráficos abaixo mostram a distribuição dos bits e grupos de oito bits em algumas linhas da imagem da Figura 14. Os bits pretos assumem valor “0”, enquanto os bits brancos assumem valor “1”. Dessa forma, grupos de oito bits pretos contínuos terão valor “0”, enquanto grupos de oito bits brancos contínuos terão valor “255”.

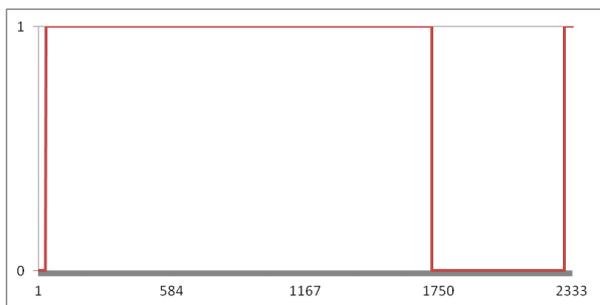


Gráfico 2 - Distribuição dos bits em uma linha na parte superior da imagem na Figura 14.

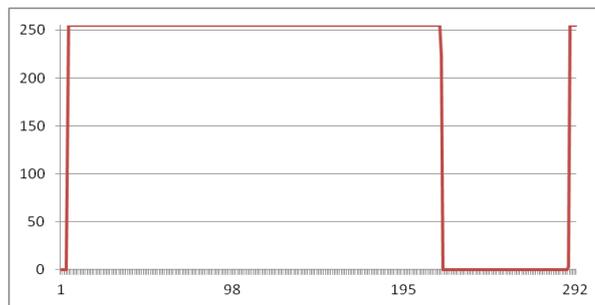


Gráfico 3 – Distribuição por grupos de oito bits em uma linha na parte superior da imagem na Figura 14.

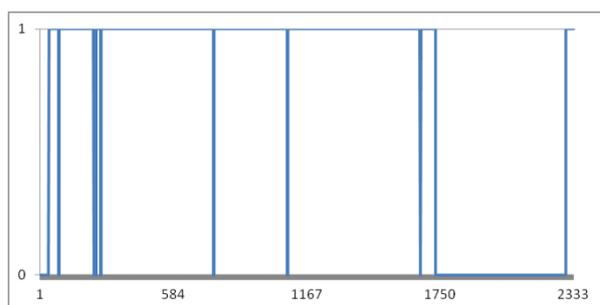


Gráfico 4 - Distribuição dos bits em uma linha na parte do meio da imagem na Figura 14.

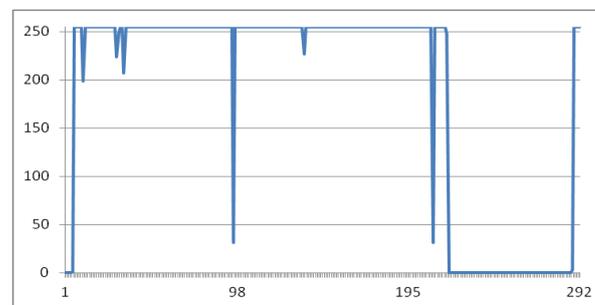


Gráfico 5 - Distribuição por grupos de oito bits em uma linha na parte do meio da imagem na Figura 14.

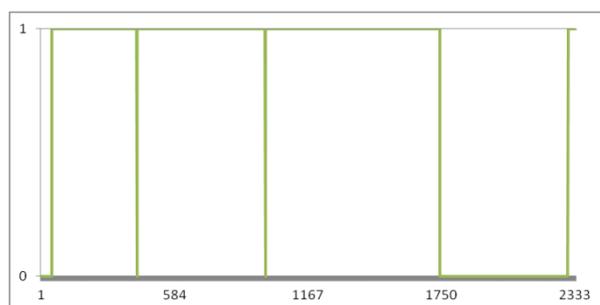


Gráfico 6 - Distribuição dos bits em uma linha na parte inferior da imagem na Figura 14.

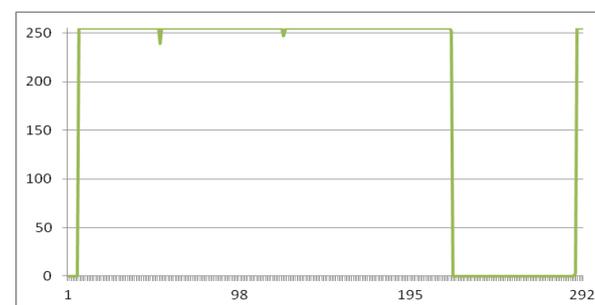


Gráfico 7 - Distribuição por grupos de oito bits em uma linha na parte do meio da imagem na Figura 14.

Os gráficos da distribuição por grupos de oito bits mostram uma melhor definição da região de interesse com a atenuação da interferência por pixels pretos no seu interior. Além da redução da largura da linha a ser analisada, reduzindo o tempo de análise de cada linha.

Devido à imensa variedade de irregularidades na região de interesse, causados por formatos variados dos papéis dos documentos, danos físicos nos documentos digitalizados, falhas na digitalização entre outros, torna-se necessária a aquisição de outros dados e comparação dos dados encontrados para validação dos mesmos, de forma a não comprometer as informações na região de interesse. Dessa forma, há uma varredura para encontrar uma largura, o que significa dois pontos, na parte superior da imagem. Uma varredura para encontrar uma largura na parte média da imagem e por último uma varredura para encontrar uma largura na parte inferior da imagem. A varredura na parte superior inicia na primeira linha e avança da esquerda para a direita, até encontrar um conjunto contíguo de blocos de pixels brancos, característica da região de interesse. Quando se encontra esse conjunto, o ponto imediatamente anterior ao início da sequência de blocos é marcado (sa), e a varredura continua até encontrar um conjunto contíguo de blocos pretos ou não completamente brancos ou mesmo a borda da imagem. Caso encontre um conjunto contíguo de blocos, o primeiro ponto do conjunto é marcado, caso chegue ao limite da imagem, este ponto do final da linha é marcado (sb). Juntamente com o primeiro ponto encontrado a esquerda (sa). A diferença entre sb e sa define a largura da região de interesse ($sb-sa$).

Uma largura válida pode ser definida como uma fração da largura total da imagem, mas, em muitos casos, a região de interesse ocupa toda a imagem. Há também casos em que a região de interesse é muito menor que a área total da imagem, o que pode implicar em uma largura e/ou altura muito inferior à largura e/ou altura da imagem total. Então, a partir da largura e altura da imagem completa são definidos os pontos de início das varreduras e a largura considerada válida para definir a região de interesse. Esta abordagem permite escalonar os parâmetros para a varredura, ou seja, é possível definir a busca por certa margem de largura, caso uma primeira varredura não encontre pontos válidos é iniciada uma nova varredura partindo do ponto inicial, porém com novos parâmetros de busca, e assim por diante, o quanto for julgado necessário. No entanto, não é aconselhável utilizar muitos escalonamentos ou o desempenho no tempo de processamento será prejudicado.

Caso não sejam encontrados conjuntos contíguos de pixels em uma linha onde está ocorrendo uma varredura, ao chegar a determinado ponto, definido pelos parâmetros do escalonamento, a varredura é interrompida e inicializada na próxima linha. Até que sejam encontrados pontos que satisfaçam os parâmetros do escalonamento.

Se uma leitura falha e não consegue encontrar nenhum ponto que satisfaça os parâmetros do escalonamento em alguma das áreas definidas na imagem, superior, média ou inferior, ainda é possível utilizar os pontos das outras duas áreas para realizar o recorte (*crop*), caso tenham sido obtidas com sucesso. Removendo mais de 50% em média a borda preta. Mas, caso haja falha na leitura em pelo menos duas áreas, não é possível efetuar o recorte da imagem com segurança. A imagem permanecerá inalterada (*bypass*).

Os parâmetros de busca da varredura são definidos de acordo com o tipo de digitalizador utilizado. Neste caso foram utilizados parâmetros para documentos de tamanho A4 digitalizados por alimentação automática em um *scanner* de produção [21] com área útil de tamanho A3. Uma largura válida foi definida como sendo maior ou igual a dois terços da largura total da imagem, tendo no máximo a largura da imagem. Para escalonamento, foi definido como segundo parâmetro uma largura válida sendo igual ou maior que metade da largura da imagem.

Além de procurar por larguras válidas, as varreduras na parte superior e inferior podem indicar as bordas horizontais do documento em suas respectivas partes (Figuras 15 e 16). Basta que seja marcada a primeira linha onde seja identificada uma largura válida. Essa linha será a primeira linha da área de interesse. No caso de documentos que estejam rotacionados na imagem, a largura da primeira linha não irá coincidir com uma largura válida, então para encontrar a fronteira horizontal nas partes superior e inferior da região de interesse é definida uma nova largura válida, apenas para detecção das bordas superior e inferior. Ou seja, durante a varredura, caso seja encontrada uma região que defina uma determinada largura, porém não alcance o mínimo de uma largura válida, esta poderá ser definida como sendo a primeira linha da região de interesse. Neste caso, como primeiro parâmetro, esta largura foi definida como metade da largura total e no escalonamento um quarto da largura total.

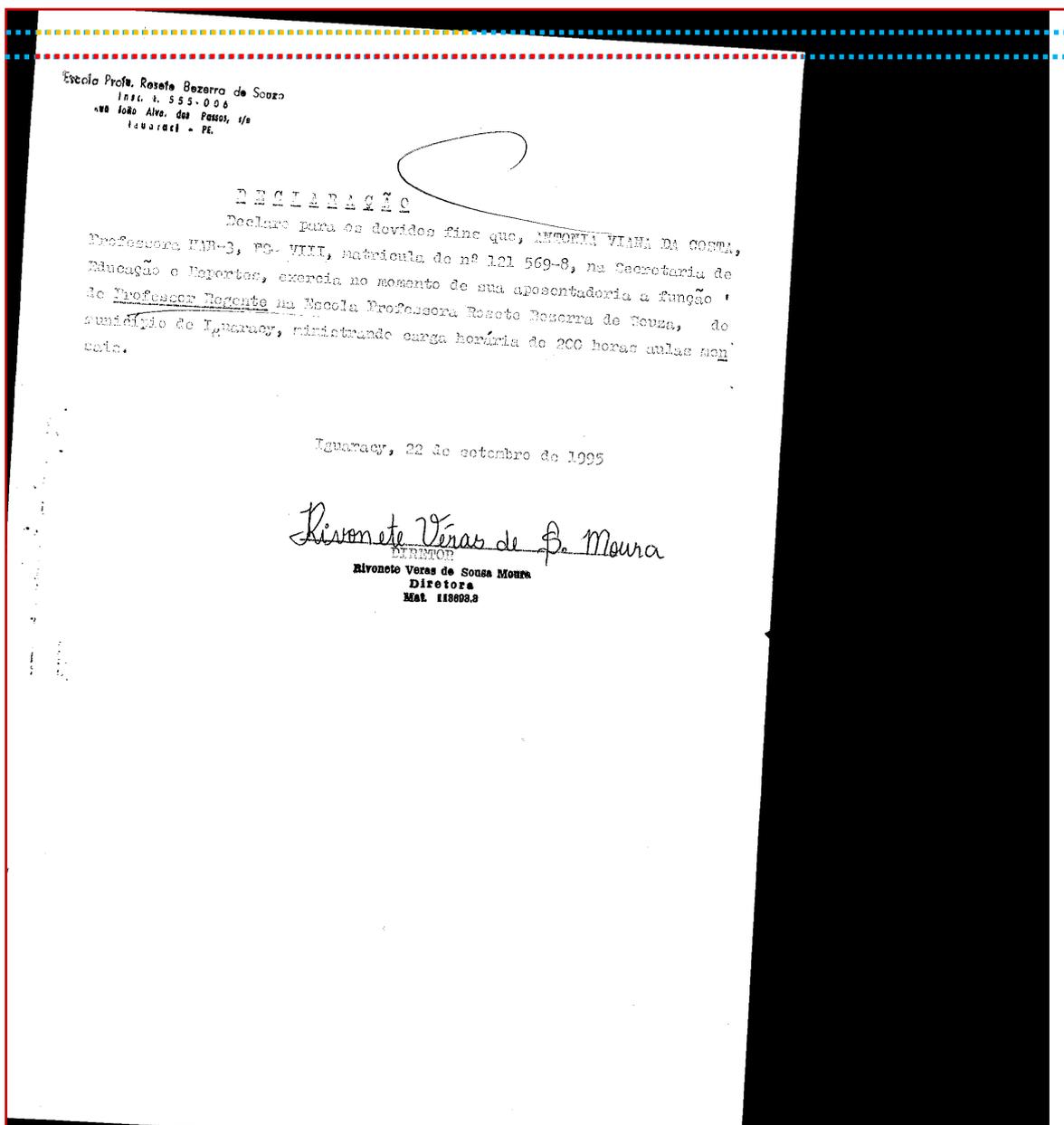


Figura 15 - Determinação das fronteiras horizontais, em amarelo, linha que define a fronteira horizontal superior.
 Em vermelho, largura válida para região de interesse definida na área superior da imagem.

O conjunto de blocos contíguos é definido como uma faixa com largura máxima de um por cento da largura total da imagem. Tal parâmetro mostrou-se bem adequado nos testes.

Em casos onde há uma faixa branca vertical no meio do ruído de borda, essa possui largura muito inferior aos parâmetros definidos, porém pode ter largura maior ou igual a um por cento da largura total da imagem. Nesses casos, como é detectada uma largura

muito inferior a uma largura válida, são definidos novos pontos de início da varredura, de forma a ignorar essa região indesejada. Os novos pontos passam a ser os pontos imediatamente após esta faixa, ou seja, a faixa passa a ser uma poderosa indicação do ruído de borda.

A varredura no meio da imagem deve começar na linha de pixels que esteja 50% da altura total da imagem ($\text{altura}/2$), e seguir em direção à parte inferior, da mesma forma e com os mesmos parâmetros da varredura para a parte superior. Com exceção da busca por limites horizontais. A definição de uma largura na parte média da imagem (ma , mb) mostrou-se muito importante, pois nesta região, em geral, há uma ocorrência menor de falhas. Sendo esses pontos cruciais para a validação dos demais pontos encontrados.

A varredura na parte inferior começa na última linha, e, ao contrário das outras varreduras, deve seguir em direção à parte superior. Os mesmos parâmetros continuam sendo usados, exceto o fato de que quando não for possível localizar dados válidos em uma linha, a próxima linha a ser analisada será a linha imediatamente acima. Para definição da largura, novamente, são localizados dois pontos (ia , ib).

Na parte inferior também é definida uma fronteira horizontal da região de interesse, da mesma forma que na parte superior (Figura 16).

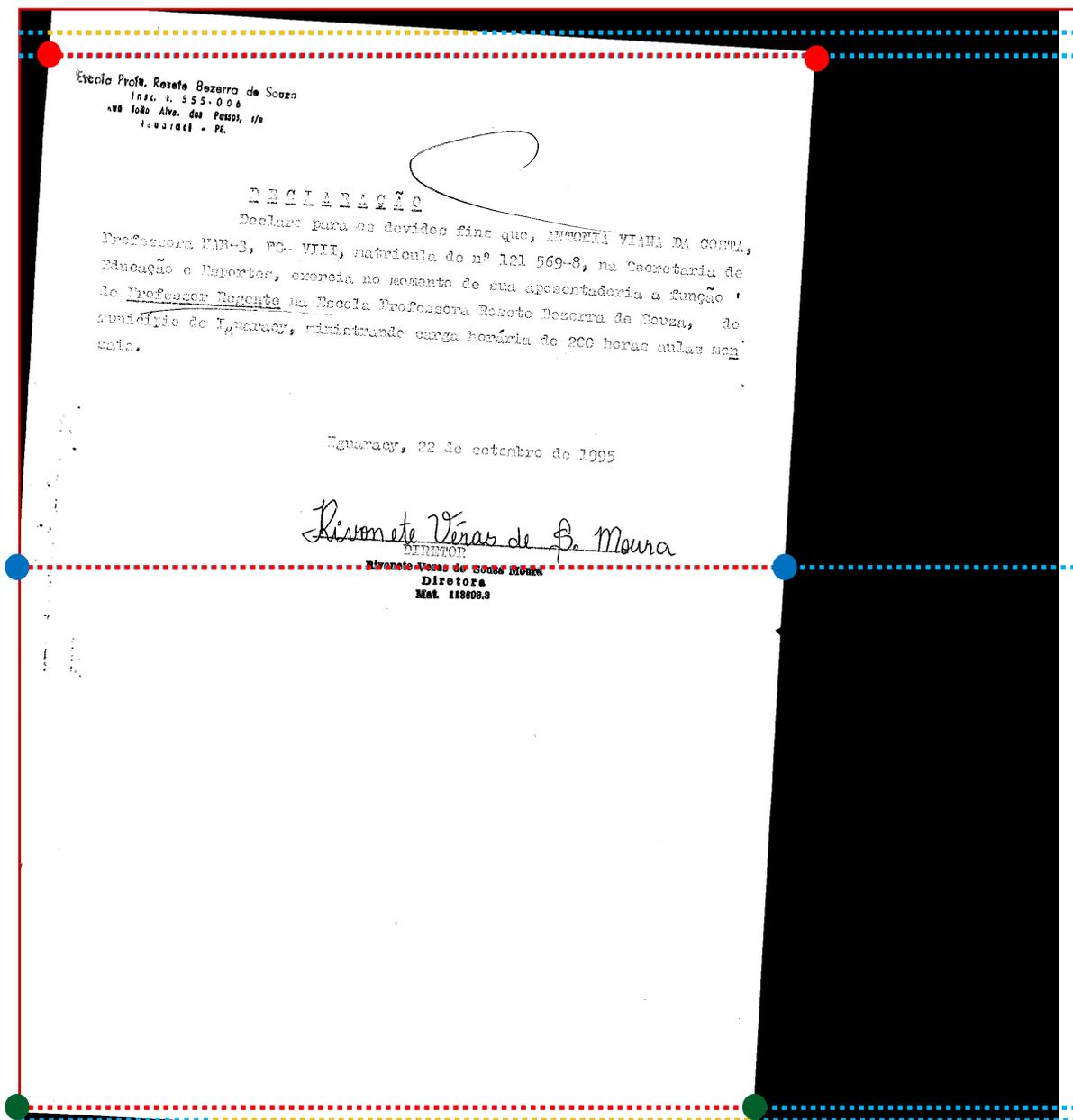


Figura 16 - Varredura em todas as faixas pré-determinadas. Em vermelho as largura encontradas; *sa* à esquerda e *sb* à direita (pontos em vermelho) na parte superior; *ma* à esquerda e *mb* à direita (pontos em azul) na parte média, e *ia* à esquerda e *ib* à direita (pontos em verde) na parte inferior. Em amarelo, limites da altura da região de interesse.

Após a definição das larguras nas três partes, superior, média e inferior, os pontos são comparados e escolhidos de forma a definir uma largura que contenha completamente a região de interesse. Todos os pontos encontrados são projetados na primeira linha da região de interesse, dessa forma é mais fácil definir quem é o maior ou o menor entre eles.

É escolhido o menor dos pontos a (sa, ma, ia), ou seja, o que está mais à esquerda, e o maior dos pontos b (sb, mb, ib), ou seja, o que está mais à direita. Dessa forma teremos a maior largura possível definida por estes pontos. Essa escolha é a melhor forma de evitar cortes indesejados na região de interesse devido a rotações, formato do papel do documento e documentos danificados que ocasionam grande irregularidade nas bordas (Figura 17).

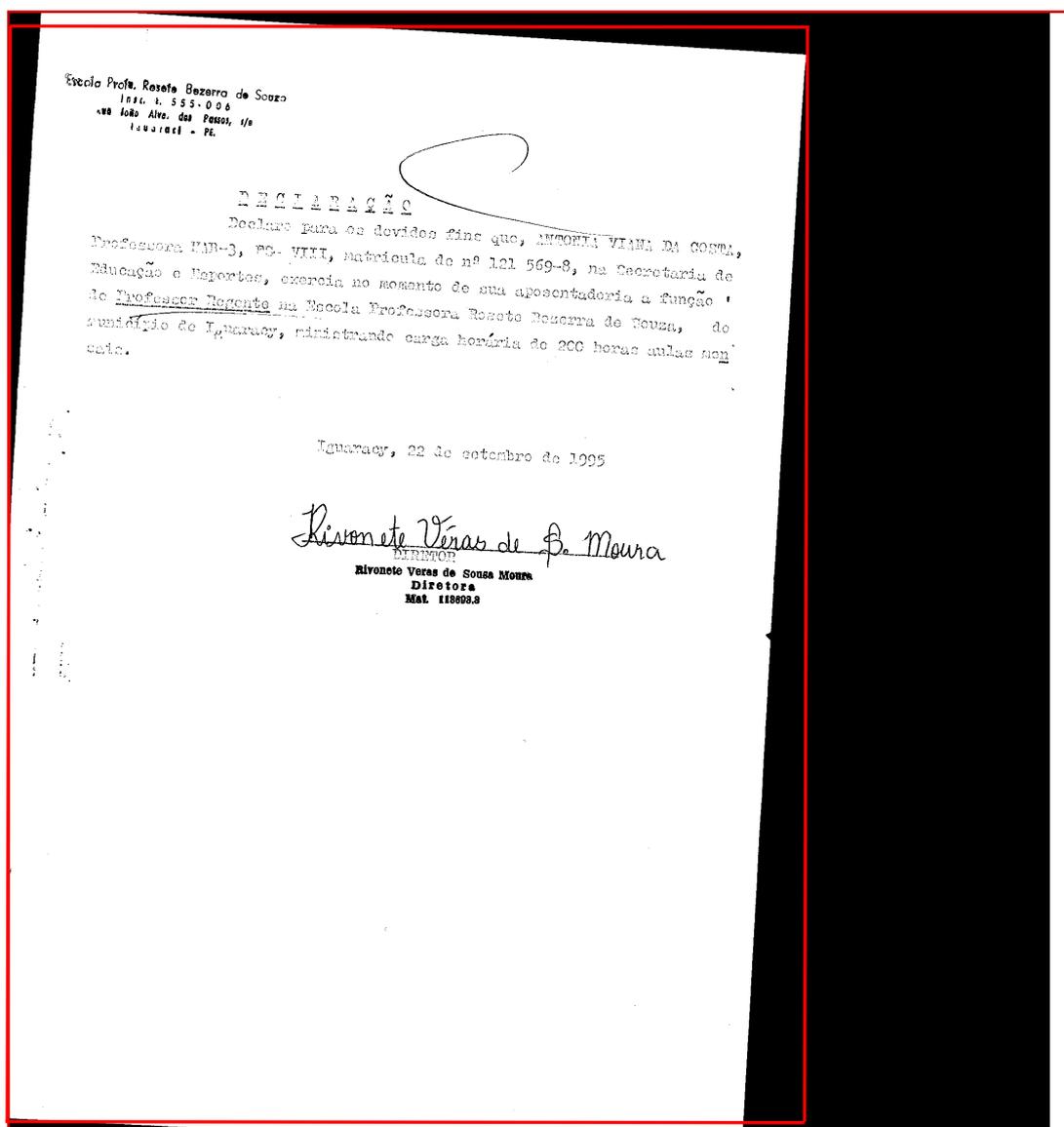


Figura 17 - Definição da região de interesse.

Com esses pontos a região de interesse é completamente definida, com largura, altura, ponto inicial de leitura e ponto final de leitura.

A região de interesse é então copiada em outro arquivo, operação que se mostrou mais rápida que a modificação do arquivo já existente.

3.1 - Implementação

O algoritmo foi implementado em linguagem C, com o uso da biblioteca LIBTiff [26] de código aberto para tratamento de imagens do tipo TIFF [25]. O código encontra-se no DVD em anexo, referenciado no apêndice A.

4 - Resultados

Os algoritmos de remoção de borda de Ávila-Lins, Formiga-Lins e o aqui apresentado foram implementados em linguagem C, compilados e executados em sistema operacional LINUX, distribuição de código aberto Ubuntu[®] 10.10. O computador utilizado tem a seguinte configuração:

- Processador Core 2 Duo T6600 Intel[®];
- 3GB de memória RAM PC 6400;
- sistema operacional LINUX, escolhido para melhor controle sobre os testes de desempenho quanto ao tempo de processamento, já que neste sistema operacional é possível ter controle total sobre os processos executados, possibilitando a medida do tempo de execução do processo de cada algoritmo individualmente para cada imagem.

A implementação de cada algoritmo foi testada em um conjunto de cinco mil imagens monocromáticas de um bit por pixel no formato TIFF com o algoritmo de compressão CCITT Group IV [25]. Para facilitar o acesso às imagens e ter uma melhor comparação, as imagens foram divididas em dez grupos de 500 imagens.

Foram realizadas análises quantitativas e qualitativas para comparação de desempenho. A análise quantitativa é relativa ao tempo de processamento da implementação de cada algoritmo e o tamanho final dos arquivos em bytes, pois um dos objetivos é redução do tamanho em bytes das digitalizações.

A análise qualitativa é relativa à qualidade final das imagens, sendo feita de forma manual, cada imagem resultante de cada algoritmo é comparadas entre si e com a imagem original. As imagens originais e resultantes encontram-se no DVD em anexo.

4.1 – Comparação dos resultados

Comparação de desempenho quanto ao tempo de processamento por imagem. O tempo de abertura e fechamento de cada arquivo não é aferido, uma vez que todas as implementações utilizam a mesma biblioteca e usam os mesmos procedimentos para realizar estas operações.

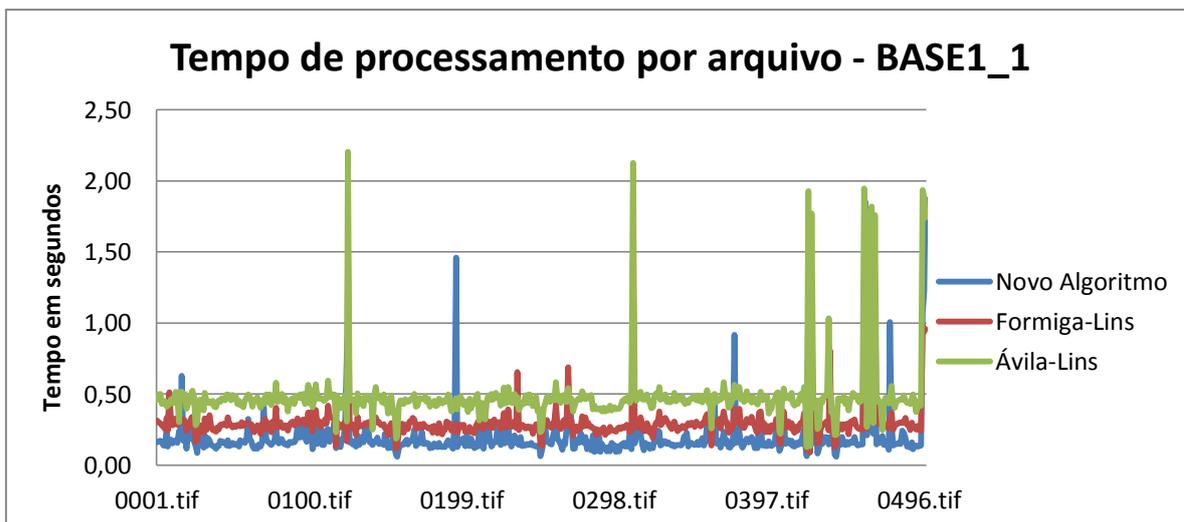


Gráfico 2 - Comparação do tempo de processamento por arquivo - grupo BASE1_1

BASE1_1			
	TEMPO MÉDIO POR ARQUIVO (s)	TEMPO TOTAL (s)	VARIÂNCIA
ÁVILA-LINS	0,484015	242,007599	0,053323
FORMIGA-LINS	0,301081	150,540320	0,019706
NOVO ALGORITMO	0,200209	100,104580	0,035463

Tabela 3 - Comparação do tempo médio de processamento por arquivo e do tempo total de processamento da BASE1_1

BASE1_1	
	REDUÇÃO DE TEMPO EM RELAÇÃO AO ÁVILA-LINS (%)
FORMIGA-LINS	37,80
NOVO ALGORITMO	58,64

Tabela 4 - Redução do tempo de processamento do Formiga-Lins e do Novo Algoritmo em relação ao Ávila-Lins na BASE1_1

BASE1_1		
	TAMANHO TOTAL (KB)	TAMANHO MÉDIO (KB)
ORIGINAL	19.660,80	39,32
ÁVILA-LINS	18.534,40	37,07
FORMIGA-LINS	18.432,00	36,86
NOVO ALGORITMO	18.124,80	36,25

Tabela 5 - Comparação do tamanho final da BASE1_1 de cada algoritmo em relação ao tamanho original em KB

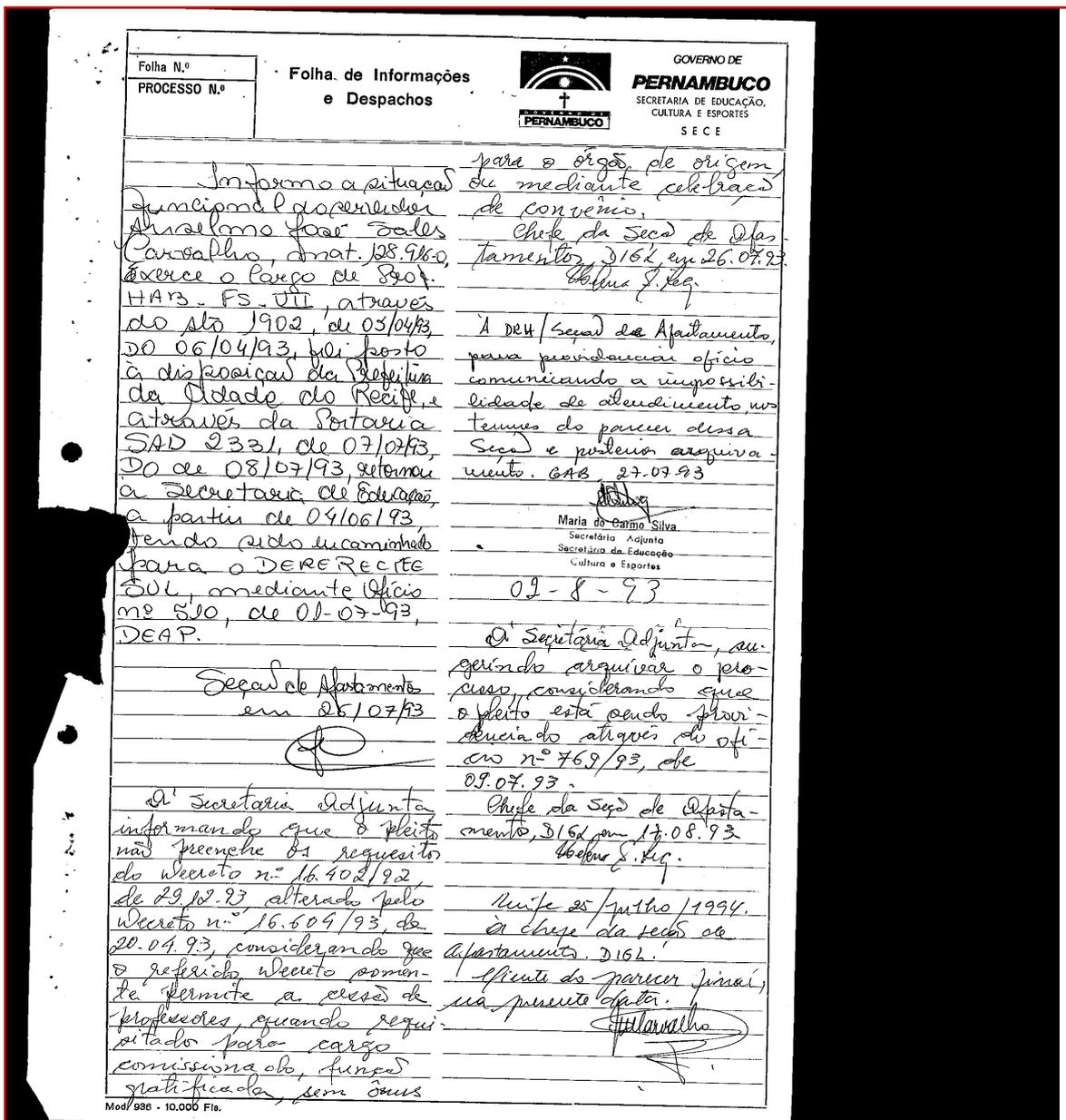


Figura 18 - Imagem 0047 da BASE1_1. Apresenta irregularidade na borda esquerda por danos no documento físico, borda ruidosa toca o texto. Tamanho de 62KB.

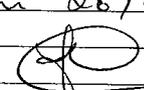
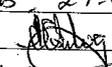
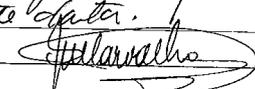
Folha N.º PROCESSO N.º	Folha de Informações e Despachos	 GOVERNO DE PERNAMBUCO SECRETARIA DE EDUCAÇÃO, CULTURA E ESPORTES S E C E
<p> Informo a situação funcional do servidor Arnaldo José Sales Carvalho, mat. 28.9160, Exerce o cargo de Prof. HAB - FS - VII, através do ato 1902, de 03/04/93, do 06/04/93, foi posto à disposição da Prefeitura da Cidade do Recife, através da Portaria SAD 2331, de 07/07/93, DO de 08/07/93, retornou a Secretaria de Educação a partir de 04/06/93, tendo sido encaminhado para o DERE RECTE SUL, mediante Ofício nº 510, de 01-07-93, DEAP. </p>		<p> para o órgão de origem de mediante celebração de convênio. chefe da Secd de Apartamentos, D162 em 26.07.93 Helma S. Riq. </p>
<p> Secd de Apartamentos em 26/07/93  </p>		<p> A DREH / Secd de Apartamentos, para providenciar ofício comunicando a impossibilidade de atendimento, nos termos do parecer dessa Secd e posterior arquivamento. CAB 27.07.93 </p>
<p> A Secretária Adjunta informando que o pleito não preenche os requisitos do Decreto nº 16.402/92, de 29.12.93, alterado pelo Decreto nº 16.604/93, de 20.04.93, considerando que o referido Decreto somente permite a cessã de professores, quando requisitado para cargo comissionado, função gratificada, sem ônus </p>		<p>  Maria do Carmo Silva Secretária Adjunta Secretaria da Educação Cultura e Esportes </p>
<p>Mod/ 936 - 10.000 Fls.</p>		<p> 02-8-93 A Secretária Adjunta, sugerindo arquivar o processo, considerando que o pleito está sendo providenciado através do ofício nº 769/93, de 09.07.93. chefe da Secd de Apartamentos, D162 em 17.08.93 Helma S. Riq. em 25/ julho /1994. in chefe da secd de Apartamentos. D162. Ofício do parecer final, na presente data.  </p>

Figura 19 - Resultado do algoritmo Ávila-Lins da imagem 0047 da BASE1_1. Tempo de processamento 0,490267s.

Informação tocada pela borda ruidosa, à esquerda, foi preservada. Tamanho 61KB.

Folha N.º PROCESSO N.º	Folha de Informações e Despachos	 PERNAMBUCO	GOVERNO DE PERNAMBUCO SECRETARIA DE EDUCAÇÃO, CULTURA E ESPORTES S E C E
<p> Informo a situação funcional do professor Arnaldo José Sales Carvalho, mat. 28.916-0, exerce o cargo de Prof. HAB - FS - VII, através do ato 1902, de 03/04/93, do 06/04/93, foi posto à disposição da Prefeitura da Cidade do Recife, através da Portaria SAD 2331, de 07/07/93, do de 08/07/93, retornou a Secretaria de Educação, a partir de 04/06/93, tendo sido encaminhado para o DERE/RECIFE SUL, mediante Ofício nº 510, de 01-07-93, DEAP. </p>		<p> para o órgão, de origem de mediante celebração de convênio. </p>	
<p> Secad de Apartamentos em 26/07/93 </p>		<p> Chefe da Secad de Apartamentos, D162 em 26.07.93 Volmar S. Reg. </p>	
<p> A Secretária Adjunta informando que o pleito não preenche os requisitos do Decreto nº 16.402/92, de 29.12.93 alterado pelo Decreto nº 16.604/93, de 20.04.93, considerando que o referido decreto somente permite a cessão de professores, quando requisitado para cargo comissionado, função gratificada, sem ônus </p>		<p> A DEH/Secad de Apartamentos, para providenciar ofício comunicando a impossibilidade de atendimento, nos termos do parecer dessa Secad e posterior arquivamento. CAB 27-07-93 </p> <p> Maria do Carmo Silva Secretária Adjunta Secretária de Educação Cultura e Esportes </p> <p> 09-8-93 </p> <p> A Secretária Adjunta, sugerindo arquivar o processo, considerando que o pleito está sendo providenciado através do ofício nº 769/93, de 09.07.93. </p> <p> Chefe da Secad de Apartamentos, D162 em 17.08.93 Volmar S. Reg. </p> <p> em 25/ julho /1994. o chefe da secad de Apartamentos. D162. Ofício do parecer final, na presente data. </p>	
Mod/936 - 10.000 Fls.		<p> [Assinatura] </p>	

Figura 20 - Resultado do algoritmo Formiga-Lins da imagem 0047 da BASE1_1. Tempo de processamento 0,337550s. Informação tocada pela borda ruidosa, à esquerda, foi preservada. Tamanho 61KB.

Folha N.º PROCESSO N.º	Folha de Informações e Despachos		GOVERNO DE PERNAMBUCO SECRETARIA DE EDUCAÇÃO, CULTURA E ESPORTES SECE
<p> Informo a situação Municipal Governador Arnaldo José Sales Carvalho, dat. 28.9.60, Exerce o Cargo de Sgo. HAB. FS. VII, através do ato 1902, de 03/04/93, DO 06/04/93, foi posto à disposição da Prefeitura da Cidade do Recife, e através da Portaria SAD 2331, de 07/07/93, DO de 08/07/93, retornou a Secretaria de Educação, a partir de 04/06/93, tendo sido encaminhado para o DERECETE SUL, mediante Ofício nº 510, de 01-07-93, DEAP. </p>		<p> para o órgão de origem de mediante celebração de convênio. Chefe da Secd de Des- tamento, 3162, em 26.07.93. Wilson S. Sá. </p>	
<p> Secad de Apartamentos em 26/07/93 </p>		<p> A DEH/Secad de Apartamentos, para providenciar ofício comunicando a impossibili- dade de atendimento, nos termos do parecer dessa Secd e posterior arquivam- ento. CAB 27.07.93 Maria do Carmo Silva Secretária Adjunta Secretaria de Educação Cultura e Esportes </p>	
<p> A Secretária Adjunta informando que o pleito não preenche os requisitos do Decreto nº 16.402/92, de 29.02.93, alterado pelo Decreto nº 16.604/93, de 20.04.93, considerando que o referido decreto comen- ta permite a cessã de professores, quando requi- sitado para cargo comissionado, função gratificada, sem ônus </p>		<p> 09-8-93 A Secretária Adjunta, su- gerindo arquivar o pro- cesso, considerando que o pleito está sendo provi- denciado através do ofí- cio nº 769/93, de 09.07.93. Chefe da Secd de Aparta- mento, 3162 em 17.08.93 Wilson S. Sá. em 25/ julho /1994. o chefe da secd de Apartamentos. 3162. Ofício do parecer final, na presente data. Arnaldo </p>	
Mod/936 - 10.000 Fls.			

Figura 21 - Resultado do Novo Algoritmo da imagem 0047 da BASE1_1. Tempo de processamento 0,174372s.

Informação tocada pela borda ruidosa, à esquerda, foi preservada.

Menor tempo de processamento Tamanho 61KB.

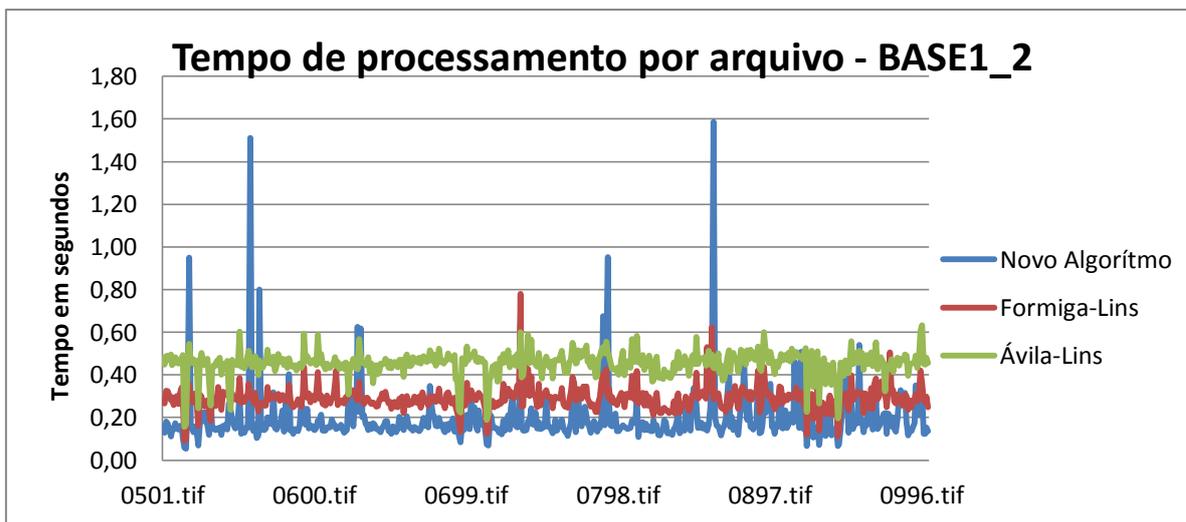


Gráfico 3 - Comparação do tempo de processamento por arquivo - grupo BASE1_2

BASE1_2			
	TEMPO MÉDIO POR ARQUIVO (s)	TEMPO TOTAL (s)	VARIÂNCIA
ÁVILA-LINS	0,454680	227,340193	0,003279
FORMIGA-LINS	0,290111	145,055573	0,003036
NOVO ALGORITMO	0,193503	96,751637	0,016324

Tabela 6 - Comparação do tempo médio de processamento por arquivo e do tempo total de processamento da BASE1_2

BASE1_2	
	REDUÇÃO DE TEMPO EM RELAÇÃO AO ÁVILA-LINS (%)
FORMIGA-LINS	36,19
NOVO ALGORITMO	57,44

Tabela 7 - Redução do tempo de processamento do Formiga-Lins e do Novo Algoritmo em relação ao Ávila-Lins na BASE1_2

BASE1_2		
	TAMANHO TOTAL (KB)	TAMANHO MÉDIO (KB)
ORIGINAL	19.660,80	39,32
ÁVILA-LINS	19.148,80	38,30
FORMIGA-LINS	19.046,40	38,09
NOVO ALGORITMO	18.944,00	37,89

Tabela 8 - Comparação do tamanho final da BASE1_2 de cada algoritmo em relação ao tamanho original em KB

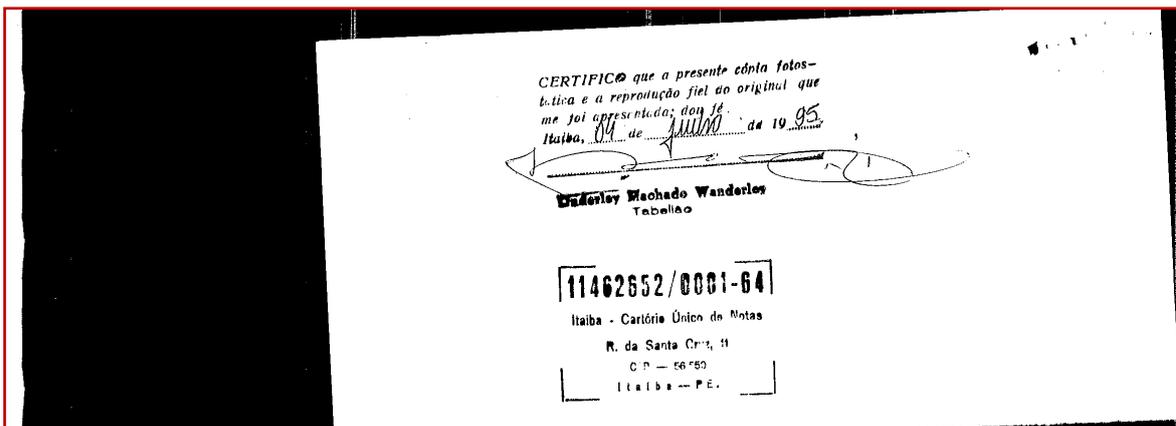


Figura 22 – Imagem 0516 da BASE1_2. Imagem de documento de tamanho menor que o padrão (uma parte da folha padrão de um documento). Apresenta rotação. Tamanho 7,71KB.

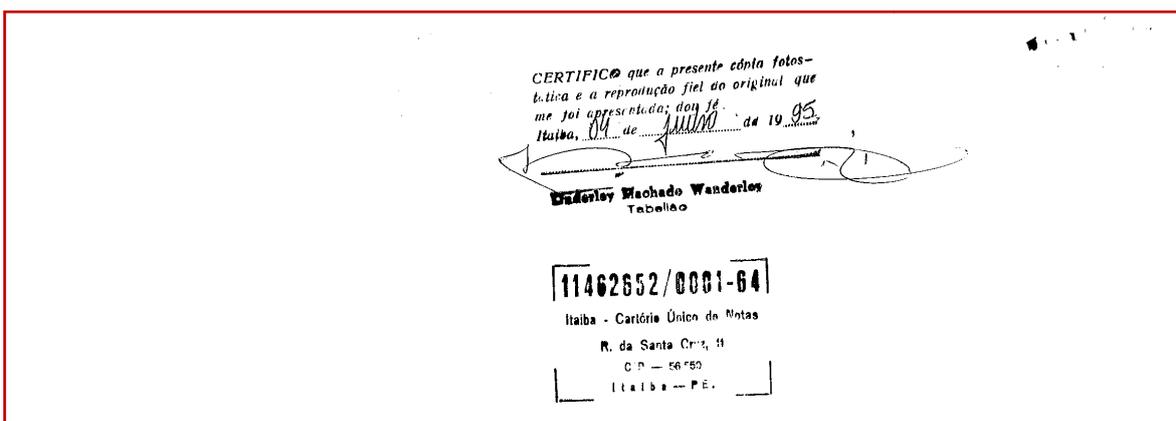


Figura 23 – Resultado do algoritmo Ávila-Lins da imagem 0516 da BASE1_2.
Tempo de processamento 0,166254s. Tamanho 6,58KB.

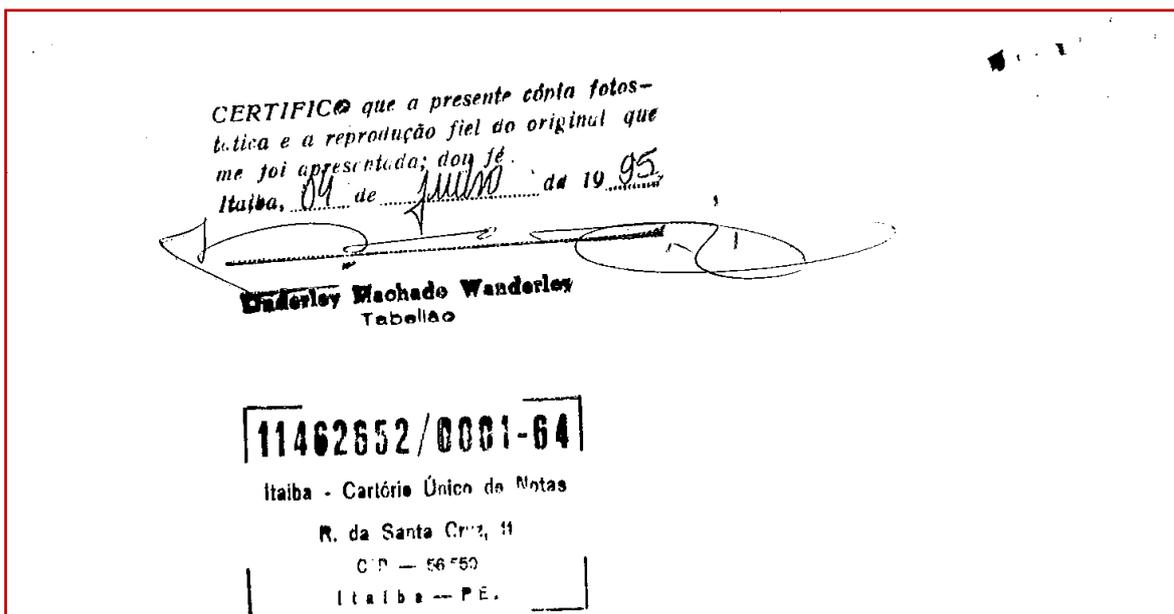


Figura 24 – Resultado do algoritmo Formiga-Lins da imagem 0516 da BASE1_2.

Tempo de processamento 0,095184s. Conseguiu reduzir o tamanho da imagem final realizando o recorte (*crop*) de parte da borda ruidosa antes de sua remoção. Tamanho 6,45KB.

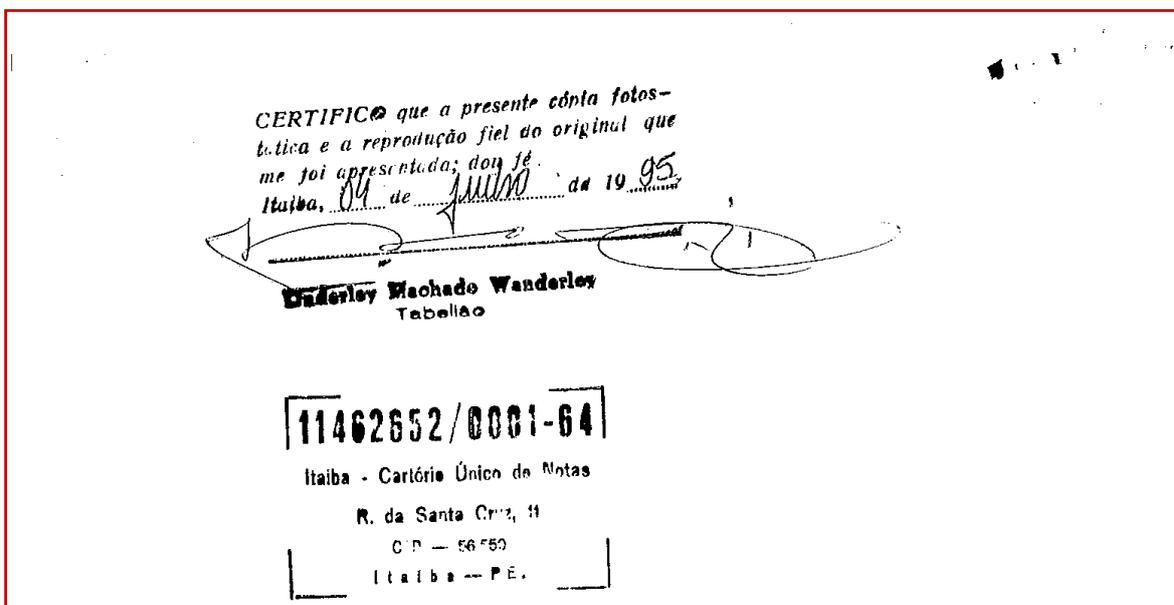


Figura 25 – Resultado do Novo Algoritmo da imagem 0516 da BASE1_2. Tempo de processamento 0,053080s. Conseguiu reduzir o tamanho da imagem final removendo realizando o recorte (*crop*) de parte da borda ruidosa antes de sua remoção. Menor tempo de processamento. Tamanho 6,45KB.

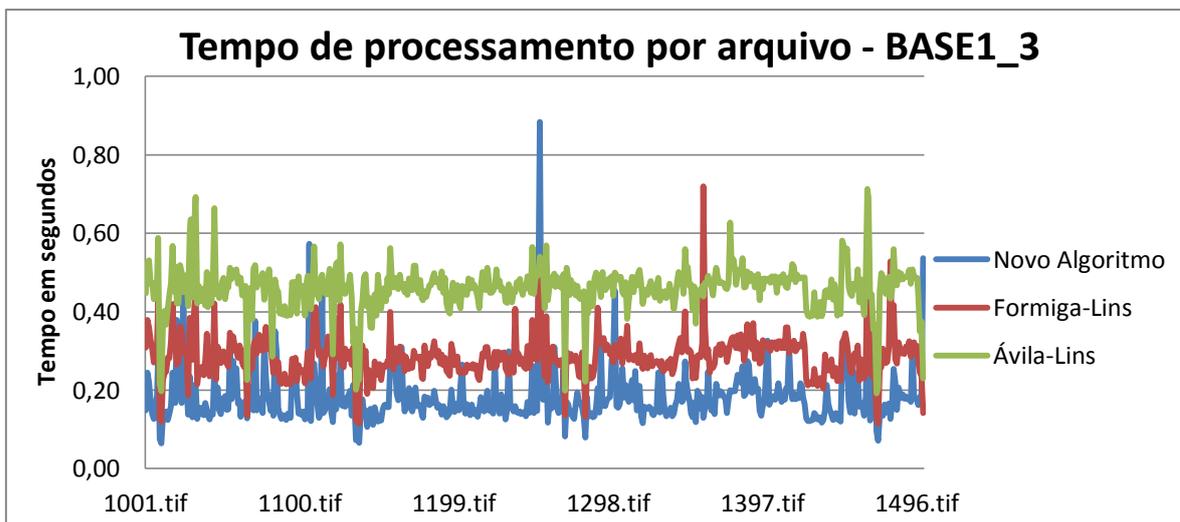


Gráfico 4 - Comparação do tempo de processamento por arquivo - grupo BASE1_3

BASE1_3			
	TEMPO MÉDIO POR ARQUIVO (s)	TEMPO TOTAL (s)	VARIÂNCIA
ÁVILA-LINS	0,458965	229,482393	0,003548
FORMIGA-LINS	0,285169	142,584454	0,003097
NOVO ALGORITMO	0,179953	89,976560	0,004571

Tabela 9 - Comparação do tempo médio de processamento por arquivo e do tempo total de processamento da BASE1_3

BASE1_3	
	REDUÇÃO DE TEMPO EM RELAÇÃO AO ÁVILA-LINS (%)
FORMIGA-LINS	37,87
NOVO ALGORITMO	60,79

Tabela 10 - Redução do tempo de processamento do Formiga-Lins e do Novo Algoritmo em relação ao Ávila-Lins na BASE1_3

BASE1_3		
	TAMANHO TOTAL (KB)	TAMANHO MÉDIO (KB)
ORIGINAL	20.787,20	41,57
ÁVILA-LINS	20.684,80	41,37
FORMIGA-LINS	20.582,40	41,16
NOVO ALGORITMO	20.582,40	41,16

Tabela 11 - Comparação do tamanho final da BASE1_3 de cada algoritmo em relação ao tamanho original em KB

Gov. do Estado de Pernambuco
 MUNICÍPIO: 8.2000 00.01734.5 - 09/08/2000
 HOME: ANTONIA DE SOUZA CAVALCANTI
 ASSUNTO: EQUIPARAÇÃO SALARIAL
 ORÇAM. DE GAR. DE DEC. DECO
 FOLG. INFORMACOES DAS 7.30 AS 13.00
 2002 TELEFONE: 4241122 RAMAIS: 2061
 AÇÃO RESERVADA PARA O DESPACHO FINAL

NOME: Antônia de Souza Cavalcanti MATRÍCULA: 76.721-2
 ENDERECO: Rua São Sebastião N° 221 CIDADE: Caruaru EST.: PE TELEFONE: 622-0324
 CEP: 007.101-134.04 SEXO: Fem. DATA DO NASCIMENTO: 22/06/31 ESTADO CIVIL: Solteira
 IDENTIDADE: 383.530 TITULO DE ELEITOR: 99697008/2002 SEÇÃO: 73 PASEP: 100102212-79
 CARGO/FUNÇÃO: Professor DEPTO. CAMPUS: HAB-3 SIMBOLO/NIVEL/CATEGORIA OU PADRAO: ES-IX NU-8
 UNIDADE DE TRABALHO: Unidade SEC. de Educação DIRETORIA: DDEE VINCULO: CLT EFET

REQUER AO EXMO. SR. SECRETÁRIO DE EDUCAÇÃO E CULTURA

BLOCO 2
 1 APROVAÇÃO DE FALTA
 2 AFASTAMENTO P/SUBMETETER-SE A PROVAS
 3 AFASTAMENTO PELO ART. 177 DA LEI 6123/68
 4 ANOTAÇÃO DE DIPLOMA E/OU CURSO
 5 CERTIDÃO P/INS ESPECÍFICOS
 6 GRATIFICAÇÃO A PROFESSORES DE EXCEPCIONAIS
 7 LICENÇA A GESTANTE
 8 LICENÇA POR MOTIVO DE DOENÇA EM PESSOA DA FAM.
 9 LICENÇA PRÊMIO (GOZO)
 10 REMOÇÃO
 11 READAPTAÇÃO DE CARGO
 12 RETIFICAÇÃO DE NOME
 13 SALÁRIO FAMÍLIA
 OUTROS

REQUER AO EXMO. SR. SECRETÁRIO DE ADMINISTRAÇÃO

BLOCO 3
 14 AFASTAMENTO P/CURSO, ESTÁGIO OU CONGRESSO
 15 ANOTAÇÃO DE TEMPO DE SERVIÇO
 16 APOSENTADORIA
 17 EXONEIAÇÃO (RECONHECER A FIRMA)
 18 LICENÇA PARA ACOMPANHAR O MARIDO
 19 LICENÇA P/TRATAMENTO DE SAÚDE
 20 LICENÇA PARA TRATO DE INTERESSE PARTICULAR
 21 LICENÇA PRÊMIO (CONCESSÃO)
 22 REAJUSTE SALARIAL
 23 RESCISÃO DE CONTRATO (RECONHECER A FIRMA)
 24 SUSPENSÃO DE CONTRATO
 OUTROS

BLOCO 4 ANEXO DOCUMENTOS NECESSÁRIOS A SOLUÇÃO DO ASSUNTO E PARA QUE PRESTA AS SEGUINTES
 INFORMAÇÕES COMPLEMENTARES:
Solicito a equiparação salarial na

ATENÇÃO
 BLOCO 5
 EM 09/08/2000
Antônia de Souza Cavalcanti
 ASSINATURA DO REQUERENTE
 MOD. 944

CIENTE ENCAMINHE-SE
 EM / /
 ASSINATURA DO RESPONSÁVEL

VISTO
 EM / /
 DIRETOR

Figura 27 – Resultado do algoritmo Ávila-Lins da imagem 1253 da BASE1_3. Tempo de processamento 0,497588s. Ruído no canto superior esquerdo permanece após o processo. Tamanho 80,1KB.

Gov. do Estado de Pernambuco
 NÚMERO: 0.2000 00.01734.3 - 09/08/2000
 HONRE. ANTONIA DE SOUZA CAVALCANTI
 AÇÃO RESERVADA PARA O DESPACHO FINAL

ASSUNTO: EQUIPARAÇÃO SALARIAL
 ORÇAM. DE GAR. DE UNIC. DECO
 FOLG. INFORMACOES DAS 7.30 AS 13.00
 2002 TELEFONE: 4241122 RAMAIS: 2061

NOME: Antonia de Souza Cavalcanti MATRÍCULA: 76.721-2
 ENDERECO: Rua São Sebastião Nº 221 CIDADE: Caruaru EST.: PE TELEFONE: 622-0324
 DATA DO NASCIMENTO: 22/06/31 ESTADO CIVIL: Solteira
 SEXO: Fem. PASEP: 100102212-79
 CH: 007.101.134.04 TÍTULO DE ELEITOR: 20 SEÇÃO: 73
 IDENTIDADE: 383.530 Nº: 99697008/20 SIMBOLO/NIVEL/CATEGORIA OU PADRAO: HAB-3, ES-IX NU-8
 CARGO/FUNÇÃO: Professor DEPTO./CAMPE: DERF - Mata Norte DIRETORIA: DDEE VINCULO: CLT EFET
 UNIDADE DE TRABALHO: Unidade SEC. de Educação

REQUER AO EXMO. SR. SECRETÁRIO DE EDUCAÇÃO E CULTURA

BLOCO 2
 1 APROVAÇÃO DE FALTA
 2 AFASTAMENTO P/SUBMETETER-SE A PROVAS
 3 AFASTAMENTO PELO ART. 177 DA LEI 6123/68
 4 ANOTAÇÃO DE DIPLOMA E/OU CURSO
 5 CERTIDÃO P/INS ESPECIFICOS
 6 GRATIFICAÇÃO A PROFESSORES DE EXCEPCIONAIS
 7 LICENÇA A GESTANTE
 8 LICENÇA POR MOTIVO DE DOENÇA EM PESSOA DA FAM.
 9 LICENÇA PRÊMIO (GOZO)
 10 REMOÇÃO
 11 READAPTAÇÃO DE CARGO
 12 RETIFICAÇÃO DE NOME
 13 SALÁRIO FAMÍLIA
 OUTROS

REQUER AO EXMO. SR. SECRETÁRIO DE ADMINISTRAÇÃO

BLOCO 3
 14 AFASTAMENTO P/CURSO, ESTÁGIO OU CONGRESSO
 15 ANOTAÇÃO DE TEMPO DE SERVIÇO
 16 APOSENTADORIA
 17 EXONEIAÇÃO (RECONHECER A FIRMA)
 18 LICENÇA PARA ACOMPANHAR O MARIDO
 19 LICENÇA P/TRATAMENTO DE SAÚDE
 20 LICENÇA PARA TRATO DE INTERESSE PARTICULAR
 21 LICENÇA PRÊMIO (CONCESSÃO)
 22 AJUSTE SALARIAL
 23 RESCISÃO DE CONTRATO (RECONHECER A FIRMA)
 24 SUSPENSÃO DE CONTRATO
 OUTROS

BLOCO 4 ANEXO DOCUMENTOS NECESSÁRIOS A SOLUÇÃO DO ASSUNTO E PARA QUE PRESTA AS SEGUINTE
 INFORMAÇÕES COMPLEMENTARES:
Solicito a equiparação salarial na

ATENÇÃO
 BLOCO 5
 EM 09/08/2000
Antonia de Souza Cav.
 ASSINATURA DO REQUERENTE
 MOD. 944 calcanti
 CLIENTE ENCAMINHE-SE EM _____
 ASSINATURA DO RESPONSÁVEL
 VISTO EM _____
 DIRETOR

Figura 28 – Resultado do algoritmo Formiga-Lins da imagem 1253 da BASE1_3. Tempo de processamento 0,440771s. Ruído no canto superior esquerdo permanece após o processo. Tamanho 80,1KB.

Governo do Estado de Pernambuco
 MUNICÍPIO DE CAROLINA DE SOUZA CAVALCANTI
 Nº 09/08/2000
 AÇÃO RESERVADO PARA O DESPACHO FINAL

ASSUNTO: EQUIPARAÇÃO SALARIAL
 ORÇÃO DE GAB. Nº 1000
 FOLHA Nº 130
 TELEFONE: 4241122 RUAIS: 2061

NOME: Antonia de Souza Cavalcanti
 ENDEREÇO: Rua São Sebastião Nº 921
 CIDADE: Carolina
 EST. PE
 MATRÍCULA: 46.421-2
 TELEFONE: 622-0324
 DATA DO NASCIMENTO: 22/06/31
 ESTADO CIVIL: Soltura
 SEXO: Fem.
 PASEP: 100102212-79
 TÍTULO DE ELEITOR: 99697008/2008
 SEÇÃO: 73
 SIMBOL/NÍVEL/CATEGORIA OU PADRÃO: HAB-3, ES-IX, NU-8
 IDENTIDADE: 383.530
 CARGO-FUNÇÃO: Professor
 DEPTO. CAMPUS: DERE - Mata Norte
 DIRETORIA: DDEE
 VINCULO: CLT EFET

UNIDADE DE TRABALHO: Unidade SEC. de Educação

REQUER AO EXMO. SR. SECRETÁRIO DE EDUCAÇÃO E CULTURA:

BLOCO 2

<input type="checkbox"/> 1 AFONO DE FALTA	<input type="checkbox"/> 8 LICENÇA POR MOTIVO DE DOENÇA EM PESSOA DA FAM.
<input type="checkbox"/> 2 AFASTAMENTO P/ SUBMETETER-SE A PROVAS	<input type="checkbox"/> 9 LICENÇA PRÊMIO (GOZO)
<input type="checkbox"/> 3 AFASTAMENTO PELO ART. 177 DA LEI 0123/68	<input type="checkbox"/> 10 REMOÇÃO
<input type="checkbox"/> 4 ANOTAÇÃO DE DIPLOMA E/OU CURSO	<input type="checkbox"/> 11 READAPTAÇÃO DE CARGO
<input type="checkbox"/> 5 CERTIFICAÇÃO P/ FINS ESPECÍFICOS	<input type="checkbox"/> 12 RETIFICAÇÃO DE NOME
<input type="checkbox"/> 6 GRATIFICAÇÃO A PROFESSORES DE EXCEPCIONAIS	<input type="checkbox"/> 13 SALÁRIO FAMILIA
<input type="checkbox"/> 7 LICENÇA A GESTANTE	<input type="checkbox"/> OUTROS

REQUER AO EXMO. SR. SECRETARIO DE ADMINISTRAÇÃO:

BLOCO 3

<input type="checkbox"/> 14 AFASTAMENTO P/ CURSO, ESTAGIO OU CONGRESSO	<input type="checkbox"/> 20 LICENÇA PARA TRATO DE INTERESSE PARTICULAR
<input type="checkbox"/> 15 ANOTAÇÃO DE TEMPO DE SERVIÇO	<input type="checkbox"/> 21 LICENÇA PRÊMIO (CONCESSÃO)
<input type="checkbox"/> 16 APOSENTADORIA	<input type="checkbox"/> 22 REAJUSTE SALARIAL
<input type="checkbox"/> 17 EXONERAÇÃO (RECONHECER A FIRMA)	<input type="checkbox"/> 23 RESCISÃO DE CONTRATO (RECONHECER A FIRMA)
<input type="checkbox"/> 18 LICENÇA PARA ACOMPANHAR O MARIDO	<input type="checkbox"/> 24 SUSPENSÃO DE CONTRATO
<input type="checkbox"/> 19 LICENÇA P/ TRATAMENTO DE SAÚDE	<input checked="" type="checkbox"/> OUTROS

BLOCO 4 ANEXO DOCUMENTOS NECESSÁRIOS A SOLUÇÃO DO ASSUNTO E PARA QUE PRESTA AS SEGUINTE
 INFORMAÇÕES COMPLEMENTARES:

Solucito a equiparação salarial ma

BLOCO 5

EM 09/08/2000
Antonia de Souza Cavalcanti
 ASSINATURA DO REQUERENTE

MOD. 944

ATENÇÃO

ASSINANTE ENCAMINHE-SE EM _____ VISTO EM _____ DIRETOR

Figura 29 – Resultado do Novo Algoritmo da imagem 1253 da BASE1_3. Tempo de processamento 0,399581s.

Ruído no canto superior esquerdo foi removido após o processo.

Menor tempo de processamento. Tamanho 79,8KB.

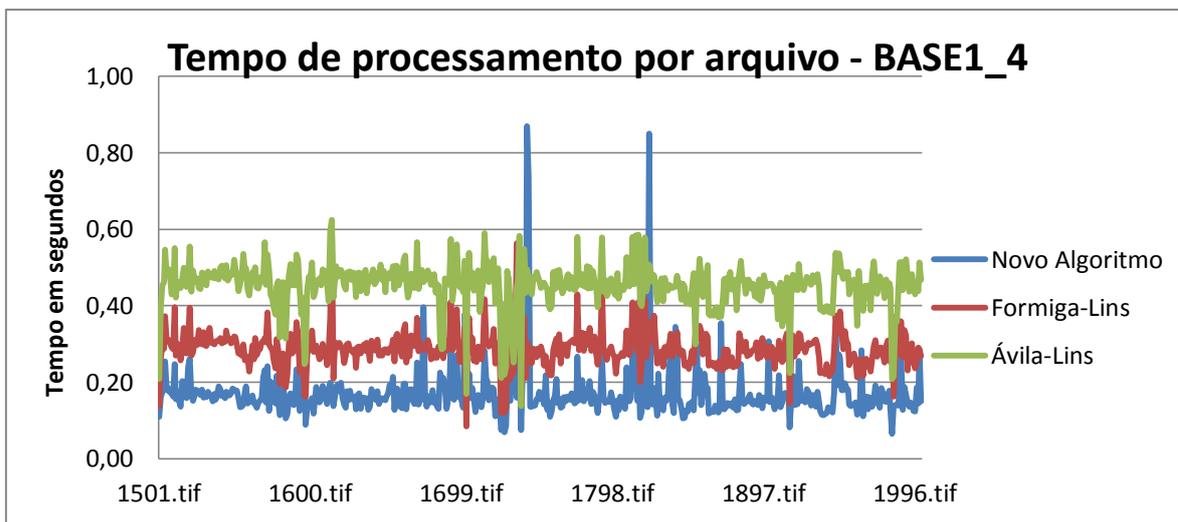


Gráfico 5 - Comparação do tempo de processamento por arquivo - grupo BASE1_4

BASE1_4			
	TEMPO MÉDIO POR ARQUIVO (s)	TEMPO TOTAL (s)	VARIÂNCIA
ÁVILA-LINS	0,453479	226,739395	0,003279
FORMIGA-LINS	0,286328	143,164045	0,003159
NOVO ALGORITMO	0,169673	84,836673	0,005163

Tabela 12 - Comparação do tempo médio de processamento por arquivo e do tempo total de processamento da BASE1_4

BASE1_4	
	REDUÇÃO DE TEMPO EM RELAÇÃO AO ÁVILA-LINS (%)
FORMIGA-LINS	36,86
NOVO ALGORITMO	62,58

Tabela 13 – Redução do tempo de processamento do Formiga-Lins e do Novo Algoritmo em relação ao Ávila-Lins na BASE1_4

BASE1_4		
	TAMANHO TOTAL (KB)	TAMANHO MÉDIO (KB)
ORIGINAL	19.660,80	39,32
ÁVILA-LINS	19.558,40	39,12
FORMIGA-LINS	19.968,00	39,94
NOVO ALGORITMO	19.558,40	39,12

Tabela 14 - Comparação do tamanho final da BASE1_4 de cada algoritmo em relação ao tamanho original em KB

Folha de Pagamento dos Professores e Funcionários

do Ginásio e Escola Normal Amália Cavalcanti
 CONTINUAÇÃO PEDRA - PERNAMBUCO Maio 1974

NOME	CARGO	Audis	Imparcialidade	PERÍODO	Imparcialidade Recibível	ASSINATURA DO RECEBEDOR
Valda Vilela de Souza	Professor	24	120,00	Maio	120,00	Valda Vilela de Souza
✓ Ana Maria Diniz Cavalcanti	"	18	72,00	"	72,00	ANA MARIA DINIZ CAVALCANTI
✓ Odete Simão	"	70	280,00	"	280,00	Odete Simão
Angela Maria Siqueira Barbosa	"	73	292,00	"	292,00	Angela Maria Siqueira Barbosa
***** T. O. T. A. I. *****	"	-	2.748,00	"	2.748,00	Edinete Diniz Paes

PAGUE-SE
P A G O
 Visto Presidente
 Visto Recebedor
 Visto Tesoureiro
 Visto Auxiliar

EM PERNAMBUCO
 TRIBUNAL DE CONTAS DO PERNAMBUCO
 SECRETARIA DE EDUCAÇÃO DO PERNAMBUCO
 AAF 96405

Figura 31 – Resultado do algoritmo Ávila-Lins da imagem 1944 da BASE1_4. Tempo de processamento 0,537512s. Tamanho 266KB.

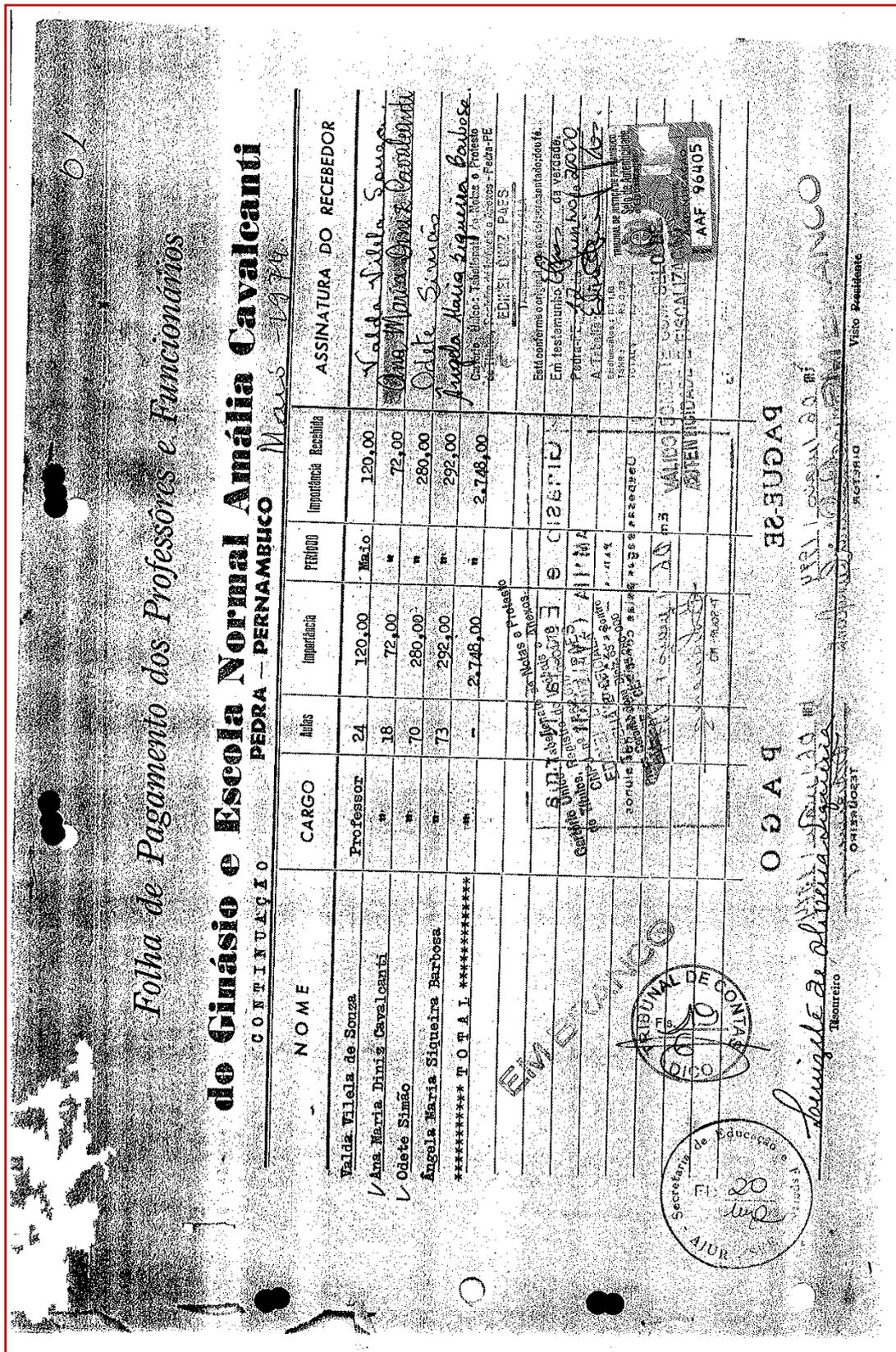


Figura 33 – Resultado do Novo Algoritmo da imagem 1944 da BASE1_4. Tempo de processamento 0,347512s. Ruído no lado esquerdo não comprometeu a detecção da fronteira à esquerda da imagem do documento. Pequena parte do ruído, no canto inferior esquerdo, foi removido. Tamanho 260KB.

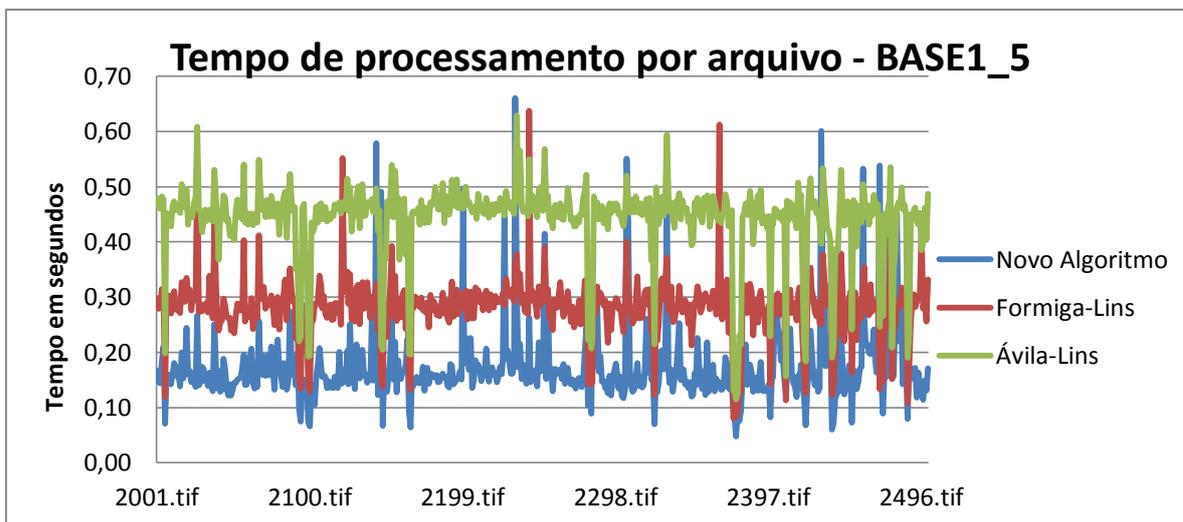


Gráfico 6 - Comparação do tempo de processamento por arquivo - grupo BASE1_5

BASE1_5			
	TEMPO MÉDIO POR ARQUIVO (s)	TEMPO TOTAL (s)	VARIÂNCIA
ÁVILA-LINS	0,441693	220,846348	0,005026
FORMIGA-LINS	0,280693	140,346449	0,003159
NOVO ALGORITMO	0,172644	86,321758	0,005163

Tabela 15 - Comparação do tempo médio de processamento por arquivo e do tempo total de processamento da BASE1_5

BASE1_5	
	REDUÇÃO DE TEMPO EM RELAÇÃO AO ÁVILA-LINS (%)
FORMIGA-LINS	36,45
NOVO ALGORITMO	60,91

Tabela 16 - Redução do tempo de processamento do Formiga-Lins e do Novo Algoritmo em relação ao Ávila-Lins na BASE1_5

BASE1_5		
	TAMANHO TOTAL (KB)	TAMANHO MÉDIO (KB)
ORIGINAL	19.558,40	39,12
ÁVILA-LINS	19.148,80	38,30
FORMIGA-LINS	19.046,40	38,09
NOVO ALGORITMO	19.353,60	38,71

Tabela 17 - Comparação do tamanho final da BASE1_5 de cada algoritmo em relação ao tamanho original em KB

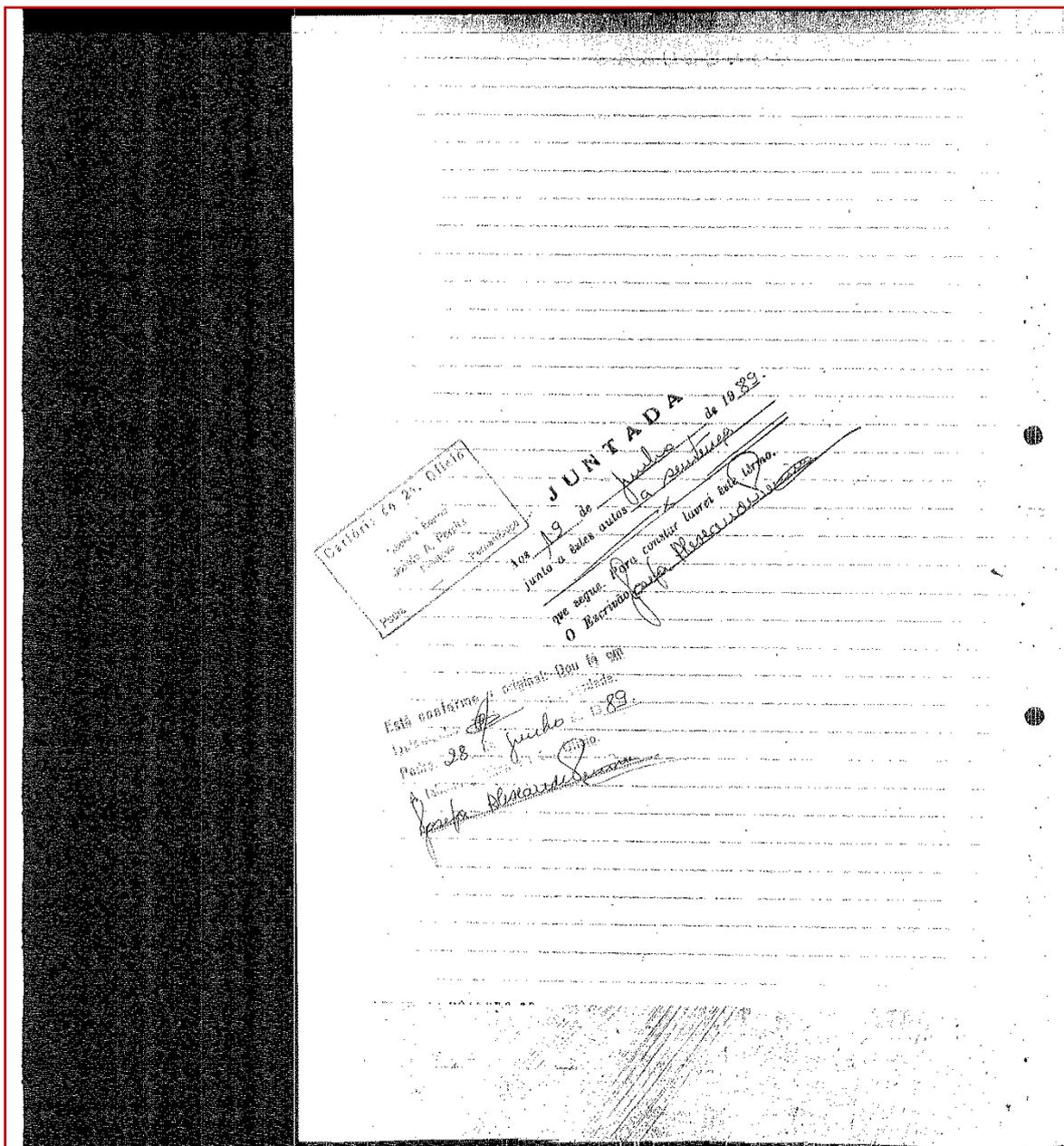


Figura 34 – Imagem 2236 da BASE1_5. Imagem de documento apresenta borda muito ruidosa à esquerda.

Tamanho 124KB.

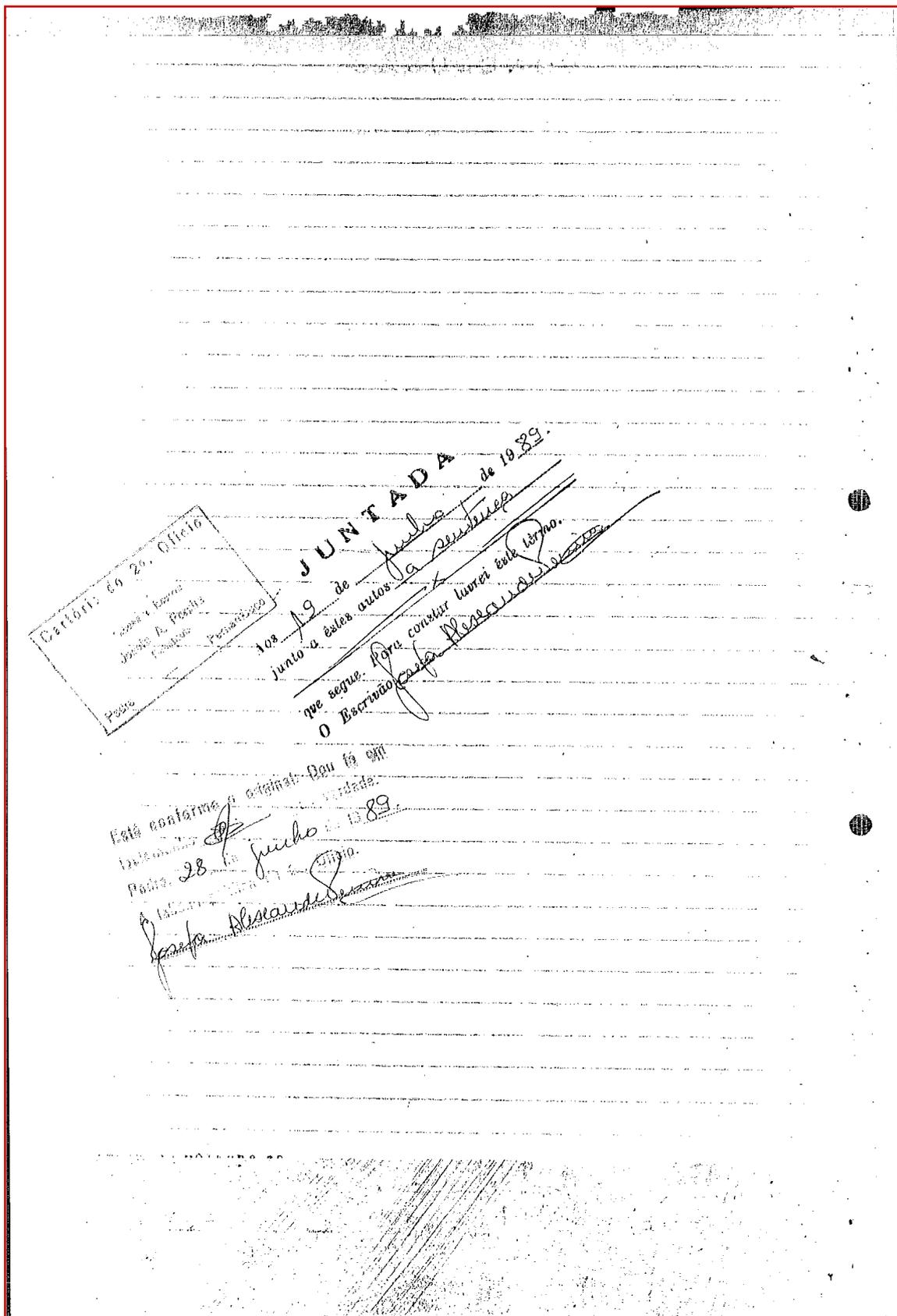


Figura 36 – Resultado do algoritmo Formiga-Lins da imagem 2236 da BASE1_5. Tempo de processamento 0,342506s. A borda ruidosa foi completamente removida. Tamanho 39,6KB.

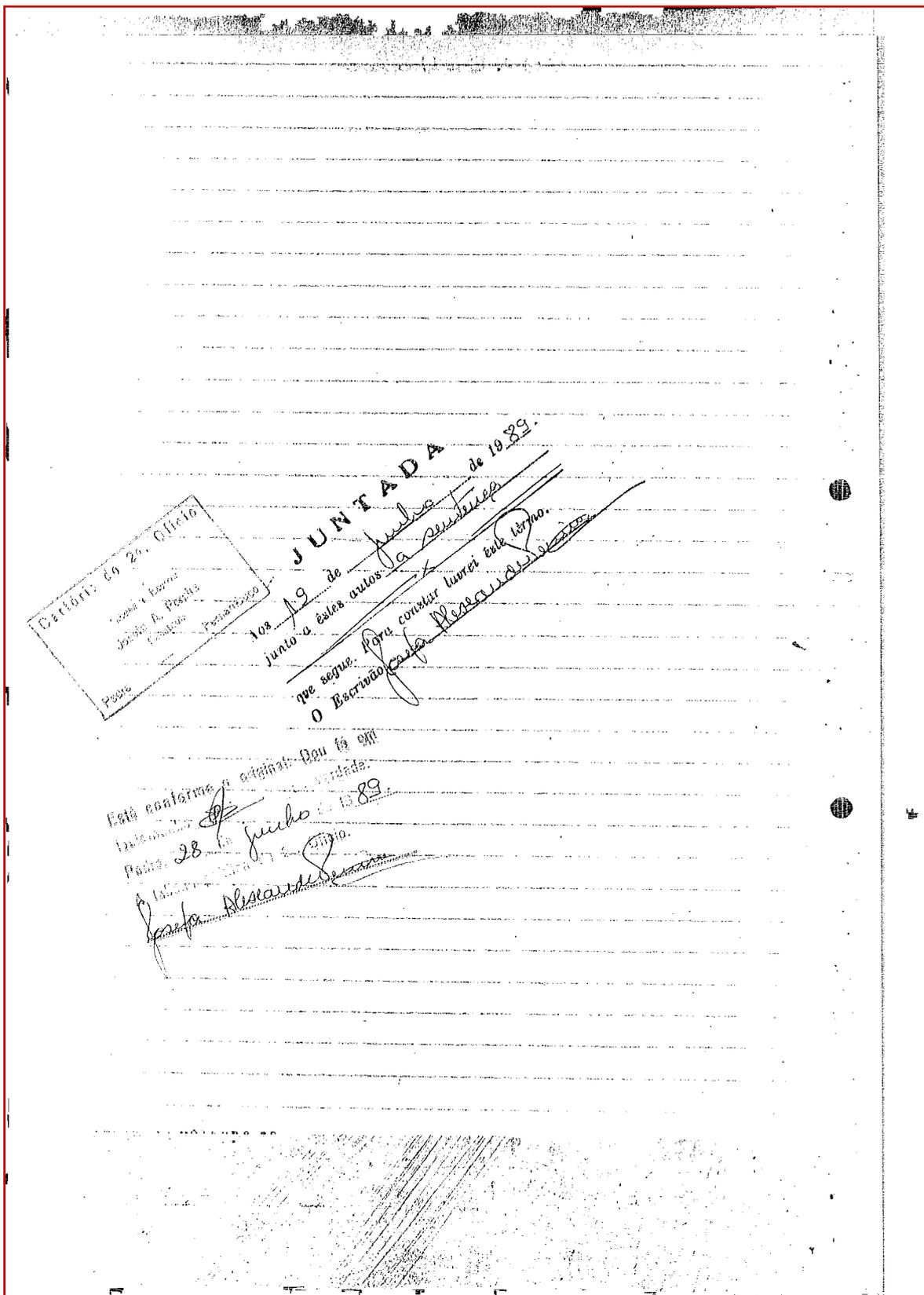


Figura 37 – Resultado do Novo Algoritmo da imagem 2236 da BASE1_5. Tempo de processamento 0,215560. A borda ruidosa foi completamente removida. Menor tempo de processamento. Tamanho 40,6KB.

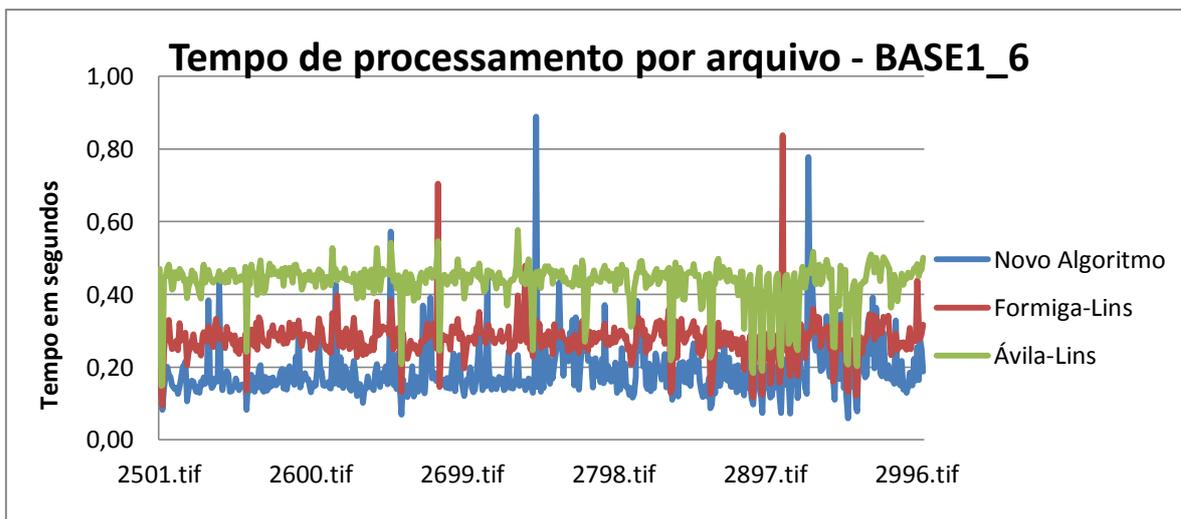


Gráfico 7 - Comparação do tempo de processamento por arquivo - grupo BASE1_6

BASE1_6			
	TEMPO MÉDIO POR ARQUIVO (s)	TEMPO TOTAL (s)	VARIÂNCIA
ÁVILA-LINS	0,433281	216,640533	0,003619
FORMIGA-LINS	0,276589	138,294338	0,002931
NOVO ALGORITMO	0,185411	92,705741	0,005596

Tabela 18 - Comparação do tempo médio de processamento por arquivo e do tempo total de processamento da BASE1_6

BASE1_6	
	REDUÇÃO DE TEMPO EM RELAÇÃO AO ÁVILA-LINS (%)
FORMIGA-LINS	36,16
NOVO ALGORITMO	57,21

Tabela 19 - Redução do tempo de processamento do Formiga-Lins e do Novo Algoritmo em relação ao Ávila-Lins na BASE1_6

BASE1_6		
	TAMANHO TOTAL (KB)	TAMANHO MÉDIO (KB)
ORIGINAL	16.384,00	32,77
ÁVILA-LINS	15.872,00	31,74
FORMIGA-LINS	15.769,60	31,54
NOVO ALGORITMO	15.462,40	30,92

Tabela 20 - Comparação do tamanho final da BASE1_6 de cada algoritmo em relação ao tamanho original em KB

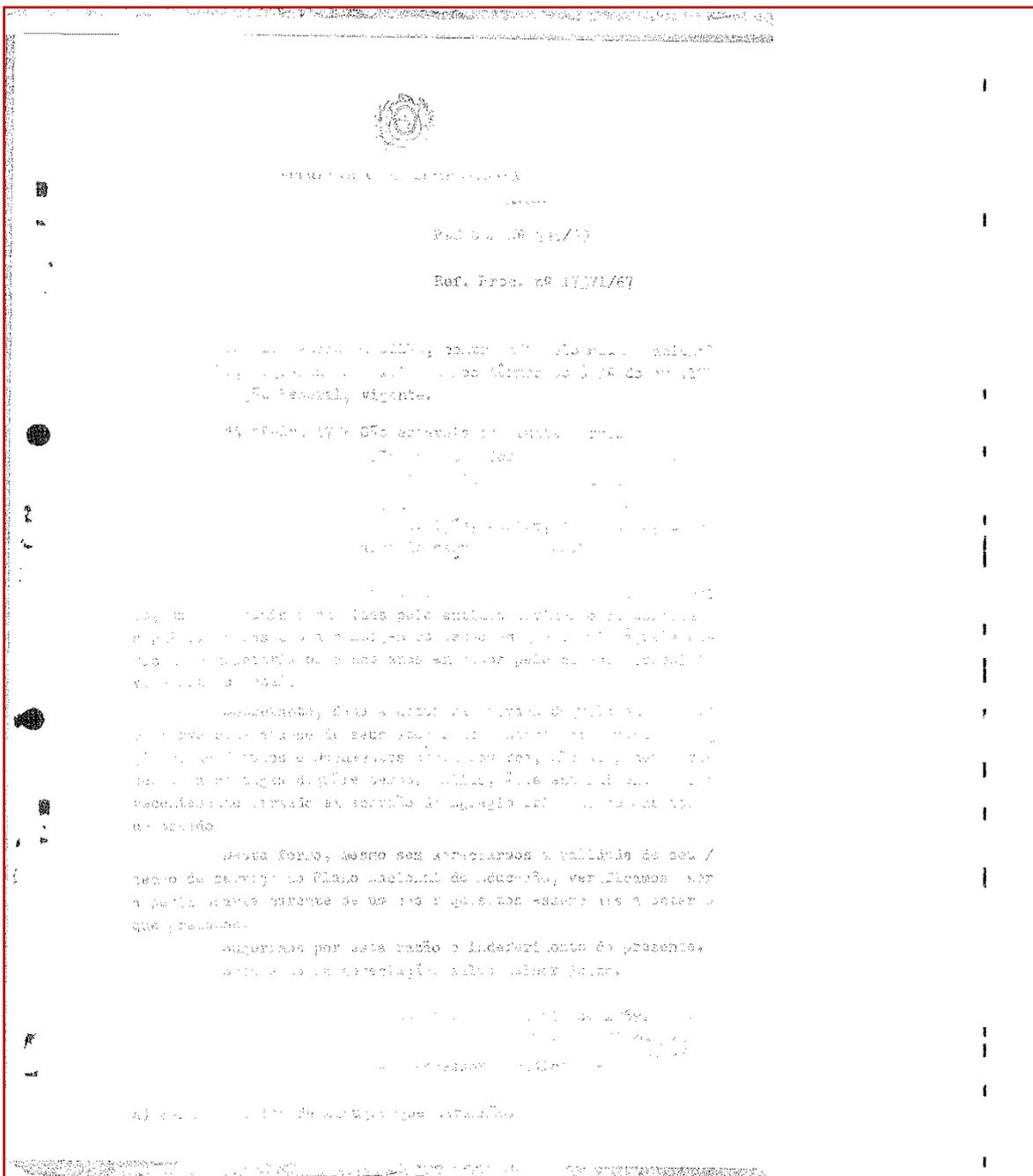


Figura 39 – Resultado do Algoritmo Ávila-Lins da imagem 2683 da BASE1_6. Tempo de processamento 0,546055s. Remoção completa da borda ruidosa. Tamanho 22,4KB.

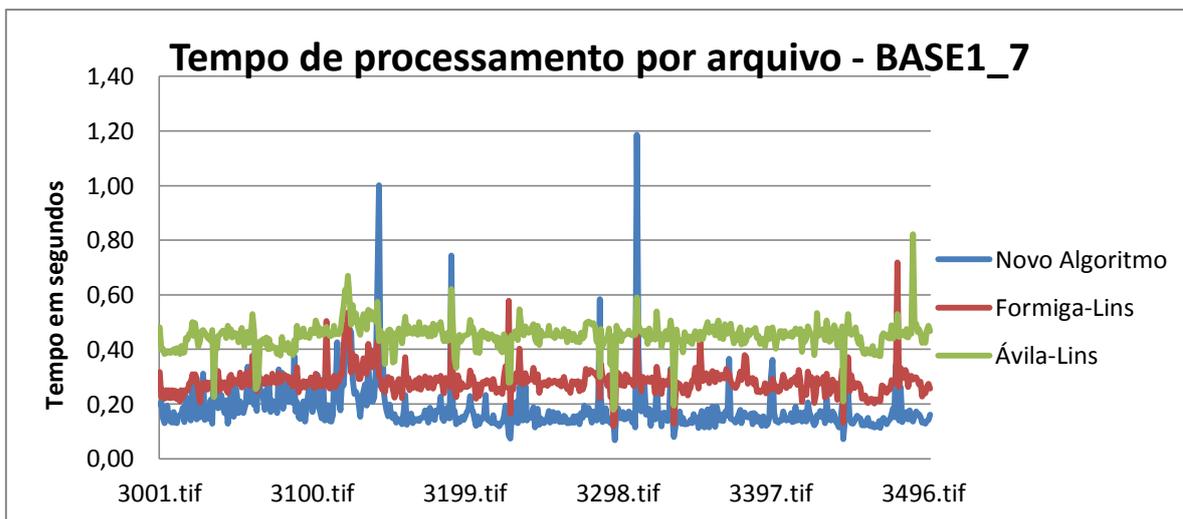


Gráfico 8 - Comparação do tempo de processamento por arquivo - grupo BASE1_7

BASE1_7			
	TEMPO MÉDIO POR ARQUIVO (s)	TEMPO TOTAL (s)	VARIÂNCIA
ÁVILA-LINS	0,449327	224,663266	0,002758
FORMIGA-LINS	0,282598	141,299101	0,002487
NOVO ALGORITMO	0,181746	90,873081	0,008315

Tabela 21 - Comparação do tempo médio de processamento por arquivo e do tempo total de processamento da BASE1_7

BASE1_7	
	REDUÇÃO DE TEMPO EM RELAÇÃO AO ÁVILA-LINS (%)
FORMIGA-LINS	37,11
NOVO ALGORITMO	59,55

Tabela 22 - Redução do tempo de processamento do Formiga-Lins e do Novo Algoritmo em relação ao Ávila-Lins na BASE1_7

BASE1_7		
	TAMANHO TOTAL (KB)	TAMANHO MÉDIO (KB)
ORIGINAL	19.353,60	38,71
ÁVILA-LINS	18.739,20	37,48
FORMIGA-LINS	18.636,80	37,27
NOVO ALGORITMO	18.944,00	37,89

Tabela 23 - Comparação do tamanho final da BASE1_7 de cada algoritmo em relação ao tamanho original em KB

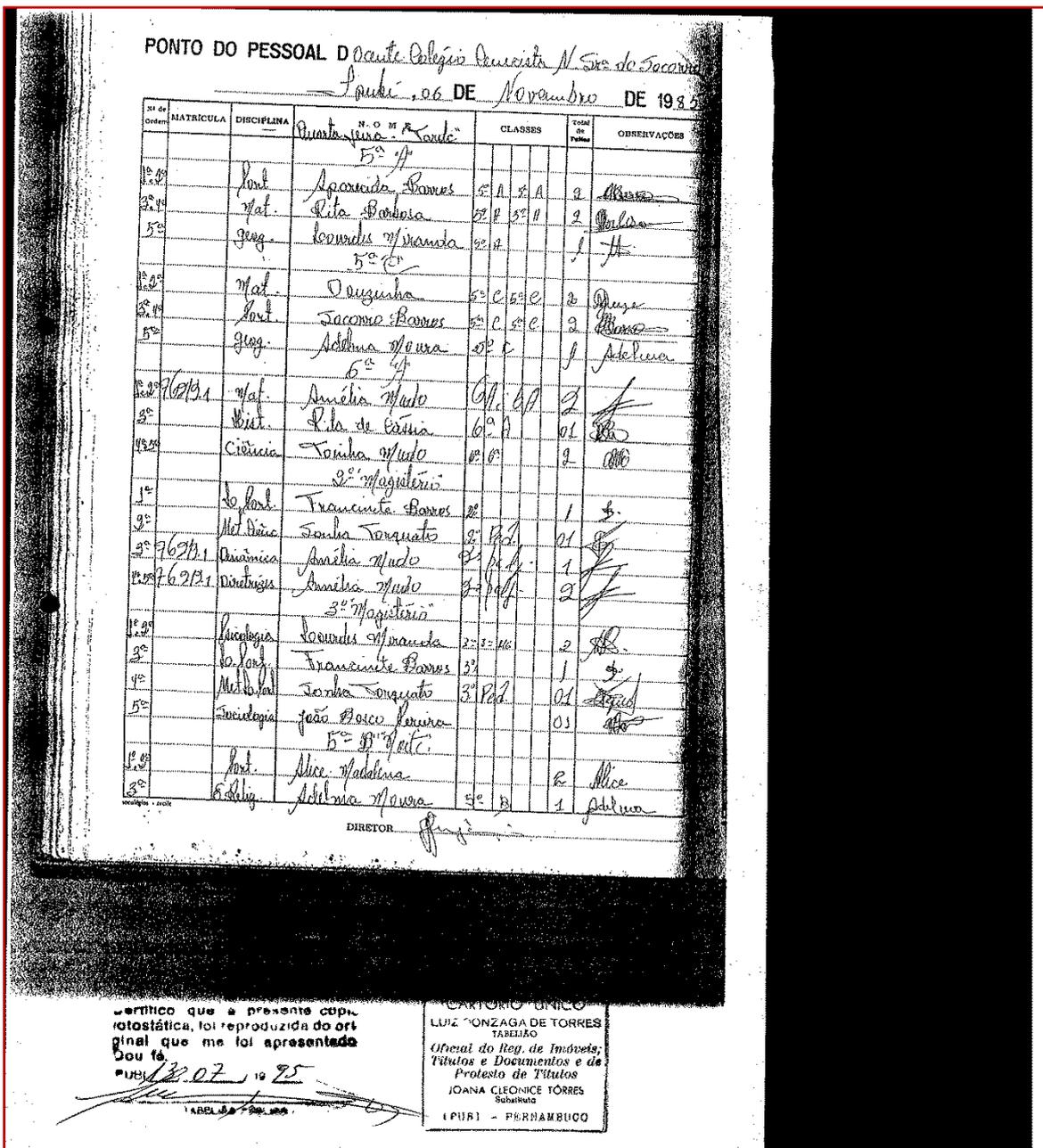


Figura 42 – Imagem 3125 da BASE1_7. Imagem com borda “preta” à direita e ruído no interior da região de interesse, circundando parte da informação do documento. Tamanho 107KB.

PONTO DO PESSOAL DO *Colégio Democrata N. Sra. do Socorro*
Ipuri, 06 DE Novembro DE 1985

Nº de Ordem	NATRICULA	DISCIPLINA	N.º M.º	CLASSES	Total de Faltas	OBSERVAÇÕES
			<i>Quarta-feira - "aula"</i>			
			<i>5º A</i>			
<i>1º 19</i>		<i>Port.</i>	<i>Aparecida Barros</i>	<i>5º A 5º A</i>	<i>2</i>	<i>Aparecida</i>
<i>3º 19</i>		<i>Mat.</i>	<i>Rita Barbosa</i>	<i>5º H 5º H</i>	<i>2</i>	<i>Rita</i>
<i>5º</i>		<i>Geog.</i>	<i>Leonilda M. Miranda</i>	<i>5º H</i>	<i>1</i>	<i>Lea</i>
			<i>5º C</i>			
<i>15 2º</i>		<i>Mat.</i>	<i>Deuzinha</i>	<i>5º C 5º C</i>	<i>2</i>	<i>Deuzinha</i>
<i>3º 19</i>		<i>Port.</i>	<i>Socorro Barros</i>	<i>5º C 5º C</i>	<i>2</i>	<i>Socorro</i>
<i>5º</i>		<i>Geog.</i>	<i>Adelma Moura</i>	<i>5º C</i>	<i>1</i>	<i>Adelma</i>
			<i>5º A</i>			
<i>1º 2º 19.1</i>		<i>Mat.</i>	<i>Amélia Mudo</i>	<i>6º A</i>	<i>2</i>	<i>Amélia</i>
<i>3º</i>		<i>Hist.</i>	<i>Rita de Cassia</i>	<i>6º A</i>	<i>01</i>	<i>Rita</i>
<i>19 5º</i>		<i>Ciências</i>	<i>Toniha Mudo</i>	<i>6º B</i>	<i>2</i>	<i>Toniha</i>
			<i>3º Magistério</i>			
<i>1º</i>		<i>Port.</i>	<i>Francinete Barros</i>	<i>3º</i>	<i>1</i>	<i>Francinete</i>
<i>3º</i>		<i>Met. Did.</i>	<i>Sonha Torquato</i>	<i>3º Ped.</i>	<i>01</i>	<i>Sonha</i>
<i>3º 7.6.2/1</i>		<i>Didática</i>	<i>Amélia Mudo</i>	<i>3º Ped.</i>	<i>1</i>	<i>Amélia</i>
<i>1º 5º 7.6.2/1</i>		<i>Didática</i>	<i>Amélia Mudo</i>	<i>3º Ped.</i>	<i>2</i>	<i>Amélia</i>
			<i>3º Magistério</i>			
<i>1º 2º</i>		<i>Psicologia</i>	<i>Leonilda M. Miranda</i>	<i>3º 5º M.</i>	<i>2</i>	<i>Lea</i>
<i>3º</i>		<i>Port.</i>	<i>Francinete Barros</i>	<i>3º</i>	<i>1</i>	<i>Francinete</i>
<i>1º</i>		<i>Met. Did.</i>	<i>Sonha Torquato</i>	<i>3º Ped.</i>	<i>01</i>	<i>Sonha</i>
<i>5º</i>		<i>Sociologia</i>	<i>João Bosco Louira</i>		<i>01</i>	<i>João</i>
			<i>5º B "Mat."</i>			
<i>1º 2º</i>		<i>Port.</i>	<i>Alice Madalena</i>		<i>2</i>	<i>Alice</i>
<i>3º</i>		<i>Est. Ling.</i>	<i>Adelma Moura</i>	<i>5º B</i>	<i>1</i>	<i>Adelma</i>

DIRETOR *[Assinatura]*

certifico que a presente copia fotostática, foi reproduzida do original que me foi apresentado Dou fé.

[Assinatura] 12.07.85

TABELÃO DE FALTAS

CARTÓRIO ÚNICO

LUIZ BONZAGA DE TORRES

TABELEÃO

Oficial do Reg. de Imóveis, Títulos e Documentos e de Protesto de Títulos

JOANA CLEONICE TORRES

Substituta

IPURI - PERNAMBUCO

Figura 43 – Resultado do algoritmo Ávila-Lins da imagem 3125 da Base1_7. Tempo de processamento 0,491096s. Borda “preta” à direita foi removida. Ruído na imagem do documento continua com ruído. Tamanho 107KB.

PONTO DO PESSOAL DA UNIDADE EDUCATIVA N.º 10 DE JACUAREMA
Jacubi, 06 DE Novembro DE 1985

N.º de Ordem	MATRICULA	DISCIPLINA	PUNTO, N.º M. e "Caric"	CLASSES	Total de Pontos	OBSERVAÇÕES
1º 1º		Port.	5ª 1ª			
3º 1º		Mat.	5ª A	5ª A	9	Maria
5º		geog.	Rita Barbosa	5ª H	2	Maria
1º 2º			Leonidas Miranda	5ª H	1	H.
3º 1º		Mat.	5ª C			
5º		Port.	Douzinha	5ª C	2	Douze
3º		geog.	Jacomo Barros	5ª C	9	Maria
5º			Adelma Moura	5ª C	1	Adelma
1º 2º	762/1	Mat.	5ª 4ª	6ª A	2	
3º		Mat.	Amélia Mudo	6ª A	01	
4º 5º		Ciência	Rita de Cássia	6ª B	2	
1º		Port.	3ª Magistério			
3º		Met. Diac.	Francinete Barros	2ª	1	S.
3º	762/1	Dinâmica	Sonha Torquato	2ª Ped.	01	
4º 5º	762/1	Dir. Turmas	Amélia Mudo	2ª Ped.	1	
			Amélia Mudo	2ª Ped.	2	
1º 2º		Psicologia	3ª Magistério			
3º		Port.	Leonidas Miranda	2ª 5ª M.	2	
4º		Met. Diac.	Francinete Barros	3ª	1	
5º		Sociologia	Sonha Torquato	3ª Ped.	01	
1º 2º		Port.	João Bosco Pereira		03	
3º		Relig.	5ª 1ª			
			Alice Madalena		2	Alice
			Adelma Moura	5ª B	1	Adelma

DIRETOR *[Assinatura]*

certifico que a presente cópia fotostática, foi reproduzida do original que me foi apresentado
 Jacubi, 07 de 1985
[Assinatura]

CARTÓRIO ÚNICO
 LUIZ MONZAGA DE TORRES
 TABELIAO
 Oficial do Reg. de Imóveis;
 Títulos e Documentos e de
 Protesto de Títulos
 JOANA CLEONICE TORRES
 Substituta
 JACUAREMA - PERNAMBUCO

Figura 44 – Resultado do algoritmo Formiga-Lins da imagem 3125 da Base1_7. Tempo de processamento 0,314593s. Borda “preta” à direita foi removida. Ruído na imagem do documento continua com ruído. Similar ao resultado do Ávila-Lins. Tamanho 107KB.

PONTO DO PESSOAL D *Acad. Políg. Militar N. Sra. do Socorro*
Louis, 06 DE Novembro DE 1982

Nº de Ordem	MATRÍCULA	DISCIPLINA	N. O M E	CLASSES	Total de Faltas	OBSERVAÇÕES
			<i>Punta para "Torre"</i> <i>5ª 1ª</i>			
1ª 1ª		Port.	<i>Aparecida Barros</i>	5ª A 5ª A	2	<i>Maria</i>
3ª 1ª		Mat.	<i>Rita Barbosa</i>	5ª B 5ª B	2	<i>Paulina</i>
5ª		Geog.	<i>Leonides Miranda</i> <i>5ª (5ª)</i>	5ª B	1	<i>J</i>
1ª 2ª		Mat.	<i>Ouzinha</i>	5ª C 5ª C	2	<i>Duse</i>
3ª 1ª		Port.	<i>Sacoro Barros</i>	5ª C 5ª C	2	<i>Maria</i>
5ª		Geog.	<i>Adelma Moura</i> <i>6ª 4ª</i>	5ª C	1	<i>Adelma</i>
1ª 1ª	76211	Mat.	<i>Amélia Mudo</i>	6ª 6ª	2	<i>J</i>
3ª		Hist.	<i>Rita de Cássia</i>	6ª A	01	<i>Rita</i>
1ª 2ª		Ciência	<i>Tainha Mudo</i> <i>3ª Magisterio</i>	6ª B	2	<i>Rita</i>
1ª		Port.	<i>Francinete Barros</i>	2ª	1	<i>J</i>
3ª		Met. Hist.	<i>Tainha Torquato</i>	2ª Ped.	01	<i>J</i>
3ª	76211	Química	<i>Amélia Mudo</i>	2ª Ped.	1	<i>J</i>
1ª 2ª	76211	Diretrizes	<i>Amélia Mudo</i> <i>3ª Magisterio</i>	2ª Ped.	2	<i>J</i>
1ª 1ª		Psicologia	<i>Leonides Miranda</i>	2ª 5ª 116	2	<i>Rita</i>
3ª		Port.	<i>Francinete Barros</i>	3ª	1	<i>J</i>
4ª		Met. Hist.	<i>Tainha Torquato</i>	3ª Ped.	01	<i>J</i>
5ª		Sociologia	<i>João Bosco Pereira</i> <i>5ª 3ª 1ª 1ª</i>		03	<i>Rita</i>
1ª 1ª		Port.	<i>Alice Madalena</i>		2	<i>Alice</i>
3ª		Ed. Ling.	<i>Adelma Moura</i>	5ª B	1	<i>Adelma</i>

DIRETOR *[Assinatura]*

certifico que a presente copia fotostática, foi reproduzida do original que me foi apresentado
 Dou fé
 UBI 22.07.19 25

LUIZ MONZAGA DE TORRES
 TABELIÃO
 Oficial do Reg. de Imóveis,
 Títulos e Documentos e de
 Protesto de Títulos
 JOANA CLEONICE TORRES
 Substituta
 CPUBI - PERNAMBUCO

Figura 45 – Resultado do Novo Algoritmo da imagem 3125 da BASE1_7. Tempo de processamento 0,462969s.

Borda “preta” removida e ruído na imagem do documento removido.

Tempo de processamento maior que o Formiga-Lins .Tamanho 68,7KB.

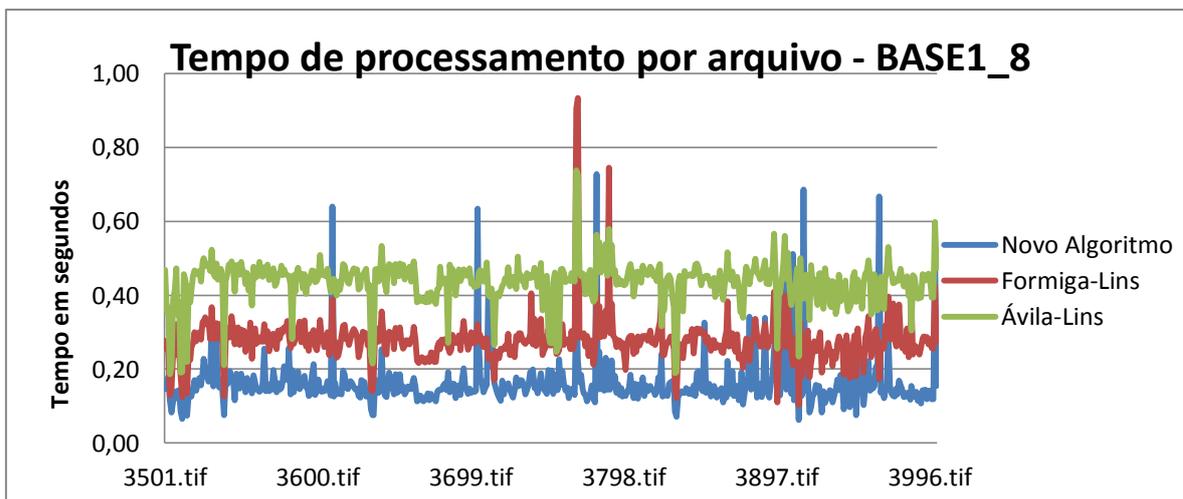


Gráfico 9 - Comparação do tempo de processamento por arquivo - grupo BASE1_8

BASE1_8			
	TEMPO MÉDIO POR ARQUIVO (s)	TEMPO TOTAL (s)	VARIÂNCIA
ÁVILA-LINS	0,433800	216,900063	0,003616
FORMIGA-LINS	0,276547	138,273510	0,003932
NOVO ALGORITMO	0,161948	80,974110	0,005093

Tabela 24 - Comparação do tempo médio de processamento por arquivo e do tempo total de processamento da BASE1_8

BASE1_8	
	REDUÇÃO DE TEMPO EM RELAÇÃO AO ÁVILA-LINS (%)
FORMIGA-LINS	36,25
NOVO ALGORITMO	62,67

Tabela 25 - Redução do tempo de processamento do Formiga-Lins e do Novo Algoritmo em relação ao Ávila-Lins na BASE1_8

BASE1_8		
	TAMANHO TOTAL (KB)	TAMANHO MÉDIO (KB)
ORIGINAL	17.510,40	35,02
ÁVILA-LINS	16.793,60	33,59
FORMIGA-LINS	16.691,20	33,38
NOVO ALGORITMO	16.896,00	33,79

Tabela 26 - Comparação do tamanho final da BASE1_8 de cada algoritmo em relação ao tamanho original em KB

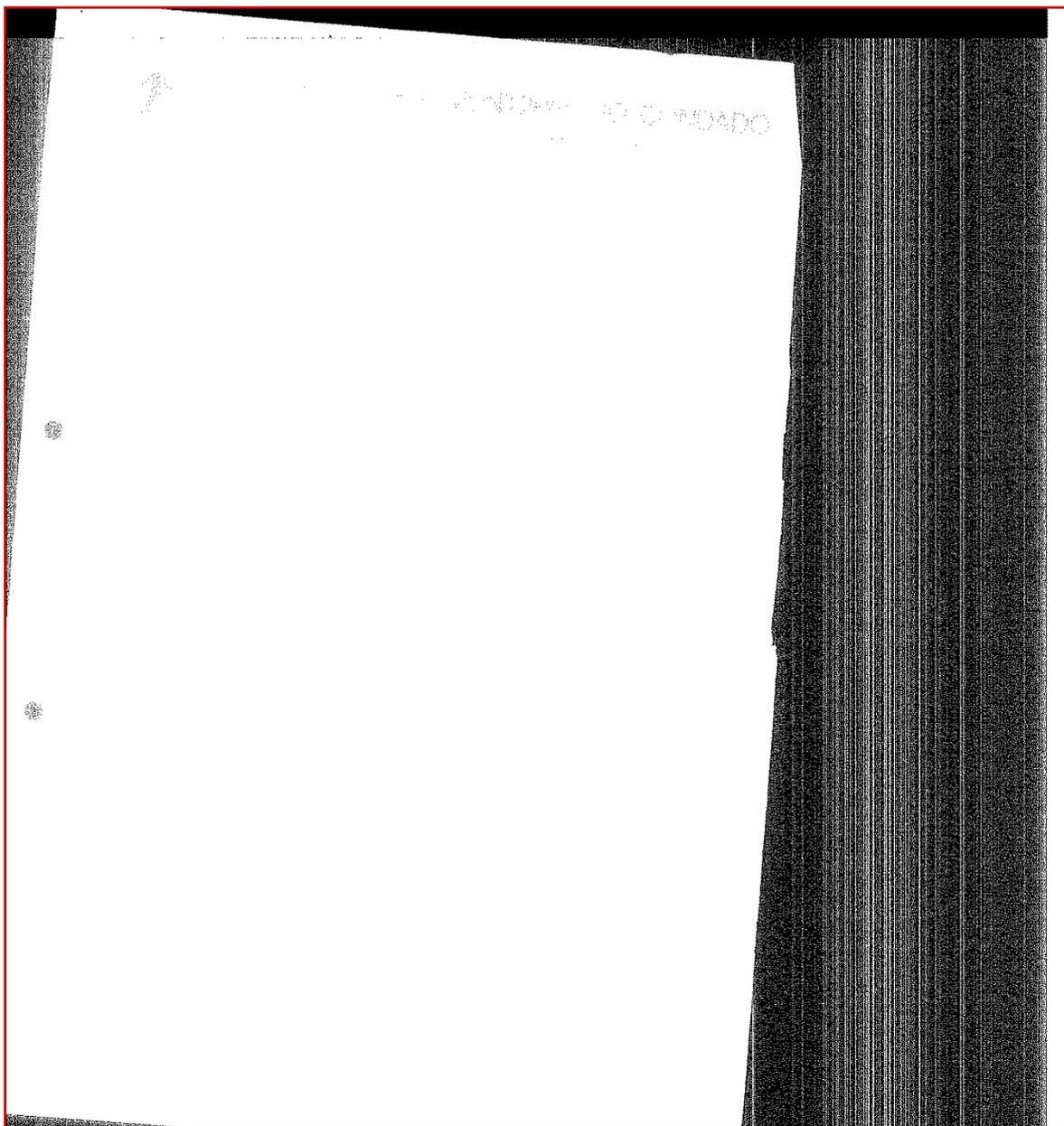


Figura 46 – Imagem 3768 da BASE1_8. Imagem com borda muito ruidosa, faixas brancas e ruído *sal e pimenta* severo à direita. Apresenta rotação. Tamanho 220KB.

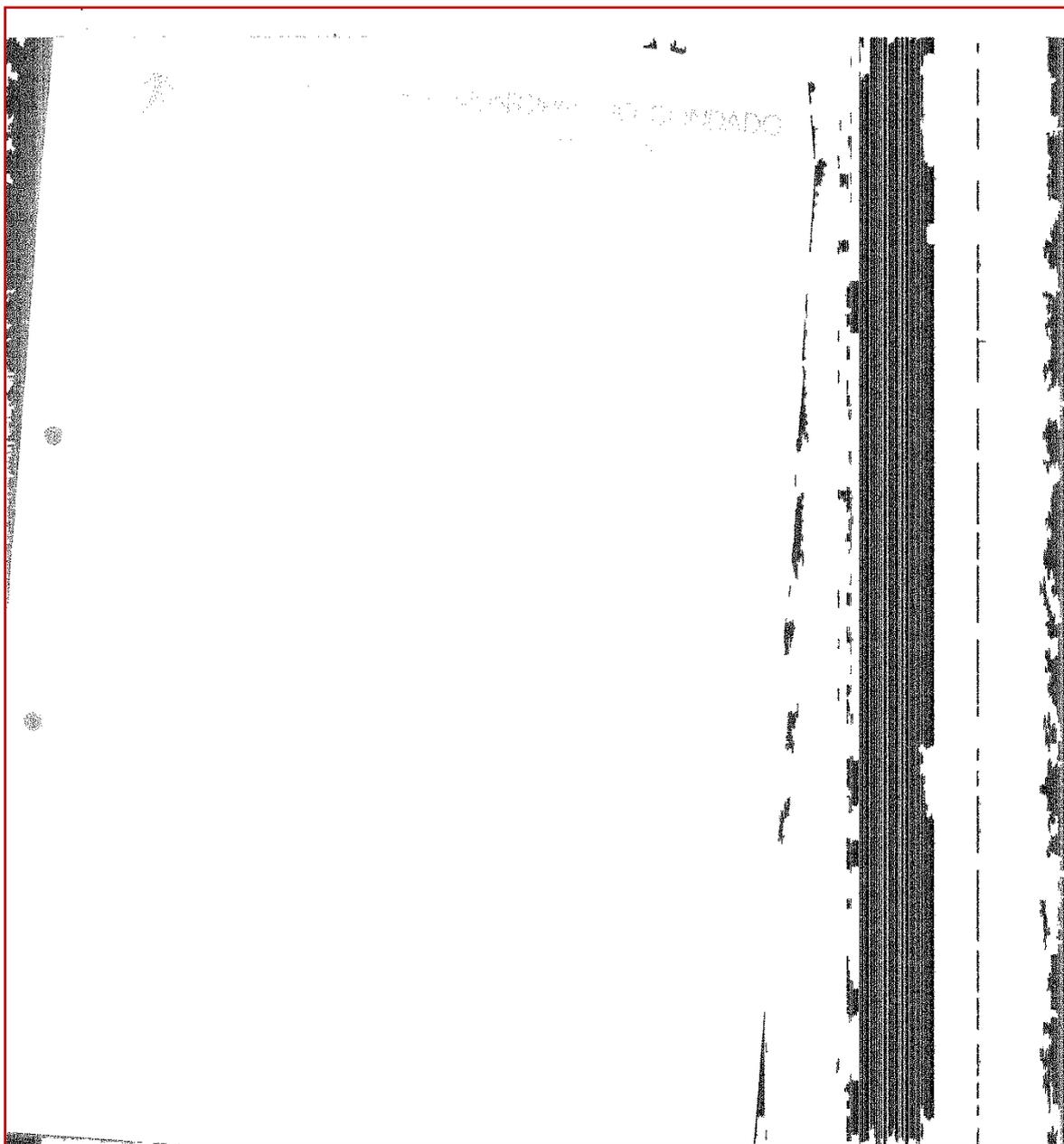
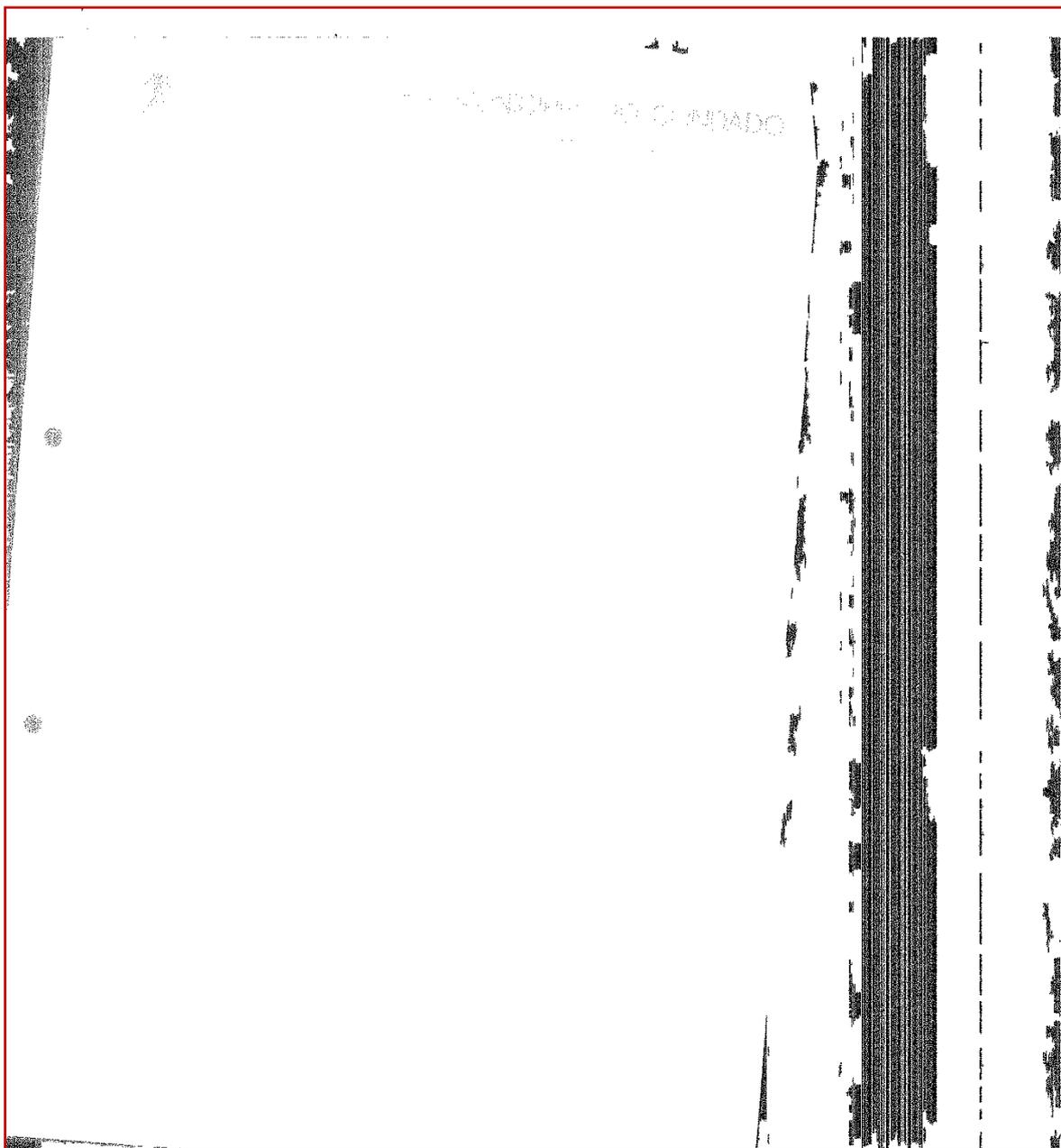
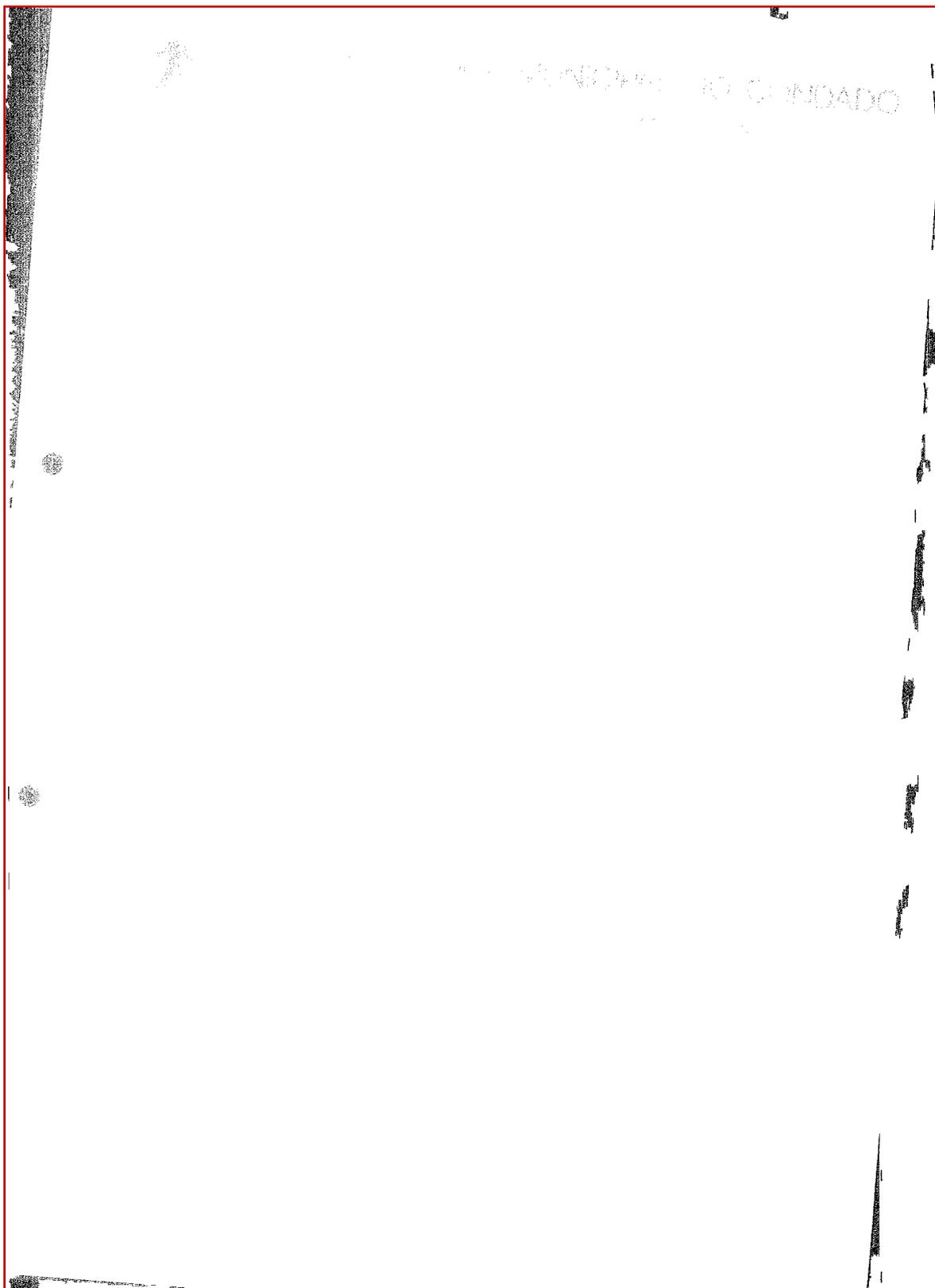


Figura 47 – Resultado do algoritmo Ávila-Lins da imagem 3768 da BASE1_8. Tempo de processamento 0,725276s.

A borda ruidosa não foi removida. Tamanho 88,8KB.



**Figura 48 – Resultado do algoritmo Formiga-Lins da imagem 3768 da BASE1_8.
Tempo de processamento 0,933549s. Resultado similar ao Ávila-Lins, borda ruidosa não foi removida.
Tempo de processamento maior que o Ávila-Lins. Tamanho 87,7KB.**



**Figura 49 – Resultado do Novo Algoritmo da imagem 3768 da BASE1_8.
Tempo de processamento 0,224066s. A borda ruidosa foi removida.
Tempo de processamento menor que o Ávila-Lins e o Formiga-Lins. Tamanho 14,1KB.**

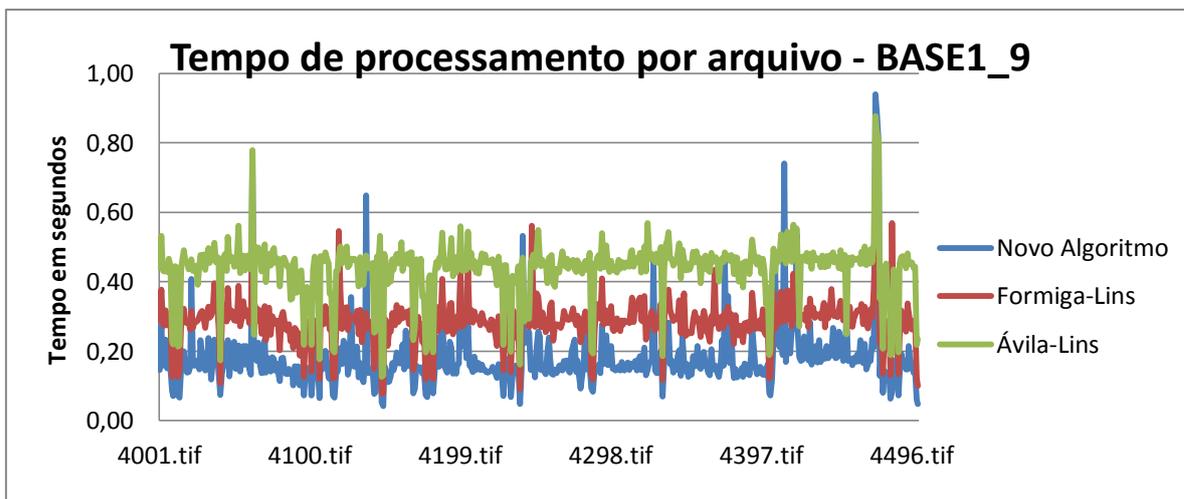


Gráfico 10 - Comparação do tempo de processamento por arquivo - grupo BASE1_9

BASE1_9			
	TEMPO MÉDIO POR ARQUIVO (s)	TEMPO TOTAL (s)	VARIÂNCIA
ÁVILA-LINS	0,437650	218,824820	0,006660
FORMIGA-LINS	0,284849	142,424698	0,003621
NOVO ALGORITMO	0,177047	88,523252	0,008293

Tabela 27 - Comparação do tempo médio de processamento por arquivo e do tempo total de processamento da BASE1_9

BASE1_9	
	REDUÇÃO DE TEMPO EM RELAÇÃO AO ÁVILA-LINS (%)
FORMIGA-LINS	34,91
NOVO ALGORITMO	59,55

Tabela 28 - Redução do tempo de processamento do Formiga-Lins e do Novo Algoritmo em relação ao Ávila-Lins na BASE1_9

BASE1_9		
	TAMANHO TOTAL (KB)	TAMANHO MÉDIO (KB)
ORIGINAL	22.528,00	45,06
ÁVILA-LINS	22.425,60	44,85
FORMIGA-LINS	22.528,00	45,06
NOVO ALGORITMO	22.220,80	44,44

Tabela 29 - Comparação do tamanho final da BASE1_9 de cada algoritmo em relação ao tamanho original em KB


 Governo do Estado de Pernambuco
 Secretaria de Educação, Cultura e Esportes

PROCESSO N. 8123802 - DATA: 13/06/95
 NOME: ANTONIA PAIVA DA SILVA GUERRA
 MAT 79968

ESPAÇO RESERVADO PARA O DESPACHO FINAL
*Anota de nos termos do
 Decreto 1402 de 22.03.67, num
 total de 05 anos e 25 dias
 MEX 1404/95*

ASSUNTO: ANOTAÇÃO
 ORIGEM: SECRETARIA DE EDUCAÇÃO CULTURA
 INFORMAÇÕES DAS 7:30 AS 13:00 PELO
 TELEFONE: 3544432 - 3544434

BLOCO 1
 NOME: ANTONIA PAIVA DA SILVA GUERRA
 ENDEREÇO: Pça. da Matriz 259
 CIDADE: B. Verde
 CEP: 744 550-604/72
 SEXO: F
 DATA DO NASCIMENTO: 23/11/46
 ESTADO CIVIL: Casada
 IDENTIDADE: 902.614 SSP-PE
 Nº: 1989857009064 LEITOR
 ZONA: 69 SEÇÃO: 35
 PÁGEP: 100512233 14
 CARGO/FUNÇÃO: PROFESSOR MATEMÁTICA
 SIMBOLO/NÍVEL/CATEGORIA OU PADRÃO: PROF. HAB. 3 * F.S.IX
 UNIDADE DE TRABALHO: Esc. Regina Pacis
 DEPTO. CAMPUS: ALTO PÁSEO
 DIRETORIA: DCE
 VINCULO: CLT EFETIVO

REQUER AO
 EXMO. SR. SECRETÁRIO DE EDUCAÇÃO E CULTURA

BLOCO 2

1 AGENO DE FALTA	8 LICENÇA POR MOTIVO DE DOENÇA EM PESSOA DA FAM.
2 AFASTAMENTO FISUBMETER-SE A PROVAS	9 LICENÇA PRÊMIO (GOZOS)
3 AFASTAMENTO PELO ART. 177 DA LEI 6123-68	10 REMOÇÃO
4 ANOTAÇÃO DE DIPLOMA E/OU CURSO	11 READAPTAÇÃO DE CARGO
5 CERTIDÃO P/INS ESPECÍFICOS	12 RETIFICAÇÃO DE NOME
6 GRATIFICAÇÃO A PROFESSORES DE EXCEPCIONAIS	13 SALÁRIO FAMILIA
7 LICENÇA A GESTANTE	14 OUTROS

REQUER AO
 EXMO. SR. SECRETÁRIO DE ADMINISTRAÇÃO

BLOCO 3

15 AFASTAMENTO P/CURSO, ESTÁGIO OU CONGRESSO	20 LICENÇA P/ TRATAMENTO DE INTERESSE PARTICULAR
16 ANOTAÇÃO DE TEMPO DE SERVIÇO	21 LICENÇA PRÊMIO (CONCESSÃO)
17 APOSENTADORIA	22 REAJUSTE SALARIAL
18 EXONERAÇÃO (REQ. REGER A FIRMA)	23 RESCISÃO DE CONTRATO (RECONHECER A FUMAM)
19 LICENÇA PARA ACOMPANHAR O MARIDO	24 SUSPENSÃO DE CONTRATO
19 LICENÇA P/ TRATAMENTO DE SAÚDE	<input checked="" type="checkbox"/> OUTROS

BLOCO 4

ANEXO DOCUMENTOS NECESSÁRIOS A SOLUÇÃO DO ASSUNTO E PARA QUE PRESTA AS SEGUINTE
 INFORMAÇÕES COMPLEMENTARES:

requerente solicita que seja anexado a sua ficha funcional, tempo de
 serviço em anexo.

BLOCO 5

EM 15/06/95
 Antonia Paiva da S. Guerra
 ASSINATURA DO REQUERENTE

EM 07/06/95
 Maria Cristina
 ASSINATURA DO RESPONSÁVEL

EM 18/05/95
 Daniel
 DIRETOR

Figura 50 – Imagem 4405 da BASE1_9. Imagem com borda “preta” à esquerda e documento com borda irregular à esquerda. Tamanho 98,6KB

70-0570


 Governo do Estado de Pernambuco
 Secretaria de Educação, Cultura e Esportes

PROCESSO N. 8123802 DATA: 13/06/95
 NOME: ANTONIA PAIVA DA SILVA GUERRA MAT70968
 ASSUNTO: ANOTAÇÃO
 ORIGEM: SECRETARIA DE EDUCAÇÃO CULTURA
 INFORMACOES DAS 7:30 AS 13:00 PELO
 TELEFONE: 4244432 DAMAT9:155454

ESPAÇO RESERVADO PARA O DESPACHO FINAL
*Anote-se nos termos do
 Decreto 1402 de 22.03.67, num
 total de 05 anos e 25 dias
 DEC 16.04.96*

BLOCO 1

NOME: ANTONIA PAIVA DA SILVA GUERRA

ENDEREÇO: Pça. da Matriz 259
 CIDADE: B. Verde ESTADO: PE TELEFONE: 846 1296

CPF: 744 550 604/72 SEXO: F DATA DO NASCIMENTO: 23/11/46 ESTADO CIVIL: Casada

IDENTIDADE: 902.614 SSP-PE Nº: 198989809 DE ALEITOR: 100512233 14 PASEP:

CARGO/FUNÇÃO: PROFESSOR SÍMBOLO/NÍVEL/CATEGORIA OU PADRÃO: PROF. HAB.3 * F.S.IX

UNIDADE DE TRABALHO: Esc. Regina Pacis DEPTO./CAMPE: ALTO PAJEU DIRETORIA: DCE VÍNCULO: CLT EFET

REQUER AO

EXMO. SR. SECRETÁRIO DE EDUCAÇÃO E CULTURA

BLOCO 2

<input type="checkbox"/> 1 ATRASO DE FALTA	<input type="checkbox"/> 8 LICENÇA POR MOTIVO DE DOENÇA EM PESSOA DA FAM.
<input type="checkbox"/> 2 AFASTAMENTO P/SUBMETER-SE A PROVAS	<input type="checkbox"/> 9 LICENÇA PRÊMIO (GOZO)
<input type="checkbox"/> 3 AFASTAMENTO PELO ART. 177 DA LEI 6123:68	<input type="checkbox"/> 10 REMOÇÃO
<input type="checkbox"/> 4 ANOTAÇÃO DE DIPLOMA E/OU CURSO	<input type="checkbox"/> 11 READAPTAÇÃO DE CARGO
<input type="checkbox"/> 5 CERTIDÃO P/FINS ESPECÍFICOS	<input type="checkbox"/> 12 RETIFICAÇÃO DE NOME
<input type="checkbox"/> 6 GRATIFICAÇÃO A PROFESSORES DE EXCEPCIONAIS	<input type="checkbox"/> 13 SALÁRIO FAMÍLIA
<input type="checkbox"/> 7 LICENÇA À GESTANTE	<input type="checkbox"/> OUTROS

REQUER AO

EXMO. SR. SECRETÁRIO DE ADMINISTRAÇÃO

BLOCO 3

<input type="checkbox"/> 14 AFASTAMENTO P/CURSO, ESTÁGIO OU CONGRESSO	<input type="checkbox"/> 20 LICENÇA P/TRATO DE INTERESSE PARTICULAR
<input checked="" type="checkbox"/> 15 ANOTAÇÃO DE TEMPO DE SERVIÇO	<input type="checkbox"/> 21 LICENÇA PRÊMIO (CONCESSÃO)
<input type="checkbox"/> 16 APOSENTADORIA	<input type="checkbox"/> 22 REAJUSTE SALARIAL
<input type="checkbox"/> 17 EXONERAÇÃO (REP. HEÇER A FIRMA)	<input type="checkbox"/> 23 RESCISÃO DE CONTRATO (RECONHECER A FIRMA)
<input type="checkbox"/> 18 LICENÇA PARA ACOM. PANHAR O MARIDO	<input type="checkbox"/> 24 SUSPENSÃO DE CONTRATO
<input type="checkbox"/> 19 LICENÇA P/TRATAMENTO DE SAÚDE	<input checked="" type="checkbox"/> OUTROS

BLOCO 4 ANEXO DOCUMENTOS NECESSÁRIOS À SOLUÇÃO DO ASSUNTO E PARA QUE PRESTA AS SEGUINTE

INFORMAÇÕES COMPLEMENTARES:

requerente solicita que seja anexado a sua ficha funcional, tempo de serviço em anexo.

BLOCO 5

EM 15.10.95
Antonia Paiva da S. Guerra
 ASSINATURA DO REQUERENTE

CIENTE ENCAMINHE-SE
 EM 07/06/95
Luciana Custosa
 ASSINATURA DO RESPONSÁVEL

VISTO
 EM 15/05/96
Antônio
 DIRETOR

15/11/95
Antônio

Figura 51 – Resultado do algoritmo Ávila-Lins da imagem 4405. Tempo de processamento 0,496218s. Borda preta removida e irregularidades da borda do documento imperceptíveis. Tamanho 93,2KB.


 Governo do Estado de Pernambuco
 Secretaria de Educação, Cultura e Esportes

PROCESSO N. 8123802 DATA: 13/06/95
 NOME: ANTONIA PAIVA DA SILVA GUERRA
 MAT 70968
 ASSUNTO: ANOTAÇÃO
 ORIGEM: SECRETARIA DE EDUCAÇÃO CULTURA
 INFORMAÇÕES DAS 7:30 AS 13:00 PELO
 TELEFONE: 4244422 RAMAL: 4554454

ESPAÇO RESERVADO PARA O DESPACHO FINAL
Anote-se nos termos do
Decreto 1402 de 22.03.67, num
total de 05 anos e 25 dias
DEF 16.04.96

BLOCO 1

NOME: ANTONIA PAIVA DA SILVA GUERRA
 ENDEREÇO: Pça. da Matriz 259
 CIDADE: B. Verde PE
 CEP: 744 550 604/72
 SEXO: F
 DATA DO NASCIMENTO: 23/11/46
 ESTADO CIVIL: Casada
 IDENTIDADE: 902.614 SSP-PE
 Nº: 198989009054 LEITOR
 ZONA: 69 SEÇÃO: 35 PASEP: 100512233 14
 CARGO/FUNÇÃO: PROFESSOR REQUERENTE
 SIMBOLO/NIVEL/CATEGORIA OU PADRÃO: PROF. HAB.3 * F.S.IX
 UNIDADE DE TRABALHO: Esc. Regina Pacis
 DEPTO./CAMPE: ALTO PAJEU
 DIRETORIA: DCE
 VINCULO: CLT EFET

REQUER AO

EXMO. SR. SECRETÁRIO DE EDUCAÇÃO E CULTURA

BLOCO 2

1 AFASTAMENTO P/SUBMETTER-SE A PROVAS	8 LICENÇA POR MOTIVO DE DOENÇA EM PESSOA DA FAM.
2 AFASTAMENTO PELO ART. 177 DA LEI 6123:68	9 LICENÇA PRÊMIO (GOZO)
3 ANOTAÇÃO DE DIPLOMA E/OU CURSO	10 REMOÇÃO
4 IDENTIDADE P/FINS ESPECÍFICOS	11 READAPTAÇÃO DE CARGO
5 GRATIFICAÇÃO A PROFESSORES DE EXCEPCIONAIS	12 RETIFICAÇÃO DE NOME
6 LICENÇA À GESTANTE	13 SALÁRIO FAMÍLIA
	OUTROS

REQUER AO

EXMO. SR. SECRETÁRIO DE ADMINISTRAÇÃO

BLOCO 3

14 AFASTAMENTO P/CURSO, ESTÁGIO OU CONGRESSO	20 LICENÇA P/TRATO DE INTERESSE PARTICULAR
15 ANOTAÇÃO DE TEMPO DE SERVIÇO	21 LICENÇA PRÊMIO (CONCESSÃO)
16 APOSENTADORIA	22 REAJUSTE SALARIAL
17 EXONERAÇÃO (RECONHECER A FIRMA)	23 RESCISÃO DE CONTRATO (RECONHECER A FIRMA)
18 LICENÇA PARA ACOA PANHAR O MARIDO	24 SUSPENSÃO DE CONTRATO
19 LICENÇA P/TRATAMENTO DE SAÚDE	X OUTROS

BLOCO 4 ANEXO DOCUMENTOS NECESSÁRIOS À SOLUÇÃO DO ASSUNTO E PARA QUE PRESTA AS SEGUINTE

INFORMAÇÕES COMPLEMENTARES:

requerente solicita que seja anexado a sua ficha funcional, tempo de
serviço em anexo.

BLOCO 5

EM: 15/05/95
 ASSINATURA DO REQUERENTE: *Antonia Paiva da S. Guerra*

CLIENTE ENCAMINHE-SE
 EM: 07/06/95
 ASSINATURA DO RESPONSÁVEL: *Guerra*

VISU
 EM: 15/06/95
 DIRETOR: *Guerra*

Figura 52 – Resultado do algoritmo Formiga-Lins da imagem 4405 da BASE1_9. Tempo de processamento 0,370507s. Resultado similar ao Ávila-Lins. Borda “preta” removida e irregularidades na borda do documento imperceptíveis. Tamanho 93,2KB.


 Governo do Estado de Pernambuco
 Secretaria de Educação, Cultura e Esportes

PROCESSO N.8123802 DATA:13/06/95
 NOME: ANTONIA PAIVA DA SILVA GUERRA
 MAT70968
 ASSUNTO: ANOTACAO
 ORIGEM: SECRETARIA DE EDUCACAO CULTURA
 INFORMACOES DAS 7:30 AS 13:00 PELO
 TELEFONE: 4244422 DAMAT:4554454

ESPAÇO RESERVADO PARA O DESPACHO FINAL
Anote-se nos termos do
Decreto 1402 de 22.03.67, num
total de 05 anos e 25 dias
DEF 16.04.96

BLOCO 1

NOME: ANTONIA PAIVA DA SILVA GUERRA
 ENDEREÇO: Pça.da Matriz 259
 CIDADE: B. Verde PE
 CEP: 744 550 604/72
 SEXO: F
 DATA DO NASCIMENTO: 23/11/46
 ESTADO CIVIL: Casada
 IDENTIDADE: 902.614 SSP-PE
 Nº: 198989009054
 ZONA: 69 SEÇÃO: 35
 PASEP: 100512233 14
 CARGO/FUNÇÃO: PROFESSOR REQUERENTE
 SIMBOLO/NIVEL/CATEGORIA OU PADRÃO: PROF. HAB.3 * F.S.IX
 UNIDADE DE TRABALHO: Esc.Regina Pacis
 DEPTO./CAMPE: ALTO PAJEU
 DIRETORIA: DCE
 VINCULO: CLT EFET

REQUER AO EXMO.SR. SECRETÁRIO DE EDUCAÇÃO E CULTURA

BLOCO 2

1 AFASTAMENTO P/SUBMETTER-SE A PROVAS	8 LICENÇA POR MOTIVO DE DOENÇA EM PESSOA DA FAM.
2 AFASTAMENTO PELO ART. 177 DA LEI 6123:68	9 LICENÇA PRÊMIO (GOZO)
3 ANOTACAO DE DIPLOMA E/OU CURSO	10 REMOCAO
4 IDENTIDAO P/FINS ESPECIFICOS	11 READAPTAÇÃO DE CARGO
5 GRATIFICACAO A PROFESSORES DE EXCEPCIONAIS	12 RETIFICACAO DE NOME
6 LICENÇA A GESTANTE	13 SALARIO FAMILIA
	OUTROS

REQUER AO EXMO. SR. SECRETÁRIO DE ADMINISTRAÇÃO

BLOCO 3

14 AFASTAMENTO P/CURSO, ESTÁGIO OU CONGRESSO	20 LICENÇA P/TRATO DE INTERESSE PARTICULAR
15 ANOTACAO DE TEMPO DE SERVIÇO	21 LICENÇA PRÊMIO (CONCESSAO)
16 APOSENTADORIA	22 REAJUSTE SALARIAL
17 EXONERACAO (RECONHECER A FIRMA)	23 RESCISAO DE CONTRATO (RECONHECER A FIRMA)
18 LICENÇA PARA ACOA PANHAR O MARIDO	24 SUSPENSAO DE CONTRATO
19 LICENÇA P/TRATAMENTO DE SAUDE	X OUTROS

BLOCO 4 ANEXO DOCUMENTOS NECESSÁRIOS À SOLUÇÃO DO ASSUNTO E PARA QUE PRESTA AS SEGUINTE

INFORMAÇÕES COMPLEMENTARES:

requerente solicita que seja anexado a sua ficha funcional, tempo de
serviço em anexo.

BLOCO 5

EM: 15/05/95
 ASSINATURA DO REQUERENTE: *Antonia Paiva da S. Guerra*

CLIENTE ENCAMINHE-SE
 EM: 07/06/95
 ASSINATURA DO RESPONSÁVEL: *Guerra*

VISU
 EM: 15/06/95
 DIRETOR: *Imolani*

Figura 53 – Resultado do Novo Algoritmo da imagem 4405 da Base1_9. Tempo de processamento 0,241360s.
 Borda “preta” removida e irregularidades da borda do documento imperceptíveis.
 Tempo de processamento menor. Tamanho 93,2KB.

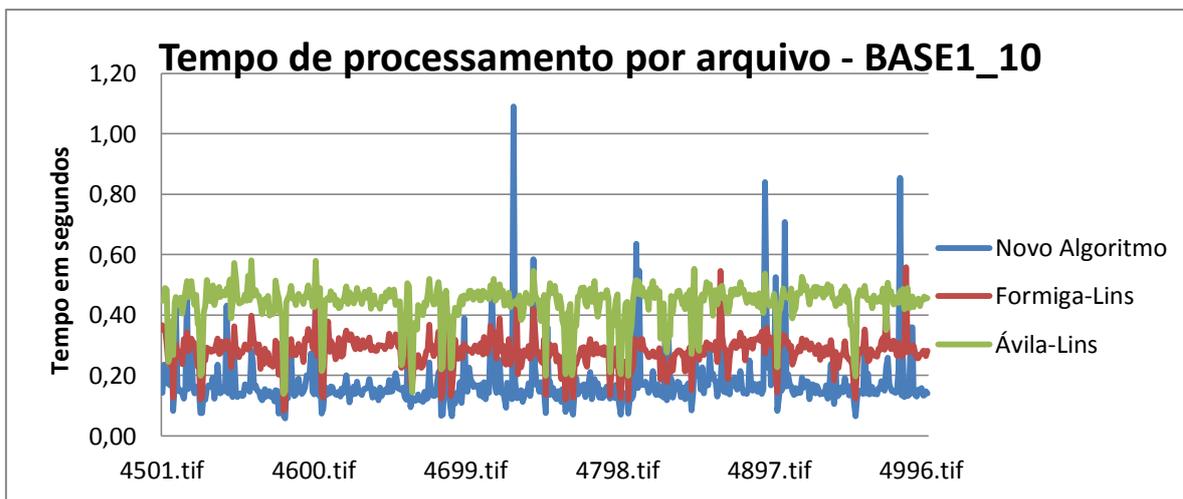


Gráfico 11 - Comparação do tempo de processamento por arquivo - grupo BASE1_10

BASE1_10			
	TEMPO MÉDIO POR ARQUIVO (s)	TEMPO TOTAL (s)	VARIÂNCIA
ÁVILA-LINS	0,438484	219,241846	0,004522
FORMIGA-LINS	0,278542	139,270787	0,002633
NOVO ALGORITMO	0,171265	85,632332	0,008655

Tabela 30 - Comparação do tempo médio de processamento por arquivo e do tempo total de processamento da BASE1_10

BASE1_10	
	REDUÇÃO DE TEMPO EM RELAÇÃO AO ÁVILA-LINS (%)
FORMIGA-LINS	36,48
NOVO ALGORITMO	60,94

Tabela 31 - Redução do tempo de processamento do Formiga-Lins e do Novo Algoritmo em relação ao Ávila-Lins na BASE1_10

BASE1_10		
	TAMANHO TOTAL (KB)	TAMANHO MÉDIO (KB)
ORIGINAL	17.100,80	34,20
ÁVILA-LINS	16.998,40	34,00
FORMIGA-LINS	16.896,00	33,79
NOVO ALGORITMO	16.896,00	33,79

Tabela 32 - Comparação do tamanho final da BASE1_10 de cada algoritmo em relação ao tamanho original.

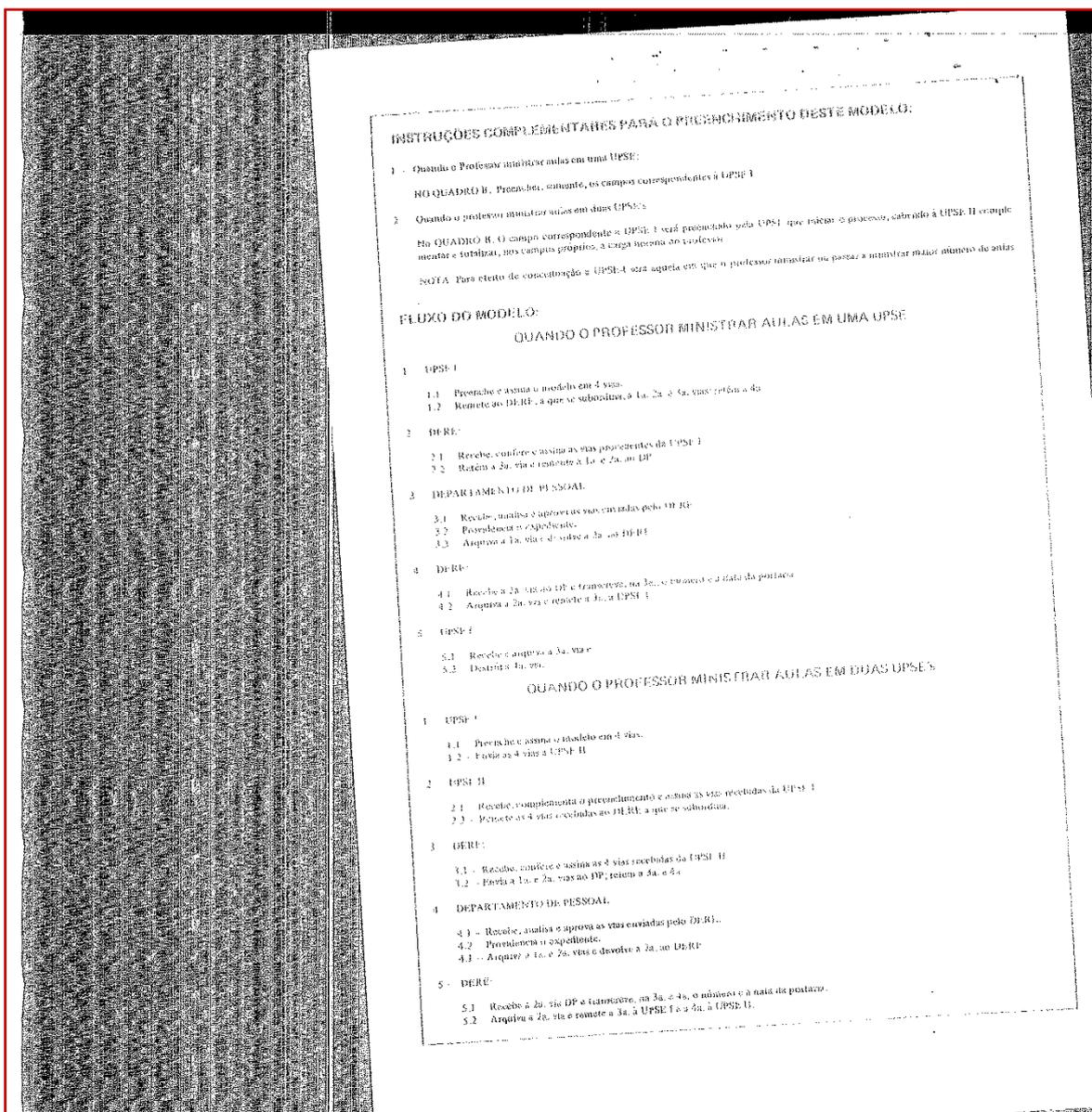


Figura 54 – Imagem 4973 da BASE1_10. Imagem com borda muito ruidosa à esquerda. Tamanho 230KB.

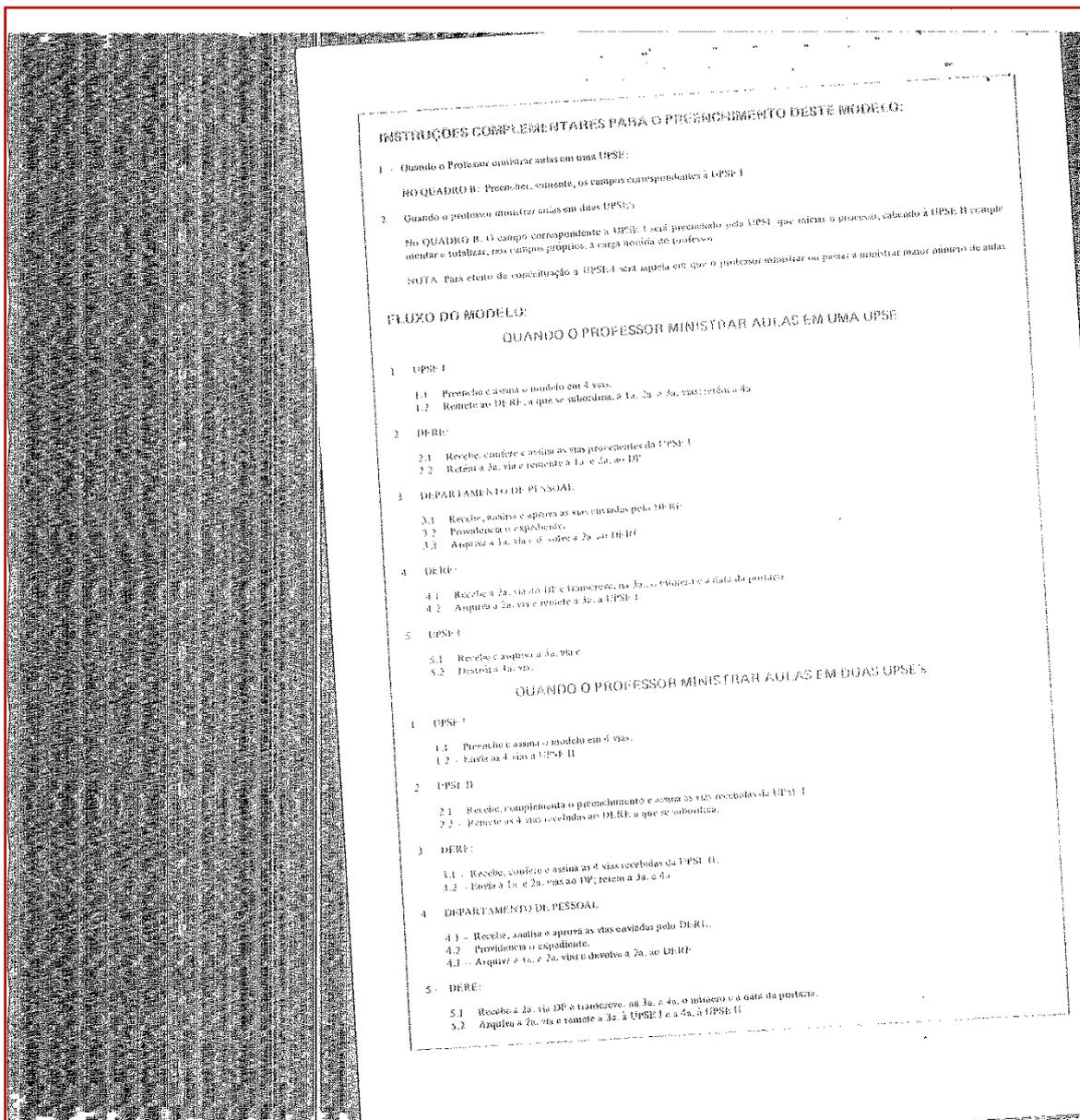


Figura 55 – Resultado do algoritmo Ávila-Lins da imagem 4973 da BASE1_10. Tempo de processamento 0,350591s. Borda ruidosa não foi removida. Tamanho 230KB.

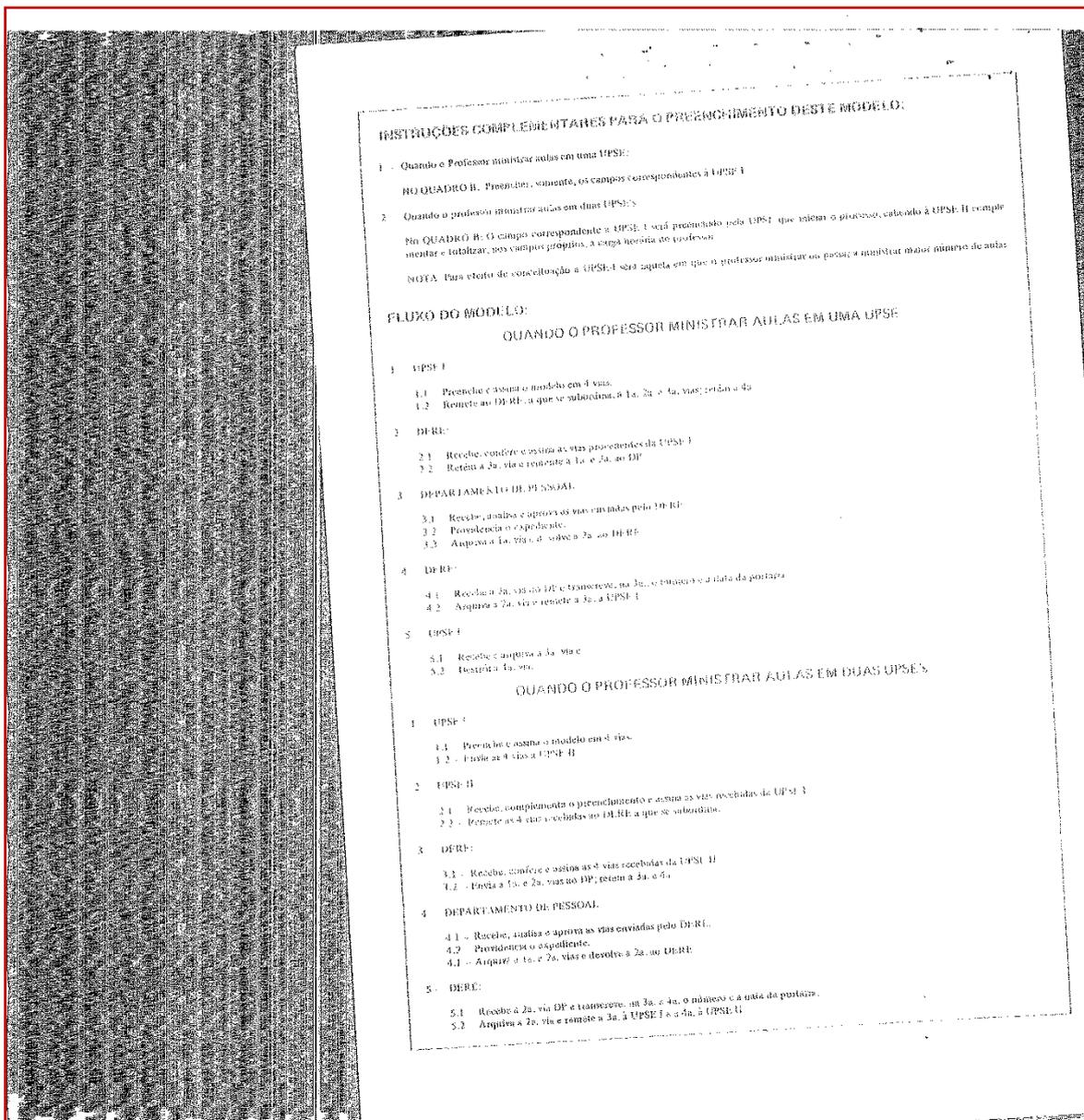


Figura 56 – Resultado do algoritmo Formiga-Lins da imagem 4973 da BASE1_10.
Tempo de processamento 0,441370s. A borda ruidosa não foi removida. Tamanho 230KB.

INSTRUÇÕES COMPLEMENTARES PARA O PREENCHIMENTO DESTES MODELOS:

1 - Quando o Professor ministrar aulas em uma UPSE I:
 NO QUADRO B: Preencher, somente, os campos correspondentes à UPSE I

2 - Quando o professor ministrar aulas em duas UPSE'S
 No QUADRO B: O campo correspondente à UPSE I será preenchido pela UPSI que iniciar o processo, cabendo à UPSE II complementar e totalizar, nos campos próprios, a carga horária do professor

NOTA: Para efeito de conceituação à UPSE I será aquela em que o professor ministrar ou passar a ministrar maior número de aulas

FLUXO DO MODELO:

QUANDO O PROFESSOR MINISTRAR AULAS EM UMA UPSE

- 1 - UPSE I
 - 1.1 - Preenche e assina o modelo em 4 vias.
 - 1.2 - Remete ao DERE, a que se subordina, a 1a, 2a e 3a vias; retém a 4a
- 2 - DERE:
 - 2.1 - Recebe, confere e assina as vias procedentes da UPSE I
 - 2.2 - Retém a 3a. via e remete a 1a. e 2a. ao DP.
- 3 - DEPARTAMENTO DE PESSOAL:
 - 3.1 - Recebe, analisa e aprova as vias enviadas pelo DERE
 - 3.2 - Providencia o expediente.
 - 3.3 - Arquivo a 1a. via e devolve a 2a. ao DERE
- 4 - DERE:
 - 4.1 - Recebe a 2a. via do DP e transcreve, na 3a., o número e a data da portaria
 - 4.2 - Arquivo a 3a. via e remete a 3a. à UPSE I
- 5 - UPSE I
 - 5.1 - Recebe e arquiva a 3a. via e
 - 5.2 - Destroi a 4a. via.

QUANDO O PROFESSOR MINISTRAR AULAS EM DUAS UPSE'S

- 1 - UPSE I
 - 1.1 - Preenche e assina o modelo em 4 vias.
 - 1.2 - Envia as 4 vias à UPSE II.
- 2 - UPSE II
 - 2.1 - Recebe, complementa o preenchimento e assina as vias recebidas da UPSE I
 - 2.2 - Remete as 4 vias recebidas ao DERE, a que se subordina.
- 3 - DERE:
 - 3.1 - Recebe, confere e assina as 4 vias recebidas da UPSE II.
 - 3.2 - Envia a 1a. e 2a. vias ao DP; retém a 3a. e 4a
- 4 - DEPARTAMENTO DE PESSOAL:
 - 4.1 - Recebe, analisa e aprova as vias enviadas pelo DERE.
 - 4.2 - Providencia o expediente.
 - 4.3 - Arquivo a 1a. e 2a. vias e devolve a 2a. ao DERE
- 5 - DERE:
 - 5.1 - Recebe a 2a. via DP e transcreve, na 3a. e 4a. o número e a data da portaria.
 - 5.2 - Arquivo a 2a. via e remete a 3a. à UPSE I e a 4a. à UPSE II.

Figura 57 – Resultado do Novo Algoritmo para a imagem 4973 da BASE1_10.

Tempo de processamento 0,226507s. Borda ruidosa parcialmente removida. Tamanho 59,6KB.

É possível observar em todos os gráficos que o Novo Algoritmo demonstra um melhor desempenho relativo ao tempo de processamento por arquivo na maioria dos casos. Com um tempo médio de processamento menor que o Formiga-Lins.

	TEMPO TOTAL (min)	TEMPO MÉDIO TOTAL (s)	VARIÂNCIA TOTAL
ÁVILA-LINS	37,378108	0,448537	0,009180
FORMIGA-LINS	23,687555	0,284251	0,004722
NOVO ALGORITMO	14,944995	0,179340	0,010298

Tabela 33 - Comparação do tempo total de processamento de todos os grupos

REDUÇÃO DE TEMPO TOTAL EM RELAÇÃO AO ÁVILA-LINS	
FORMIGA-LINS	36,63%
NOVO ALGORITMO	60,02%

Tabela 34 - Porcentagem da redução do tempo de processamento em relação ao Ávila-Lins

Observando os tempos totais de processamento para cada grupo, o Novo Algoritmo reduziu o tempo de processamento em relação ao Ávila-Lins em torno de 60%. Em relação ao Formiga-Lins apresentou uma redução entorno de 36,91%.

O tamanho final da base também foi reduzido em relação ao Formiga-Lins.

	Tamanho em MB da base completa
ORIGINAL	187,70
ÁVILA-LINS	183,50
FORMIGA-LINS	183,20
NOVO ALGORITMO	182,60

Tabela 35 – Comparação do tamanho total da base, em MB, original e após a aplicação de cada algoritmo.

Significa que houve um alcance maior do Novo Algoritmo em relação ao Formiga-Lins, ou seja, arquivos que antes não eram completamente processados ou simplesmente não eram processados, passaram a ser processados pela remoção de borda.

Na maioria dos testes o Novo Algoritmo apresentou uma maior variância no tempo de processamento em relação ao Ávila-Lins e ao Formiga-Lins. Tal fato se dá devido à capacidade de escalonamento do Novo Algoritmo, e aos casos onde o documento apresenta elementos pictóricos que são representados por elementos pretos de textura sólida, linhas ou outros elementos que produzem sequências de pixels pretos contíguos. Essas sequências podem, coincidentemente, estar na trajetória de varredura do Novo Algoritmo, e acabam

“enganando-o”, de forma que uma largura não válida seja detectada e force a troca de linha para varredura várias vezes. Esta situação é demonstrada pelas imagens cujo tempo de processamento introduziu nos gráficos picos anômalos. Como em:

- Base1_1:

Imagem	Tempo de Processamento
0195	1,461073
0376	0,915427
0461	1,854128
0500	1,872314

- Base 1_2:

Imagem	Tempo de Processamento
0518	0,949102
0558	1,511627
0564	0,801051
0791	0,951415
0860	1,587102

- Base 1_3:

Imagem	Tempo de Processamento
1254	0,883725

- Base 1_4:

Imagem	Tempo de Processamento
1742	0,869036
1822	0,850750

- Base 1_5:

Imagem	Tempo de Processamento
2143	0,578410
2199	0,497031
2233	0,660449
2305	0,550515
2431	0,601057

- Base 1_6:

Imagem	Tempo de Processamento
2652	0,572245
2747	0,889811
2925	0,777961

- Base 1_7:

Imagem	Tempo de Processamento
3143	1,000407
3190	0,744495
3286	0,584652
3310	1,186557

- Base 1_8:

Imagem	Tempo de Processamento
3609	0,639583
3703	0,634315
3780	0,726955
3914	0,685635
3963	0,666635

- Base 1_9:

Imagem	Tempo de Processamento
4137	0,648199
4412	0,741269
4472	0,939934
4473	0,878564
4474	0,797316

- Base 1_10:

Imagem	Tempo de Processamento
4730	1,090880
4810	0,635941
4894	0,840047
4907	0,707896
4982	0,854633

4.2 – Análise qualitativa das imagens filtradas

Os resultados da comparação quantitativa mostram que houve significativa melhoria no processamento das imagens pelo Ávila-Lins com o pré-processamento aqui proposto. E a análise qualitativa comprova a melhoria do processamento, primeiramente pelo processamento de imagens com intenso ruído *sal e pimenta*.

Alguns exemplos de imagens e resultados do Ávila-Lins, Formiga-Lins, de algumas ferramentas comerciais e o Novo Algoritmo são apresentados a seguir.

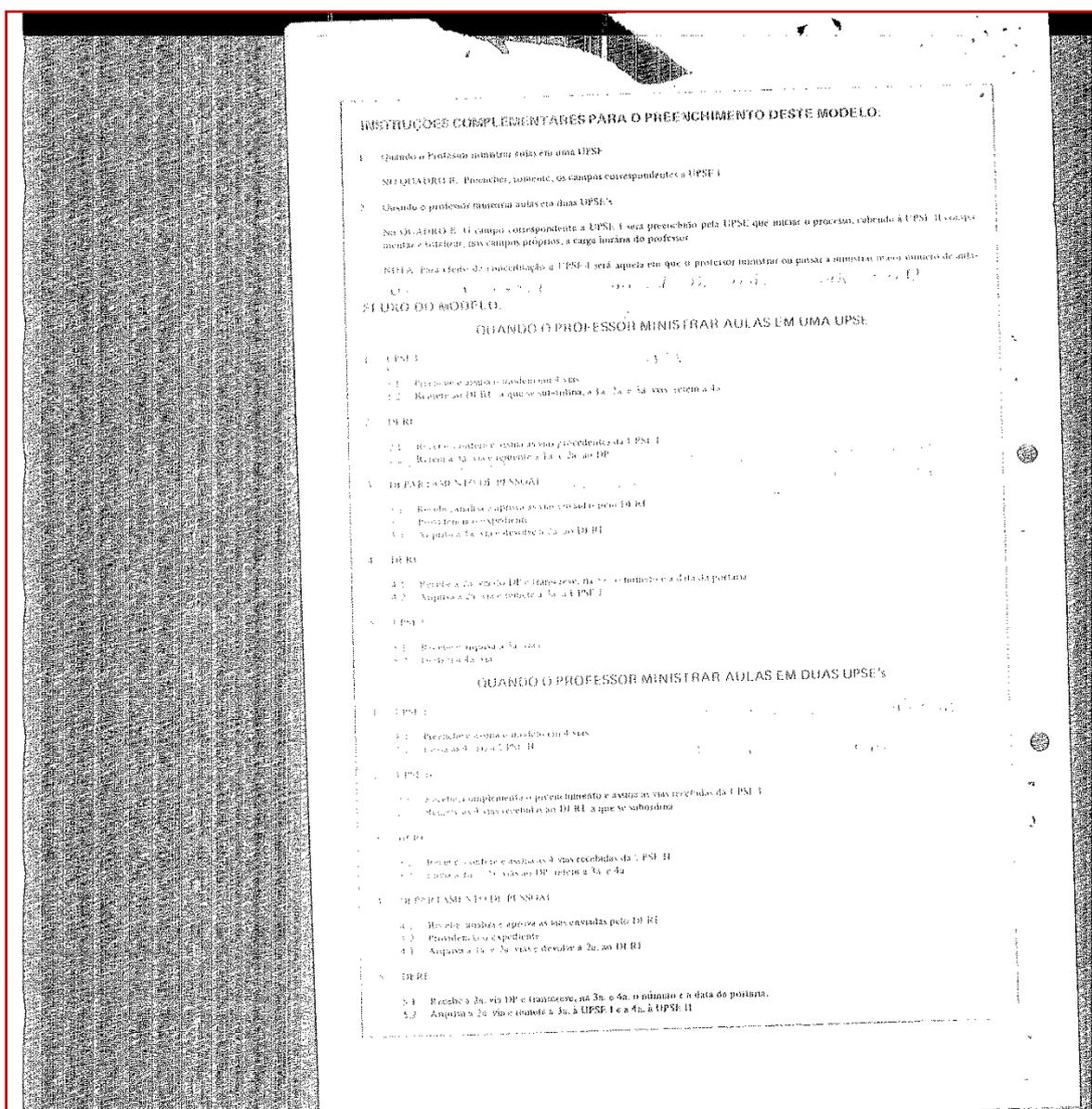


Figura 58 - Imagem 0141 da Base1_1, imagem original. Forte ruído *sal e pimenta* nas bordas. Tamanho 214KB.

INSTRUÇÕES COMPLEMENTARES PARA O PREENCHIMENTO DESTA MODELO:

1. Quando o Professor ministrar aulas em uma UPSE

No QUADRO B. Preencher, somente, os campos correspondentes a UPSE I

2. Quando o professor ministrar aulas em duas UPSE's

No QUADRO B. O campo correspondente a UPSE I será preenchido pela UPSE que iniciar o processo, cabendo à UPSE II o campo assinar e rubricar, nos campos próprios, a carga horária do professor

NOTA: Esta cota de concessão a UPSE I será aquela em que o professor ministrar ou passar a ministrar o maior número de aulas

TIPO DO MODELO.

QUANDO O PROFESSOR MINISTRAR AULAS EM UMA UPSE

1. UPSE I

1.1. Preencher assinar e rubricar em 4 vias

1.2. Rubricar a DE-RI a que se subordinar, a 1a, 2a e 3a vias remetendo

2. DE-RI

2.1. Rubricar, assinar e rubricar as vias recebidas da UPSE I

2.2. Rubricar a 3a via referente à 1a, 2a ou 3a DP

3. DEPARTAMENTO DE PESSOAL

3.1. Rubricar, assinar e rubricar as vias remetidas pelo DE-RI

3.2. Preencher o Expediente

3.3. Assinar a 1a e 2a vias e devolver à 1a do DE-RI

4. DE-RI

4.1. Rubricar a 2a via do DE-RI (Rubricar a 1a via numerando e a data de portaria)

4.2. Assinar a 2a via e remeter à 1a da UPSE I

5. UPSE II

5.1. Rubricar e rubricar a 1a via

5.2. Devolver a 1a via

QUANDO O PROFESSOR MINISTRAR AULAS EM DUAS UPSE'S

1. UPSE I

1.1. Preencher assinar e rubricar em 4 vias

1.2. Rubricar a 1a via referente à 1a, 2a ou 3a DP

2. UPSE II

2.1. Rubricar, complementar o preenchimento e assinar as vias recebidas da UPSE I

2.2. Rubricar as 1a vias remetidas ao DE-RI a que se subordinar

3. DE-RI

3.1. Rubricar, assinar e rubricar as 4 vias recebidas da UPSE I

3.2. Rubricar a 1a, 2a via ou 3a DP remetendo à 3a e 4a

4. DEPARTAMENTO DE PESSOAL

4.1. Rubricar, assinar e rubricar as vias remetidas pelo DE-RI

4.2. Preencher o Expediente

4.3. Assinar a 1a e 2a vias e devolver à 1a do DE-RI

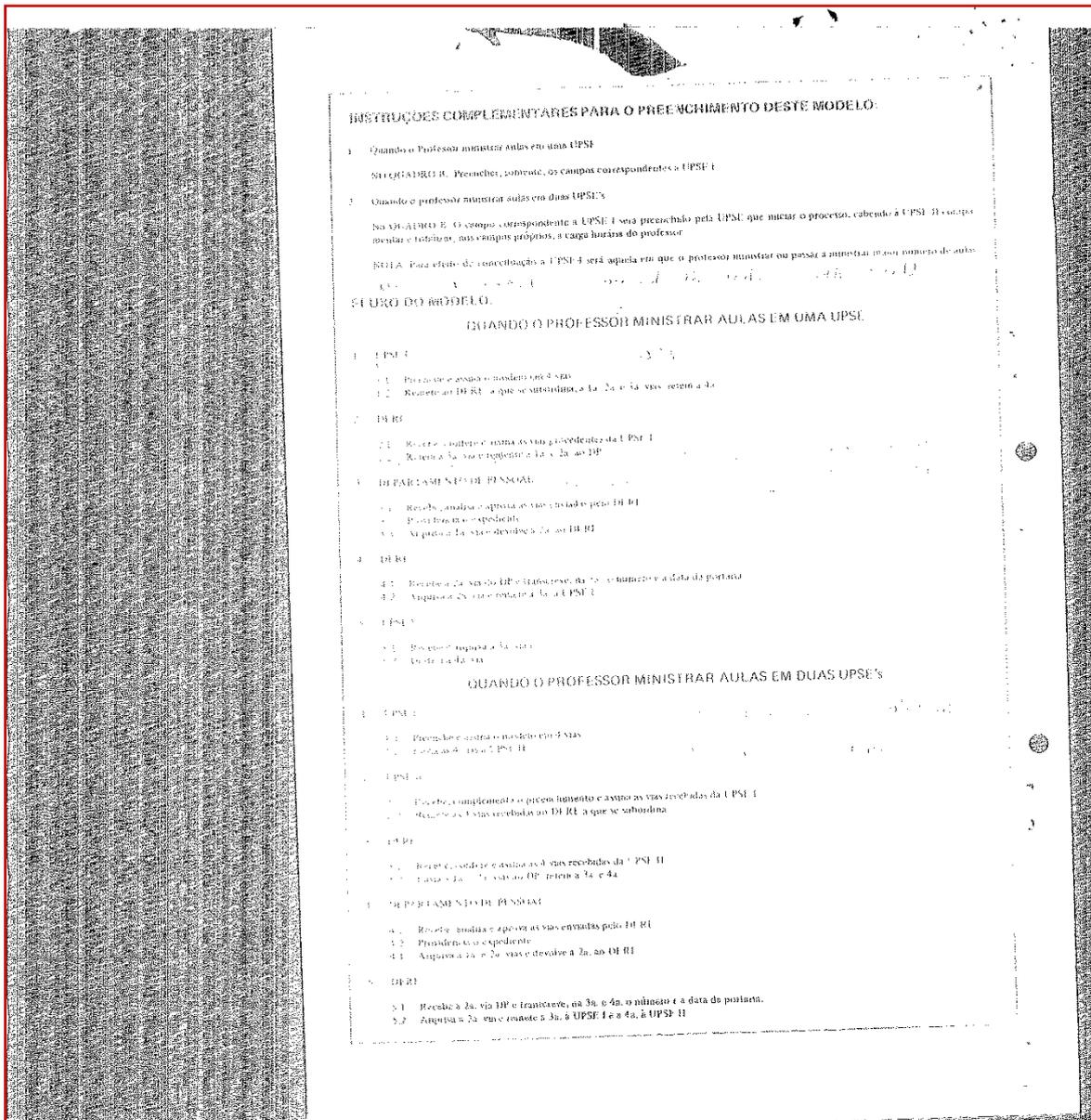
5. DE-RI

5.1. Rubricar a 2a via DP e remeter, na 1a e 3a, o primeiro e a data de portaria.

5.2. Assinar a 2a via e remeter à 3a da UPSE I e à 4a da UPSE II

Figura 59 - Imagem 0141 da Base1_1, resultado do Ávila-Lins. Borda ruidosa não foi removida.

Tempo de processamento 0,249909s. Tamanho 215KB.



**Figura 60 - Imagem 0141 da Base1_1, resultado do Formiga-Lins. Borda ruidosa não foi removida.
Tempo de processamento 0,303501s. Tamanho 215KB.**

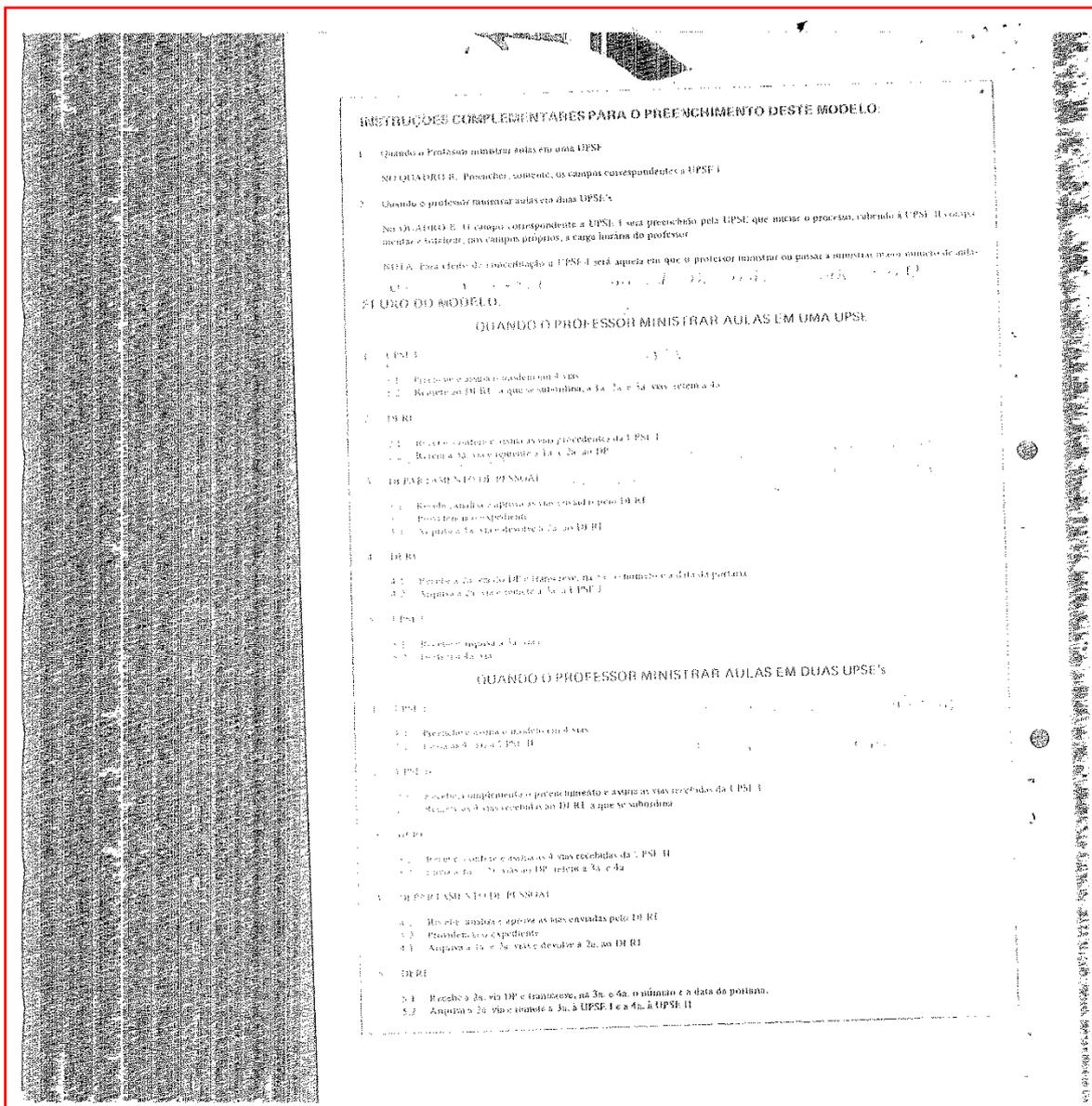


Figura 61 - Imagem 0141 da Base1_1, resultado da ferramenta comercial BlackIce. Borda ruidosa não foi removida. Tamanho 200KB.

INSTRUÇÕES COMPLEMENTARES PARA O PREENCHIMENTO DESTES MODELOS:

1. Quando o Professor ministrar aulas em uma UPSE
 NO QUADRO B, Preencher, somente, os campos correspondentes a UPSE I

2. Quando o professor ministrar aulas em duas UPSE's
 No QUADRO B, O campo correspondente a UPSE I será preenchido pela UPSE que iniciar o processo, cabendo à UPSE II o campo
 desta e totalizar, nos campos próprios, a carga horária do professor

NOTA: Para efeito de concentração a UPSE I será aquela em que o professor ministrar ou passar a ministrar maior número de aulas

QUADRO DO MODELO:

QUANDO O PROFESSOR MINISTRAR AULAS EM UMA UPSE

1. UPSE I

1.1. Preencher e assinar o modelo em 4 vias

1.2. Remeter as 4 vias expedidas à 1a e 2a vias para a

2. DE RI

2.1. Receber, conferir e assinar as vias recebidas da UPSE I

2.2. Remeter as 2 vias expedidas à 1a e 2a vias para a

3. DEPARTAMENTO DE PESSOAL

3.1. Receber, analisar e aprovar as vias expedidas pelo DE RI

3.2. Preencher o expediente

3.3. Arquivar a 1a via expedida à 2a do DE RI

4. DE RI

4.1. Receber a 2a via do DE e transcrever, na 3a, o número e a data da portaria

4.2. Arquivar a 2a via expedida à 1a a UPSE I

5. UPSE II

5.1. Receber a 2a via da 1a via

5.2. Preencher o modelo

QUANDO O PROFESSOR MINISTRAR AULAS EM DUAS UPSE'S

1. UPSE I

1.1. Preencher e assinar o modelo em 4 vias

1.2. Remeter as 4 vias para a UPSE II

2. UPSE II

2.1. Preencher o complemento e preencher o modelo e assinar as vias recebidas da UPSE I

2.2. Remeter as 4 vias expedidas ao DE RI a que se subordina

3. DE RI

3.1. Receber, conferir e assinar as 4 vias recebidas da UPSE II

3.2. Remeter as 2 vias ao DE RI para a 1a e 2a

4. DEPARTAMENTO DE PESSOAL

4.1. Receber, analisar e aprovar as vias expedidas pelo DE RI

4.2. Preencher o expediente

4.3. Arquivar a 1a e 2a vias expedidas à 2a do DE RI

5. DE RI

5.1. Receber a 2a via do DE e transcrever, na 3a, o número e a data da portaria

5.2. Arquivar a 2a via expedida à 2a, à UPSE I e a 4a, à UPSE II

Figura 62 - Imagem 0141 da Base1_1, resultado da ferramenta comercial ClearImage.

Borda ruidosa não foi removida. Tamanho 211KB.

INSTRUÇÕES COMPLEMENTARES PARA O PREENCHIMENTO DESTES MODELOS.

1. Quando o Professor ministrar aulas em uma UPSE:

NO QUADRO B. Pesquisar, somente, os campos correspondentes a UPSE I.

2. Quando o professor ministrar aulas em duas UPSE'S:

NO QUADRO B. O campo correspondente a UPSE I será preenchido pela UPSE que iniciar o processo, cabendo a UPSE II os complementares e totais, nos campos próprios, a carga horária do professor.

NOTA: Para efeito de identificação a UPSE I será aquela em que o professor iniciar ou passar a ministrar as aulas (início de aula).

QUANDO O PROFESSOR MINISTRAR AULAS EM UMA UPSE

1. CPSE I

1.1. Preencher campo correspondente em 3 vias.

1.2. Retornar ao DE RE a que se subordina, a 1a, 2a e 3a vias, respectivamente.

2. DE RE

2.1. Receber, conferir e assinar as vias recebidas da CPSE I.

2.2. Retornar a 2a via e registrar a 1a e 2a ao DP.

3. DEPARTAMENTO DE PSNOMIA

3.1. Receber, analisar e aprovar as vias recebidas pelo DE RE.

3.2. Preencher as respectivas.

3.3. Arquivar a 1a via e devolver a 2a ao DE RE.

4. DE RA

4.1. Receber a 2a via do DE e transcrever, na 1a via, o número e a data da portaria.

4.2. Arquivar a 2a via e registrar a 1a a CPSE I.

5. UPSE II

5.1. Receber e assinar a 1a via.

5.2. Devolver a 2a via.

QUANDO O PROFESSOR MINISTRAR AULAS EM DUAS UPSE'S

1. CPSE I

1.1. Preencher campo correspondente em 4 vias.

1.2. Retornar ao DE RE a que se subordina, a 1a, 2a e 3a vias, respectivamente.

2. CPSE II

2.1. Preencher o complemento o preenchimento e assinar as vias recebidas da CPSE I.

2.2. Retornar ao DE RE a que se subordina, a 1a, 2a e 3a vias, respectivamente.

3. DE RE

3.1. Receber e conferir e assinar as 4 vias recebidas da CPSE II.

3.2. Retornar a 2a via ao DP, respectivamente a 3a e 4a.

4. DEPARTAMENTO DE PSNOMIA

4.1. Receber, analisar e aprovar as vias recebidas pelo DE RE.

4.2. Preencher as respectivas.

4.3. Arquivar a 1a e 2a vias e devolver a 3a ao DE RE.

5. DE RE

5.1. Receber a 2a via do DP e transcrever, na 3a e 4a, o número e a data da portaria.

5.2. Arquivar a 2a via e registrar a 3a à UPSE I e a 4a à UPSE II.

Figura 63 - Imagem 0141 da Base1_1, resultado da ferramenta comercial LeadTools.

Borda ruidosa não foi removida. Tamanho 205KB.

INSTRUÇÕES COMPLEMENTARES PARA O PREENCHIMENTO DESTES MODELOS:

1. Quando o Professor ministrar aulas em uma UPSE:
 No QUADRO B, Preencher, conforme, os campos correspondentes a UPSE I.

2. Quando o professor ministrar aulas em duas UPSE'S:
 No QUADRO B, O campo correspondente a UPSE I será preenchido pelo UPSE que iniciar o processo, cabendo a UPSE II o campo medido e totalizar, nos campos próprios, a carga horária do professor.

NOTA: Para efeito de convocação a CPSE, terá prioridade em que o professor ministrará ou prestará ministrará o maior número de aulas.

FLUXO DO MODELO:

QUANDO O PROFESSOR MINISTRAR AULAS EM UMA UPSE:

1. CPSE I
 - 1.1. Preencher o campo "CARGA HORÁRIA" em 4 vias.
 - 1.2. Remeter as 4 vias recolhidas ao DE RI, a que se subordina, a 2a e 3a vias, retornar a 4a.
2. DE RI
 - 2.1. Receber e centralizar as vias recolhidas da CPSE I.
 - 2.2. Remeter a 3a via recolhida a 2a e 3a vias ao DE.
3. DEPARTAMENTO DE PESSOAL
 - 3.1. Receber, analisar e aprovar as vias enviadas pelo DE RI.
 - 3.2. Providenciar o expediente.
 - 3.3. Arquivar a 2a via e devolver a 2a ao DE RI.
4. DE RI
 - 4.1. Receber a 2a via do DE e transmitir, na 3a, o número e a data de portaria.
 - 4.2. Arquivar a 2a via e retornar a 2a a CPSE I.
5. CPSE I
 - 5.1. Receber o impasse a 2a via.
 - 5.2. Retornar a 4a via.

QUANDO O PROFESSOR MINISTRAR AULAS EM DUAS UPSE'S:

1. CPSE I
 - 1.1. Preencher o campo "CARGA HORÁRIA" em 4 vias.
 - 1.2. Retornar a 4a via a CPSE II.
2. CPSE II
 - 2.1. Preencher, complementar o preenchimento e assinar as vias recolhidas da CPSE I.
 - 2.2. Remeter as 4 vias recolhidas ao DE RI, a que se subordina.
3. DE RI
 - 3.1. Receber e centralizar as vias recolhidas da CPSE II.
 - 3.2. Remeter a 3a via recolhida ao DE, retornar a 3a e 4a.
4. DEPARTAMENTO DE PESSOAL
 - 4.1. Receber, analisar e aprovar as vias enviadas pelo DE RI.
 - 4.2. Providenciar o expediente.
 - 4.3. Arquivar a 2a e 3a vias e devolver a 2a ao DE RI.
5. DE RI
 - 5.1. Receber a 2a via do DE e transmitir, na 3a, o número e a data de portaria.
 - 5.2. Arquivar a 2a via e retornar a 3a à UPSE I e a 4a, à UPSE II.

Figura 64 - Imagem 0141 da Base1_1, resultado da ferramenta comercial ScanFix.

Borda ruidosa não foi removida. Tamanho 200KB.

INSTRUÇÕES COMPLEMENTARES PARA O PREENCHIMENTO DESTES MODELOS:

1 Quando o Professor ministrar aulas em uma UPSE
 NO QUADRO B. Preencher, somente, os campos correspondentes a UPSE I

2 Quando o professor ministrar aulas em duas UPSE's
 No QUADRO B. O campo correspondente a UPSE I será preenchido pela UPSE que iniciar o processo, cabendo à UPSE II o campo
 orientar e totalizar, nos campos próprios, a carga horária do professor

NOTA: Para efeito de concessão a UPSE 4 será aquela em que o professor ministrou ou passar a ministrar maior número de aulas

FLUXO DO MODELO.

QUANDO O PROFESSOR MINISTRAR AULAS EM UMA UPSE

1 - UPSE I

1.1 Preencher e assinar o modelo em 4 vias

1.2 Remeter ao DE-RI a que se subordinar, a 1a, 2a e 3a vias retidas a 4a

2 - DE-RI

2.1 Receber, analisar e assinar as vias recebidas da UPSE I

2.2 Remeter a 3a via e remeter a 1a e 2a ao DP

3 - DEPARTAMENTO DE PEDAGOGIA

3.1 Receber, analisar e assinar as vias enviadas pelo DE-RI

3.2 Providenciar o expediente

3.3 Arquivar a 1a via e devolver a 2a ao DE-RI

4 - DE-RI

4.1 Receber a 2a via do DP e transcrever, na 3a via, o número e a data da portaria

4.2 Arquivar a 2a via e remeter a 3a à UPSE I

5 - UPSE I

5.1 Receber a 3a via

5.2 Distribuir a 4a via

QUANDO O PROFESSOR MINISTRAR AULAS EM DUAS UPSE'S

1 - UPSE I

1.1 Preencher e assinar o modelo em 4 vias

1.2 Remeter ao DE-RI a que se subordinar, a 1a, 2a e 3a vias retidas a 4a

2 - UPSE II

2.1 Receber, complementar o preenchimento e assinar as vias recebidas da UPSE I

2.2 Remeter as 4 vias recebidas ao DE-RI a que se subordinar

3 - DE-RI

3.1 Receber, analisar e assinar as 4 vias recebidas da UPSE II

3.2 Remeter a 1a e 2a vias ao DP retidas a 3a e 4a

4 - DEPARTAMENTO DE PEDAGOGIA

4.1 Receber, analisar e assinar as vias enviadas pelo DE-RI

4.2 Providenciar o expediente

4.3 Arquivar a 1a e 2a vias e devolver a 2a ao DE-RI

5 - DE-RI

5.1 Receber a 2a via do DP e transcrever, na 3a, o número e a data da portaria.

5.2 Arquivar a 2a via e remeter a 3a à UPSE I e a 4a à UPSE II

Figura 65 - Imagem 0141 da Base1_1, resultado da ferramenta comercial Atalasoft.

Borda ruidosa não foi removida. Tamanho 94KB.

INSTRUÇÕES COMPLEMENTARES PARA O PREENCHIMENTO DESTES MODELOS:

1. Quando o Professor ministrar aulas em uma UPSE
 NO QUADRO B. Preencher, somente, os campos correspondentes a UPSE I

2. Quando o professor ministrar aulas em duas UPSE'S
 No QUADRO B. O campo correspondente a UPSE I será preenchido pela UPSE que iniciar o processo, cabendo a UPSE II complementar e totalizar, nos campos próprios, a carga horária do professor

NOTA: Para efeito de consecutividade a UPSE-I será aquela em que o professor ministrar ou passar a ministrar maior número de aulas

FLUXO DO MODELO.

QUANDO O PROFESSOR MINISTRAR AULAS EM UMA UPSE

1. UPSE I

1.1. Preenche e assina o modelo em 3 vias
 1.2. Remete ao DE RI a que se subordina, a 1a, 2a e 3a vias retém a 4a

2. DE RI

2.1. Recebe e confere e assina as vias procedentes da UPSE I
 2.2. Retém a 3a via e repete a 1a e 2a ao DP

3. DEPARTAMENTO DE PESSOAL

3.1. Recebe, analisa e aprova as vias enviadas pelo DE RI
 3.2. Providencia o expediente
 3.3. Arquiva a 1a via e devolve a 2a ao DE RI

4. DE RI

4.1. Recebe a 2a via do DP e transcreve, na 3a, o número e a data da portaria
 4.2. Arquiva a 2a via e remete a 3a à UPSE I

5. UPSE I

5.1. Recebe e insere a 3a via
 5.2. Destina a 4a via

QUANDO O PROFESSOR MINISTRAR AULAS EM DUAS UPSE'S

1. UPSE I

1.1. Preenche e assina o modelo em 3 vias
 1.2. Envia as 3 vias à UPSE II

2. UPSE II

2.1. Recebe, complementa o preenchimento e assina as vias recebidas da UPSE I
 2.2. Remete as 3 vias recebidas ao DE RI a que se subordina

3. DE RI

3.1. Recebe e confere e assina as 3 vias recebidas da UPSE II
 3.2. Envia a 1a e 2a vias ao DP, retém a 3a e 4a

4. DEPARTAMENTO DE PESSOAL

4.1. Recebe, analisa e aprova as vias enviadas pelo DE RI
 4.2. Providencia o expediente
 4.3. Arquiva a 1a e 2a vias e devolve a 2a ao DE RI

5. DE RI

5.1. Recebe a 2a via DP e transcreve, na 3a e 4a, o número e a data da portaria.
 5.2. Arquiva a 2a via e remete a 3a, à UPSE I e a 4a, à UPSE II

Figura 66 - Imagem 0141 da Base1_1, resultado do Novo Algoritmo.

Tempo de processamento 0,143480s. Tamanho 29,2KB.

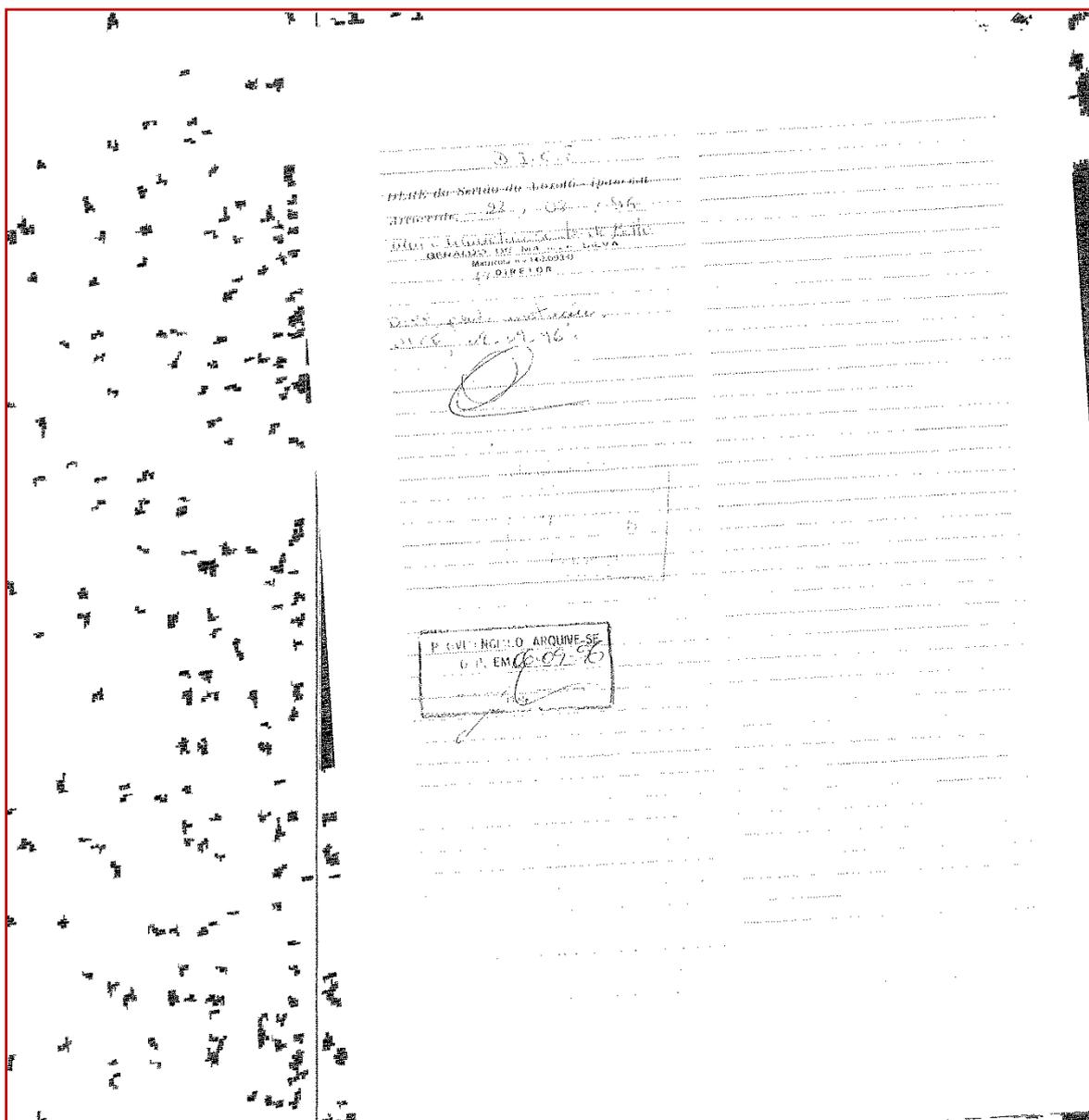


Figura 73 -0438 da Base1_1, resultado do Formiga-Lins. Borda ruidosa não foi completamente removida.

Tempo de processamento 0,806837s. Tamanho 30,6KB.

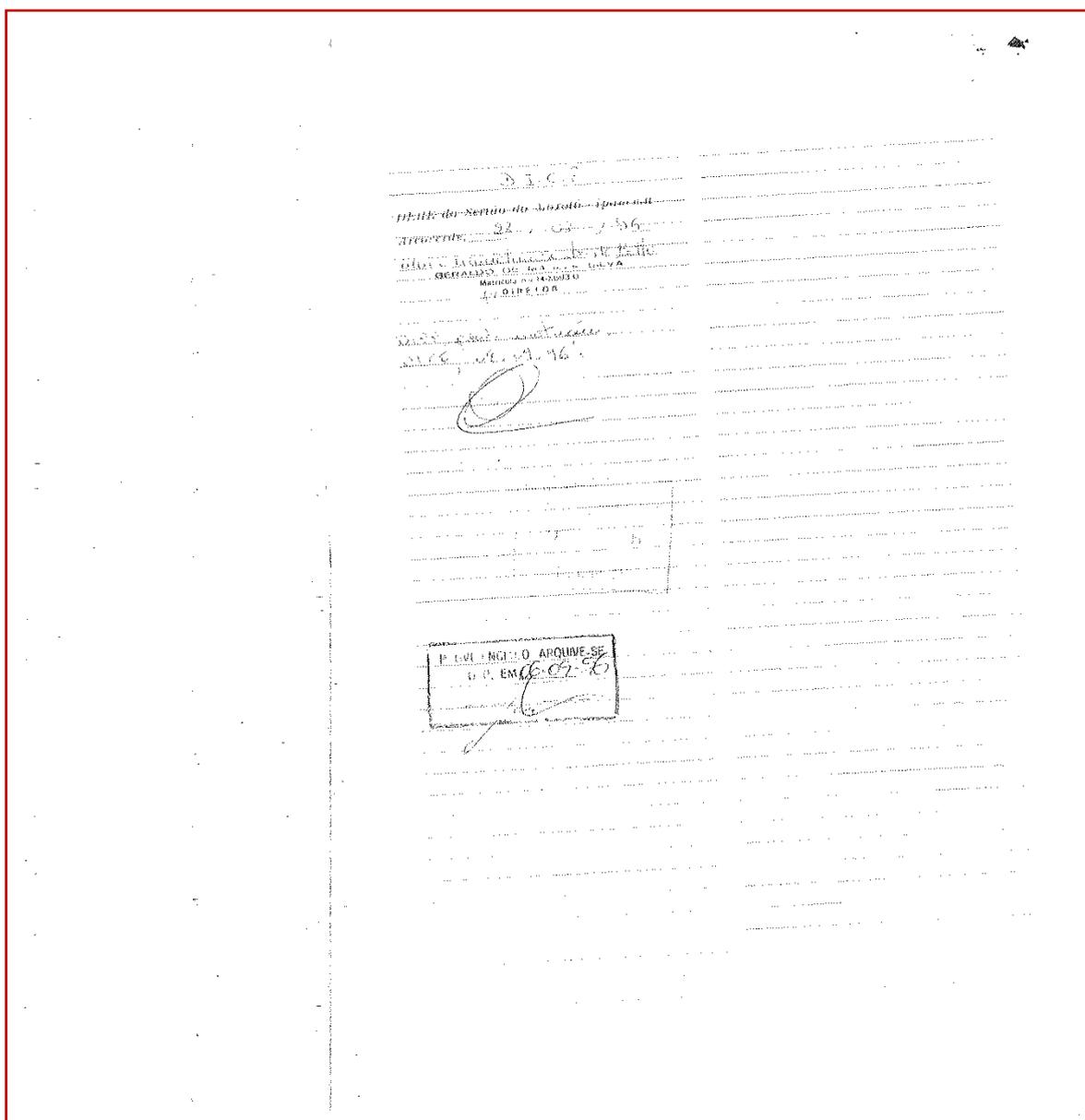


Figura 74 - Imagem 0438 da Base1_1, resultado da ferramenta comercial BlackIce.

Borda ruidosa foi removida. Não foi efetuado o *crop*. Tamanho 16,2KB.

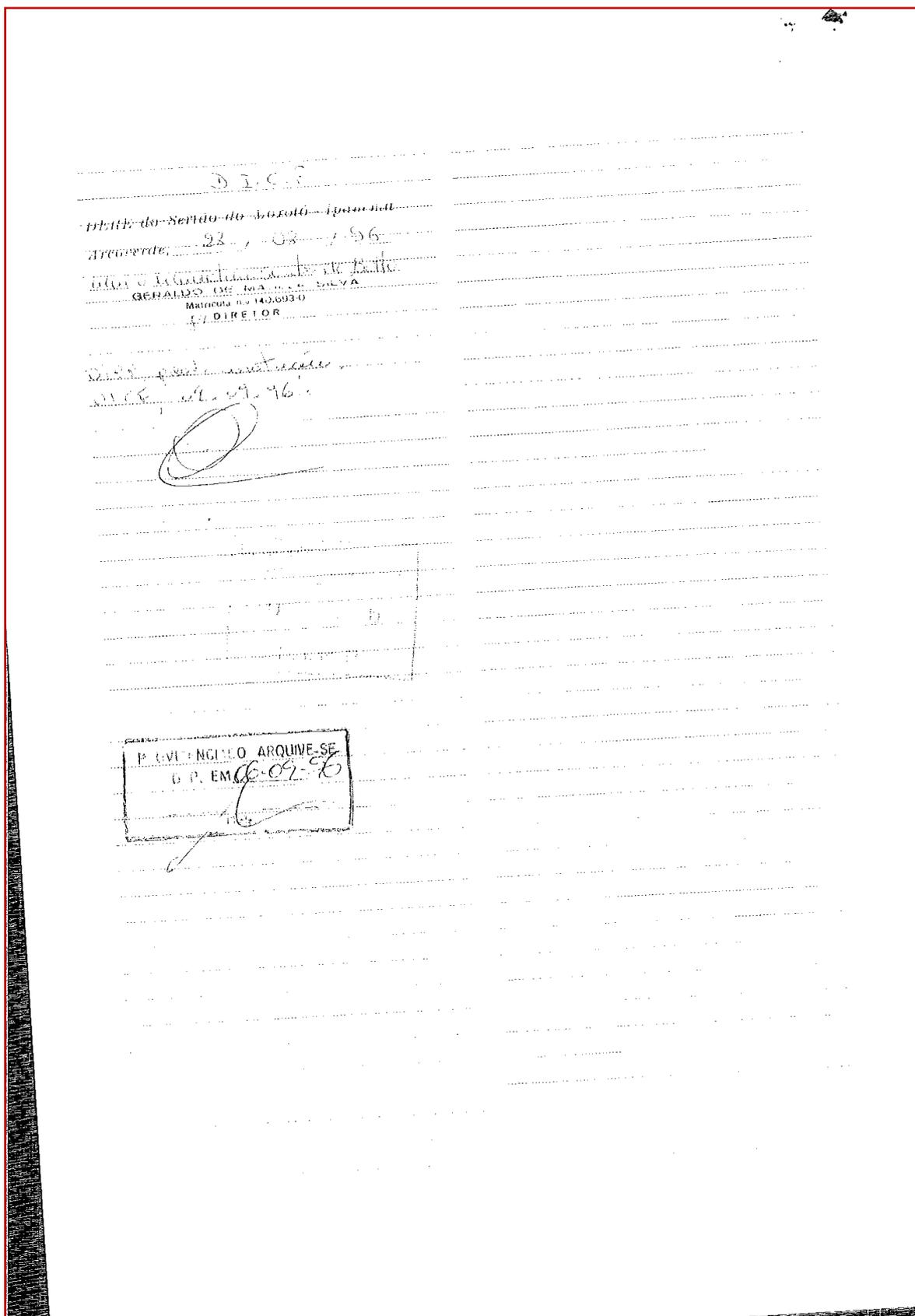


Figura 75 - Imagem 0438 da Base1_1, resultado da ferramenta comercial ClearImage.
Borda ruidosa removida e crop efetuado. Tamanho 21,4KB.

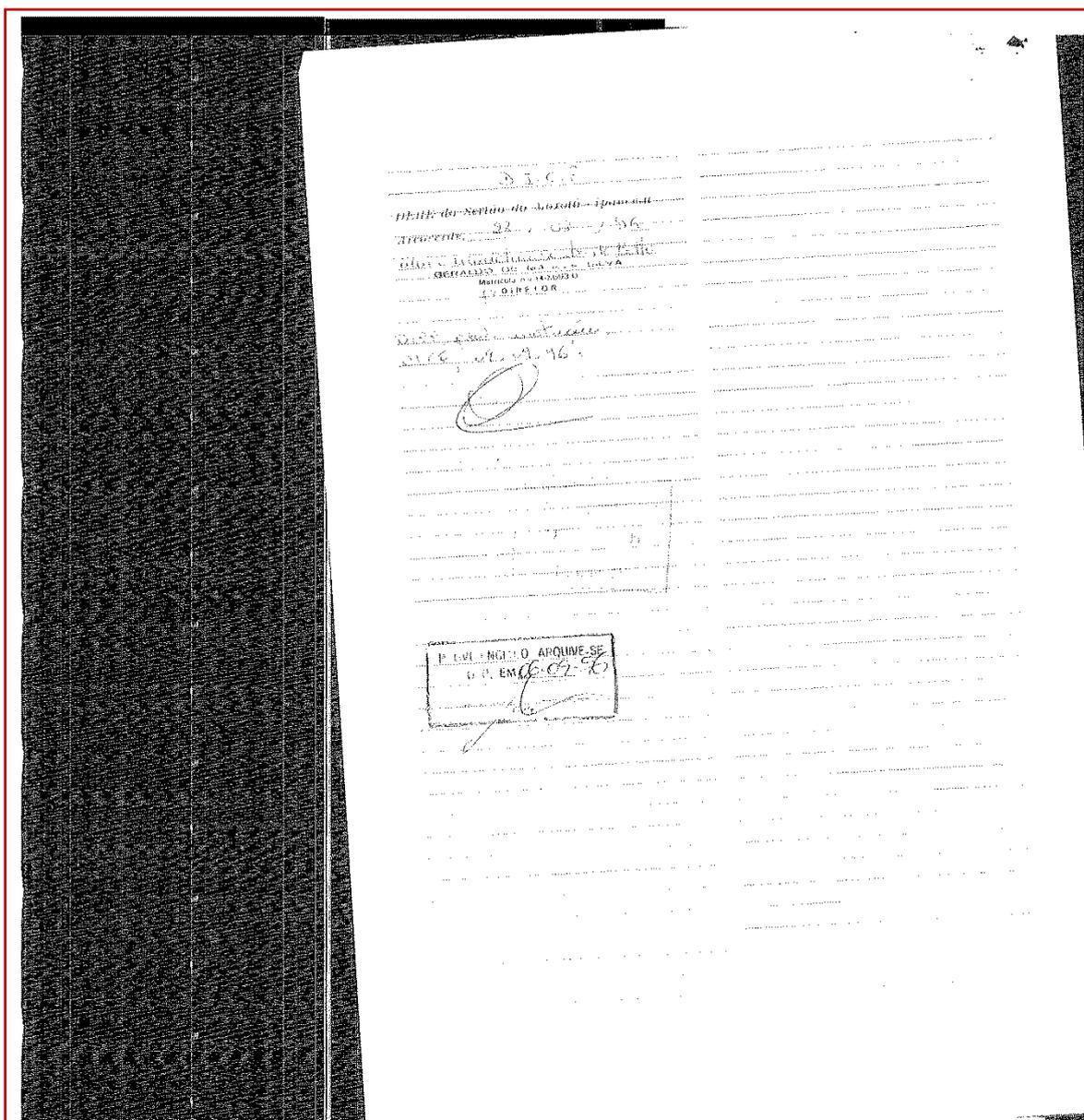


Figura 76 - Imagem 0438 da Base1_1, resultado da ferramenta comercial LeadTools.

Borda ruidosa não foi removida. Tamanho 147KB.

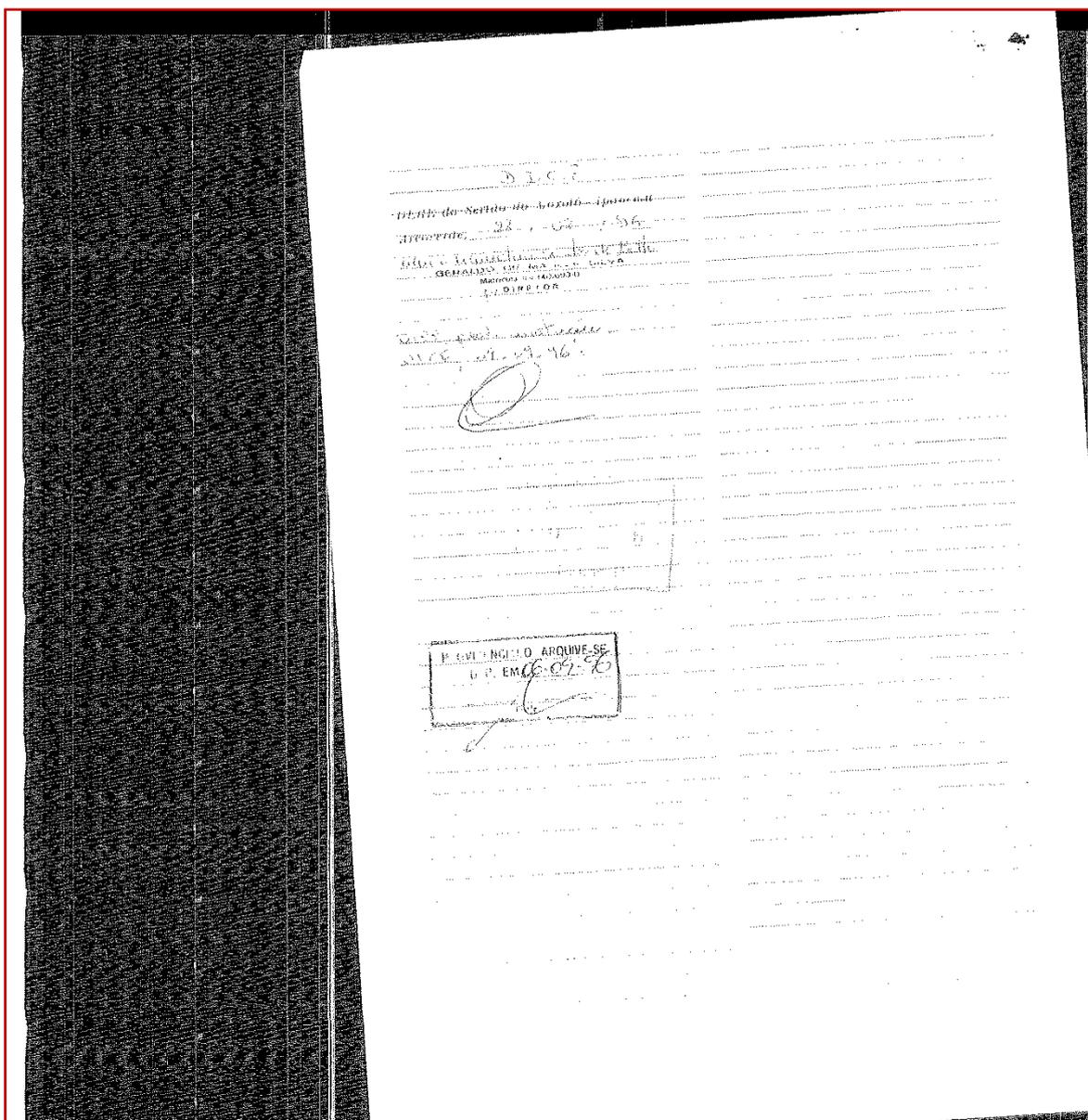


Figura 77 - Imagem 0438 da Base1_1, resultado da ferramenta comercial ScanFix.

Borda ruidosa não foi removida. Tamanho 150KB.

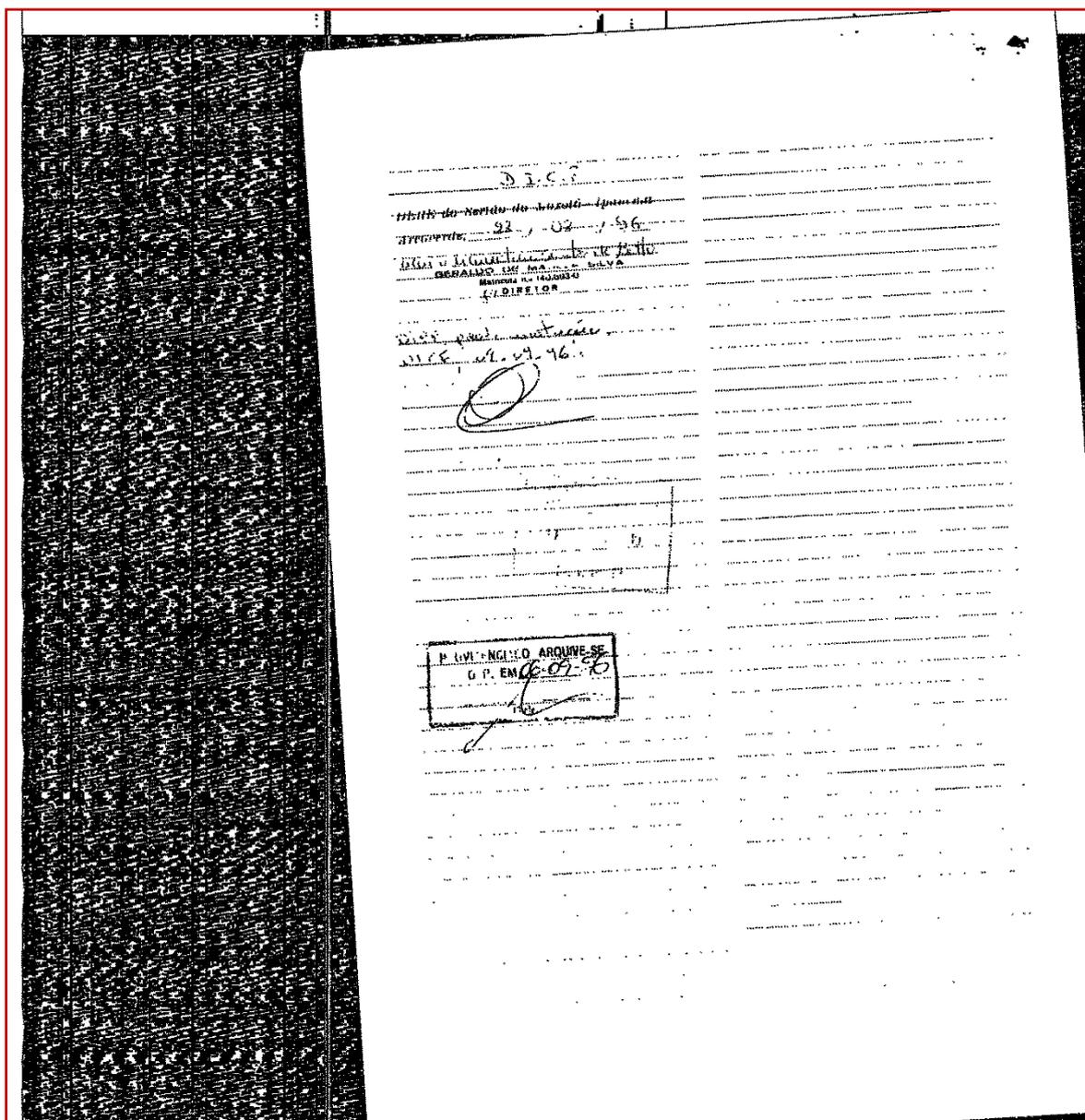


Figura 78 - Imagem 0438 da Base1_1, resultado da ferramenta comercial Atalasoft.
 Borda ruidosa não foi removida. Tamanho 84,5KB.

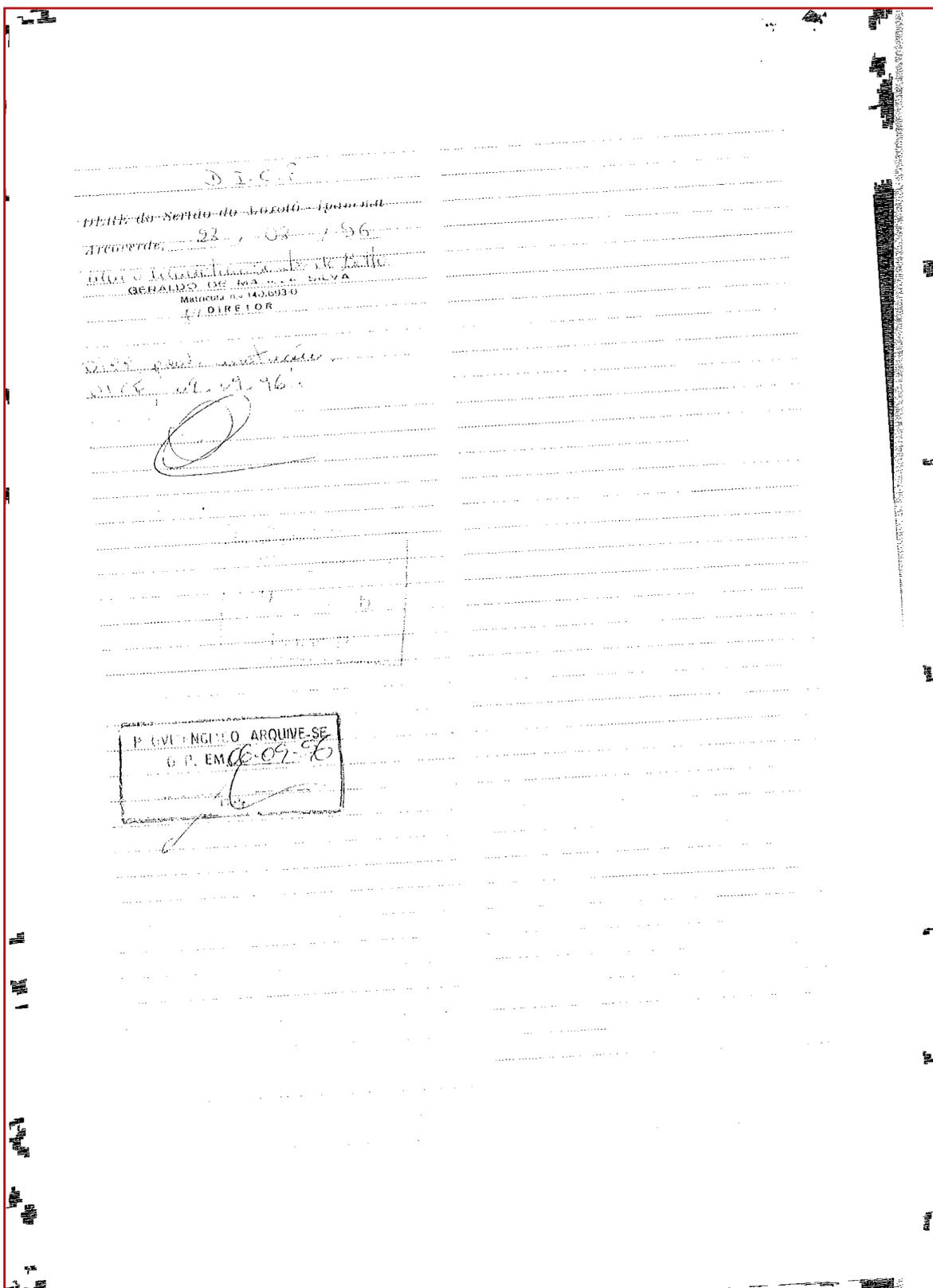


Figura 79 - Imagem 0438 da Base1_1, resultado do Novo Algoritmo. Borda ruidosa removida e *crop* efetuado. Tempo de processamento 0,196103s. Tamanho 21,8KB.

Estes são alguns resultados que demonstram a eficácia do Novo Algoritmo em imagens com bordas ruidosas. Tanto na remoção do ruído quanto na redução do tempo de processamento.

Há ainda casos, como o da imagem 3125 da BASE1_7, em que a remoção mais eficiente da borda ruidosa, com conseqüente maior redução da área do documento, proporcionou ao Ávila-Lins a possibilidade de remoção de ruído no interior do documento que antes não era efetuado.

Porém, houve casos onde os parâmetros para detecção da borda inferior não foram suficientes, ocasionando como resultado uma imagem com altura maior do que a área de interesse. Mas essas bordas inferiores ficaram completamente sem ruído. Possibilitando o uso de uma técnica simples e rápida para remoção de bordas brancas [02], não interferindo no tempo de processamento.

4.2.1 – Comparação com ferramentas comerciais

Foram efetuadas comparações com os resultados de alguns aplicativos para melhoria de imagem de documentos disponíveis no mercado em um pequeno grupo de imagens. As mesmas imagens consideradas nesta dissertação.

Não foi possível uma análise quantitativa destes softwares uma vez que são versões de demonstração (*shareware*). Não permitindo o uso para várias imagens em sequência automática, nem reproduzir as mesmas condições de controle utilizadas nos algoritmos Ávila-Lins, Formiga-Lins e o Novo Algoritmo.

As imagens utilizadas, os resultados para cada software e os arquivos de instalação dos softwares encontram-se na pasta “Comercial” no DVD em anexo.

Os softwares utilizados foram:

- **BlackIce Tiff Viewer 8.5, da BlackIce Software Inc [33].**

Obteve boa qualidade em seus resultados, mas não efetuou o recorte (*crop*) das imagens melhoradas (imagens 0002, 0047, 0125, 0229, 0268, 0310, 0438, 0583, 1253). Apresentou problemas na remoção de bordas muito ruidosas (imagem 0141) e causou perda de informação do documento (imagem 0229).

- **ClearImage 7. Inlite Research Inc.[34].**

Obteve os melhores resultados dos softwares comerciais testados. Porém não foi capaz de remover partes da borda ruidosa que tocam a informação do documento (imagens 0047, 0029). Apresentou problemas na remoção de bordas muito ruidosas (imagem 0141). Em documentos com bordas irregulares (imagem 0583), as irregularidades permaneceram bastante visíveis.

- **Leadtools 17.5, da Leadtools Inc. [35].**

Apresentou péssimos resultados, não conseguindo remover a borda ruidosa por completo em nenhuma das imagens, apenas parcialmente em algumas imagens (imagens 0125, 0310, 1253). Não efetuou o recorte (*crop*) em nenhuma das imagens.

- **ScanFix Xpress V6, da Accusoft Pegasus [36].**

Removeu a borda ruidosa e fez o recorte (*crop*) apenas nas imagens com documentos de bordas regulares e não rotacionadas (imagens 0002, 0047, 0125, 0229, 0310). As outras imagens, que apresentam documentos com bordas irregulares, rotação e borda muito ruidosa (imagens 0141, 0268, 0438, 0583, 1253), não foram alteradas.

- **Atalasoft DotImage 10 da Atalasoft [37].**

O pior resultado nesta comparação. Removeu a borda ruidosa em algumas das imagens, mas não fez o recorte (imagens 0002, 0047, 0125, 0229, 0310, 0583). A qualidade de todas as imagens foi fortemente prejudicada.

Como foi visto os algoritmos Ávila-Lins e Formiga-Lins não conseguiram remover as bordas ruidosas apenas nas imagens com forte ruído, e realizaram o recorte nas imagens em que a borda ruidosa foi removida. Não prejudicaram a qualidade de nenhuma imagem e não provocaram a perda de informações dos documentos.

O Novo Algoritmo conseguiu remover as bordas ruidosas que não foram removidas pelas ferramentas comerciais nem pelos algoritmos Ávila-Lins e Formiga-Lins.

5 – Conclusões e Trabalhos Futuros

Esta dissertação teve a finalidade de estudar e aprimorar alguns algoritmos para processamento e tratamento de documentos digitalizados monocromáticos. A principal motivação deste estudo é a necessidade de automação dos projetos de digitalização devido ao grande volume de documentos. Um novo algoritmo para remoção de bordas ruidosas foi proposto baseado no algoritmo morfológico de detecção de fronteiras com o objetivo de melhorar o desempenho do algoritmo Ávila-Lins. Reduzindo o tempo de processamento e aumentando o alcance e confiabilidade do Ávila-Lins. O novo algoritmo foi comparado com o próprio Ávila-Lins [2] e com o Formiga-Lins [1], algoritmo proposto com a mesma finalidade. Além da comparação qualitativa com alguns softwares comerciais para a remoção de bordas ruidosas (*document cleanup*) [33][34][35][36][37].

Foram efetuadas comparações com base em cinco mil imagens de testes, onde o tempo de processamento por imagem com o Novo Algoritmo, em comparação com o Ávila-Lins, foi reduzido em mais de 60%. E em comparação com o Formiga-Lins a redução verificada foi maior que 30%.

O alcance da nova proposta, quantidade de imagens que tiveram sua qualidade melhorada, também se mostrou maior, com a remoção de bordas ruidosas que não eram removidas por completo ou simplesmente eram ignoradas pelo Ávila-Lins, pelo Formiga-Lins e até por softwares comerciais.

A ideia de reduzir a resolução de leitura (analisar 8 pixels por vez ao invés de um por um) com a finalidade de diminuir o tempo de processamento mostrou-se um ideia válida que não compromete o resultado final, dados o tamanho e resolução das imagens analisadas. Esta ideia pode ser aplicada para imagens maiores, que comprometeriam o desempenho no tempo de processamento. Bastando aumentar o numero de pixels analisados de 8 para 12, 14 ou 16, dependendo do tamanho e resolução da imagem.

O Novo Algoritmo mostrou que é possível a remoção de ruídos na imagem do documento, como na imagem 3125 da BASE1_7 (Figura 45), através de modificações no algoritmo Ávila-Lins, antes designado apenas para remoção de bordas com textura sólida.

5.1 – Trabalhos Futuros

Este trabalho abriu muitas possibilidades de trabalhos futuros. O primeiro é a melhoria e expansão dos níveis de escalonamento de forma que não somente os parâmetros de varredura mudem com as dimensões do documento, mas também a extensão da faixa de pixels analisados seja modificada, aumentando proporcionalmente em relação às dimensões do documento analisado. Isso possibilitará a manutenção do desempenho no tempo de processamento mesmo para documentos de grandes dimensões.

Outra forma de melhorar o Novo Algoritmo é a partir da modificação da forma de varredura. Ao invés de seguir linha por linha na sequência de varredura no sentido vertical, as próximas 2 ou mais linhas na sequência seriam ignoradas, devido a alta probabilidade de ocasionarem a mesma falha de leitura da linha anterior.

Outra possibilidade de trabalho futuro é modificação do algoritmo Ávila-Lins de forma que possa ser aplicado, não apenas para a remoção do ruído de borda, mas também para a remoção de ruído na parte interna do documento. Através da aplicação do Ávila-Lins em subimagens da imagem original, ajustando os parâmetros de remoção de ruído para cada subimagem analisada.

6 – Referências

- 01 FORMIGA, A. A., LINS, R. D.: Efficient Removal of Noisy Borders of Monochromatic Documents. In: International Conference on Image Analysis and Recognition, 2009, Springer Verlag, 2009. v.5627. p.158 - 167
- 02 Ávila, B.T., Algoritmos e arquiteturas para processamento de documentos digitalizados monocromáticos, Dissertação de Mestrado em Engenharia Elétrica, Departamento de Eletrônica e Sistemas, Universidade Federal de Pernambuco, 2006.
- 03 Ávila, B.T., Lins, R.D.: Efficient removal of noisy borders from monochromatic documents. In: Campilho, A.C., Kamel, M.S. (eds.) ICIAR 2004. LNCS, vol. 3212, pp. 249–256. Springer, Heidelberg, 2004.
- 04 Ávila, B.T., Lins, R.D.: A New Algorithm for Removing Noisy Borders from Monochromatic Documents. In: ACM-SAC 2004, Março 2004, pp. 1219–1225, 2004.
- 05 Mello, C.A.B., Lins, R.D. Image Segmentation of Historical Documents. Visual 2000, México, agosto, 2000.
- 06 Lins, R.D., Machado, D.S.A. A comparative study of file formats for image storage and transmission. Journal of Electronic Imaging, v. 13(1), p. 175-183, 2004.
- 07 Shafait, F., Breuel, T.M.: A simple and effective approach for border noise removal from document images. In: Multitopic Conference, 2009. INMIC 2009. IEEE 13th International, Dezembro 2009.
- 08 Fan, K.C., Wang, Y.K., Lay, T.R.: Marginal noise removal of document images. Patt. Recog. 35, 2593–2611 ,2002.
- 09 Alves, N.F. Estratégias para melhoria do desempenho de ferramentas comerciais de reconhecimento óptico de caracteres. Dissertação de Mestrado em Engenharia Elétrica, Departamento de Eletrônica e Sistemas, Universidade Federal de Pernambuco, 2003.
- 10 Kanungo, T., Haralick, R.M., Phillips, I.: Global and local document degradation models. In: Proc. Snd Int. Conf. Doc. Analysis and Recognition, pp. 730–734, 1993.
- 11 O’Gorman, L., Kasturi, R. Document Image Analysis. IEEE Computer Society Executive Briefing, IEEE, 1997;
- 12 Lins, R.D., Ávila, B.T. BigBatch - A Toolbox for Monochromatic Documents. In: ACM International Conference on Document Engineering, ACM Press, Bristol, Inglaterra, 2005;
- 13 Lins, R.D., Ávila, B.T., Formiga, A.A. BigBatch: An Environment for Processing

- Monochromatic Documents. In: International Conference on Image Analysis and Recognition, v. 41142. p. 886-896, LNCS, Springer, Póvoa de Varzim, Portugal, 2006;
- 14 Shapiro, L.G., Stockman, G.C. Computer Vision, março, 2000 – <http://www.cse.msu.edu/~stockman/Book/book.html>;
 - 15 O’Gorman, L. The Document Spectrum for Page Layout Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 15(11), p. 1162-1173, 1993;
 - 16 R. Gonzalez, R. Woods.: Digital Image Processing, 3° Edition, Prentice Hall, 2008;
 - 17 O’Gorman, L., Kasturi, R.: Document Image Analysis, IEEE Computer Society Executive Briefing ,2009;
 - 18 Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to algorithms, 2nd edn. MIT Press, Cambridge (2001)
 - 19 Berger, M.: Computer Graphics with Pascal. Addison-Wesley, Reading ,1986
 - 20 Schildt, H., C completo e total, 3° Edition, Makron Books, 2006, pp. , pp. ,
 - 21 Scanner 1500 and Scanner 2500 - www.kodak.com/US/plugins/acrobat/en/business/scanners/Scanner1500and2500.pdf
 - 22 i1800Scanner - http://graphics.kodak.com/docimaging/uploadedfiles/en_i1800Scanner.pdf
 - 23 Cattoni, R., Coianiz, T., Messelodi, S., Modena, C.M. Geometric Layout Analysis Techniques for Document Image Understanding: a review. ITC_IRST, janeiro, 1998;
 - 24 BRACELPA - Associação Brasileira de Celulose e Papel.: Relatório Anual 2009/2010, pp. 27-37, pp. 38-41.
 - 25 TIFF, Revision 6.0 final, 1992 <http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf>,
 - 26 LibTIFF - TIFF Library and Utilities, <http://www.libtiff.org>, ultimo acesso em 01/11/2011 as 14:30.
 - 27 IBM - Graphics programming with libtiff Part 1 <https://www.ibm.com/developerworks/linux/library/l-libtiff/>, ultimo acesso em 01/11/2011;
 - 28 IBM - Graphics programming with libtiff, Part 2 <http://www.ibm.com/developerworks/linux/library/l-libtiff2/> ultimo acesso em 01/11/2011;
 - 29 The C++ Resources Network, <http://cplusplus.com>, ultimo acesso em 01/11/2011 as 13:10.

- 30 Environmental Paper Network - <http://www.environmentalpaper.org/> - ultimo acesso em 01/11/2011 as 15:10.
- 31 RISI - Essential information for the forest products industry.: <http://www.risiinfo.com/>, ultimo acesso em 01/11/2011.
- 32 Nou-Rau Software Livre - Rau-Tu - UNICAMP - Sistema de Arquivamento e Indexação de Documentos.: <http://www.rau-tu.unicamp.br/nou-rau/softwarelivre/>, ultimo acesso em 01/11/2011.
- 33 BlackIce Tiff Viewer 8.5. BlackIce Software Inc., <http://www.blackice.com/>, ultimo acesso em 01/12/2011.
- 34 ClearImage 7. Inlite Research Inc., <http://www.inliteresearch.com>, ultimo acesso em 01/12/2011.
- 35 Leadtools 17.5. Leadtools Inc., <http://www.leadtools.com>, ultimo acesso em 01/12/2011.
- 36 ScanFix Xpress V6. Accusoft Pegasus, <http://www.accusoft.com/index.htm>, ultimo acesso em 07/01/2012.
- 37 Atalasoft DotImage 10. Atalasoft, <http://www.atalasoft.com/>, ultimo acesso em 07/01/2012

APÊNDICE A - Código fonte dos algoritmos e programas

Os algoritmos implementados geraram muitas linhas de código o que, portanto, inviabilizou as impressões de listagens nesta dissertação. Um DVD com as implementações de cada Algoritmo, os arquivos para instalação dos softwares comerciais para *document cleanup*, a base de imagens original utilizada para testes e todas as imagens resultantes está em anexo a esta dissertação. Sendo organizado da seguinte maneira:

Diretório	Descrição
Ávila-Lins	Código fonte da implementação do algoritmo Ávila-Lins.
Formiga-Lins	Código fonte da implementação do algoritmo Formiga-Lins.
Novo Algoritmo	Código fonte da implementação do Novo Algoritmo.
Comercial	Arquivos de instalação dos softwares comerciais.
BASE1	Contém as imagens originais, separadas em 10 pastas (BASE1_1 a BASE1_10) com 500 arquivos cada.
Resultados Ávila-Lins	Contem as imagens resultantes do processamento pelo algoritmo Ávila-Lins.
Resultados Formiga-Lins	Contem as imagens resultantes do processamento pelo algoritmo Formiga-Lins.
Resultados Novo Algoritmo	Contem as imagens resultantes do processamento pelo Novo Algoritmo.
Resultados Comercial	Contem as imagens resultantes do processamento pelos softwares comerciais.

APÊNDICE B - Publicações

- [1] Mariano, Ednardo ; Lins, Rafael Dueire ; Silva, Gabriel de França Pereira e ; Fan, Jian ; Majewicz, Peter ; Thielo, Marcelo . Correcting Specular Noise in Multiple Images of Photographed Documents. In: ICDAR 2011 - International Conference on Document Analysis and Recognition, 2011, Pequim. ICDAR 2011. New York : IEEE Press, 2011. p. 111-116.