

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DEPARTAMENTO ACADÊMICO DE COMPUTAÇÃO
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

DOUGLAS VINICIUS DE ABREU

**RECONHECIMENTO DE ESTILOS DE ESCRITA EM
DOCUMENTOS MANUSCRITOS: UMA
ABORDAGEM EM NÍVEL DE DOCUMENTOS,
LINHAS E PALAVRAS.**

MONOGRAFIA

CAMPO MOURÃO

2019

DOUGLAS VINICIUS DE ABREU

**RECONHECIMENTO DE ESTILOS DE ESCRITA EM
DOCUMENTOS MANUSCRITOS: UMA
ABORDAGEM EM NÍVEL DE DOCUMENTOS,
LINHAS E PALAVRAS.**

Trabalho de Conclusão de Curso de graduação apresentado à disciplina de Trabalho de Conclusão de Curso 2, do Curso de Bacharelado em Ciência da Computação do Departamento Acadêmico de Computação da Universidade Tecnológica Federal do Paraná, como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

Orientador: Prof. Dr. Diego Bertolini Gonçalves

CAMPO MOURÃO

2019



ATA DE DEFESA DO TRABALHO DE CONCLUSÃO DE CURSO

Às **21:20** do dia **28 de novembro de 2019** foi realizada na sala **D103** da UTFPR-CM a sessão pública da defesa do Trabalho de Conclusão do Curso de Bacharelado em Ciência da Computação do(a) acadêmico(a) **Douglas Vinicius De Abreu** com o título **Reconhecimento de estilos de Escrita em Documentos Manuscritos: Uma Abordagem em Nível de Documentos, Linhas e Palavras**. Estavam presentes, além do(a) acadêmico(a), os membros da banca examinadora composta por: **Prof. Dr. Diego Bertolini Gonçalves** (orientador(a)), **Prof. Dr. Andre Luiz Satoshi Kawamoto** e **Profa. Dra. Aretha Barbosa Alencar**. Inicialmente, o(a) acadêmico(a) fez a apresentação do seu trabalho, sendo, em seguida, arguido(a) pela banca examinadora. Após as arguições, sem a presença do(a) acadêmico(a), a banca examinadora o(a) considerou _____ na disciplina de Trabalho de Conclusão de Curso 2 e atribuiu, em consenso, a nota _____ (_____). Esse resultado foi comunicado ao (à) acadêmico(a) e aos presentes na sessão pública. A banca examinadora também comunicou ao (à) acadêmico(a) que este resultado fica condicionado à entrega da versão final dentro dos padrões e da documentação exigida pela UTFPR ao professor responsável do TCC no prazo de **onze dias**. Em seguida foi encerrada a sessão e, para constar, foi lavrada a presente Ata que segue assinada pelos membros da banca examinadora, após lida e considerada conforme.

Observações:

Campo Mourão, **28 de novembro de 2019**

Prof. Dr. Andre Luiz Satoshi Kawamoto
Membro 1

Profa. Dra. Aretha Barbosa Alencar
Membro 2

Prof. Dr. Diego Bertolini Gonçalves
Orientador

A ata de defesa assinada encontra-se na coordenação do curso.

Resumo

Abreu, Douglas Vinicius. Reconhecimento de Estilos de Escrita em Documentos Manuscritos: Uma abordagem em nível de Documentos, Linhas e Palavras. . 2019. 41. f. Monografia (Curso de Bacharelado em Ciência da Computação), Universidade Tecnológica Federal do Paraná. Campo Mourão, 2019.

O reconhecimento de estilos de escrita ou reconhecimento de *script* é uma etapa necessária para automatizar o reconhecimento de sistemas de Reconhecimento Óptico de Caracteres (OCR). A tarefa de reconhecimento de estilos de escrita em documentos manuscritos é desafiadora devido à similaridade entre alguns estilos de escritas. Nosso principal objetivo neste trabalho é analisar e avaliar o desempenho de técnicas robustas em onze diferentes estilos de escrita. Além do desempenho, analisamos o impacto ao se trabalhar com reconhecimento em nível de documentos, linhas e palavras. Foram utilizados os descritores de texturas Padrões Binários Locais (LBP), Quantização Local de Fase (LPQ) e Características Robustas Aceleradas (SURF) em conjunto com o classificador Máquina de Vetores de Suporte (SVM). Os experimentos foram realizados em uma base de dados com onze classes, de forma a alcançarmos taxas de acerto similares às apresentadas na literatura. A melhor taxa de acerto foi de 97,73% utilizando o descritor de textura SURF trabalhando a nível de documento, nos níveis de linhas e palavras o melhor desempenho foi de 96% e 94,37%, respectivamente, utilizando o descritor LPQ. Por meio dos experimentos realizados, concluímos que o melhor desempenho se aplica ao nível de documentos seguido por linhas e palavras. Observamos também que a melhora no desempenho de aplicações de reconhecimento de estilo de escritas não está relacionado somente com a quantidade de amostras utilizadas no processo de treinamento.

Palavras-chaves: Reconhecimento Óptico de Caracteres. Máquina de vetores de Suporte. Características Robustas Aceleradas. Quantização Local de Fase. Padrões Binários Locais. Estilos de escritas

Abstract

Abreu, Douglas Vinicius. Script Recognition from handwritten documents : document, line and word-level approach. 2019. 41. f. Monograph (Undergraduate Program in Computer Science), Federal University of Technology – Paraná. Campo Mourão, PR, Brazil, 2019.

Script recognition is a necessary step in automating the recognition of Optical Character Recognition (OCR) systems. The task of script recognition in handwritten documents is challenging due to the similarity between some of these styles. Our main objective in this paper is to analyze and evaluate the performance of robust techniques in eleven different script recognition. In addition to performance, we analyzed the impact of working with document, line, and word-level recognition. The Local Binary Patterns (LBP), Local Phase Quantization (LPQ), and Speeded Up Robust Features (SURF) texture descriptors were used in conjunction with the Support Vector Machine (SVM) classifier. The experiments were performed in a database with eleven classes, in order to achieve hit rates similar to those presented in the literature. The best hit rate was 97.73% using the document-level SURF texture descriptor. In line and word levels the best performance was 96% and 94.37%, respectively, using the LPQ descriptor. Through our experiments, we find that the best performance applies to the document level, followed by lines and words. We also noted that the performance improvement of handwriting recognition applications is not only related to the amount of samples used in the training process.

Keywords: Optical Character Recognition. Support Vector Machine. Local Binary Patterns. Local Phase Quantization. Speeded Up Robust Features. Script recognition

Lista de figuras

1.1	Modelo para reconhecimento de padrões. Adaptada de Passos (2019)	9
1.2	Consolantes manuscritas de algumas formas de escritas regionais da Índia. Adaptada de Pal et al. (2012)	11
1.3	Textos manuscritos dos principais idiomas regionais da Índia. Adaptada de (PAL et al., 2012)	12
1.4	Evolução das formas de escritas indianas	12
2.1	Exemplo de operação básica do Padrões Binários Locais (LBP) com valor resultante 184. adaptada de (AMARAL, 2011)	16
2.2	Representação gráfica do descritor <i>Speeded Up Robust Features</i> (SURF). . .	18
2.3	Ideia geral da técnica <i>Support-vector Machine</i> (SVM) com os vetores de suporte circulados. Retirado de (GONÇALVEZ, 2008).	18
4.1	Visão geral do método proposto	25
4.2	Etapa de Extração de Características	27
4.3	Resumo do processo de segmentação da imagem em palavras.	28
4.4	Reconhecimento do estilo de escrita (Abordagem I)	28
5.1	Resultado dos experimentos em nível de Páginas	34

Lista de tabelas

1.1	Relação entre Idiomas e Estilo de Escrita. Adaptada de (OBAIDULLAH et al., 2018)	13
3.1	Resumo da Revisão Bibliográfica	23
4.1	Distribuição dos dados na base PHDIndic_11	26
5.1	Taxa de Reconhecimento (%) para nível de Páginas - Experimento 1	32
5.2	Taxa de Reconhecimento (%) para nível de Páginas - Experimento 2	32
5.3	Taxa de Reconhecimento (%) para nível de Páginas - Experimento 3	33
5.4	Taxa de Reconhecimento (%) para nível de Linhas - Experimento 1	34
5.5	Taxa de Reconhecimento (%) para nível de Linhas - Experimento 2	34
5.6	Taxa de Reconhecimento (%) para nível de Palavras - Experimento 1	35
5.7	Taxa de Reconhecimento (%) para nível de Palavras - Experimento 2	35
5.8	Estudo comparativo	36

Siglas

CNN:	<i>Convolutional Neural Network</i>
KNN:	K Vizinhos mais Próximos
LBP:	Padrões Binários Locais
LPQ:	Quantização Local de Fase
MLP:	<i>Multi-layer Perceptron</i>
OCR:	<i>Optical Character Recognition</i>
RBF:	<i>Radial Basis Function</i>
RF:	<i>Random forest</i>
SIFT:	<i>Scale-Invariant Feature Transform</i>
STFT:	<i>Short-Term Fourier Transform</i>
SURF:	<i>Speeded Up Robust Features</i>
SVM:	<i>Support-vector Machine</i>

Sumário

1	Introdução	9
1.1	Análise de Documentos em Imagens	10
1.2	Problemática	11
1.3	Objetivos e Questões de Pesquisa	13
1.4	Organização do Texto	14
2	Referencial Teórico	15
2.1	Descritores de Textura	15
2.1.1	Padrões Binários Locais (LBP)	15
2.1.2	Quantização Local de Fase (LPQ)	16
2.1.3	Características Robustas Aceleradas (SURF)	17
2.2	Máquina de Vetores de Suporte (SVM)	17
2.3	Considerações Finais	18
3	Trabalhos Relacionados	19
3.1	Discussão	21
4	Método	25
4.1	Base de Dados	25
4.2	Segmentação	26
4.3	Extração de Características	28
4.4	Classificação	29
5	Experimentos e Resultados	31
5.1	Reconhecimento a Nível de Documentos	32
5.2	Reconhecimento a Nível de Linha	33
5.3	Reconhecimento a Nível de Palavra	34
5.4	Considerações Finais	35
6	Conclusões	37
6.1	Contribuição	37
6.2	Trabalhos Futuros	38

Introdução

Segundo Theodoridis e Koutroubas (2008), reconhecimento de padrões é uma área de pesquisa que tem por objetivo a classificação de objetos (padrões) em categorias ou classes. O reconhecimento de padrões visa construir uma representação mais simples de um conjunto de dados considerando suas características mais relevantes possibilitando sua partição em classes (DUDA et al., 2000). A Figura 1.1 apresenta as principais etapas de um sistema de reconhecimento de padrões.

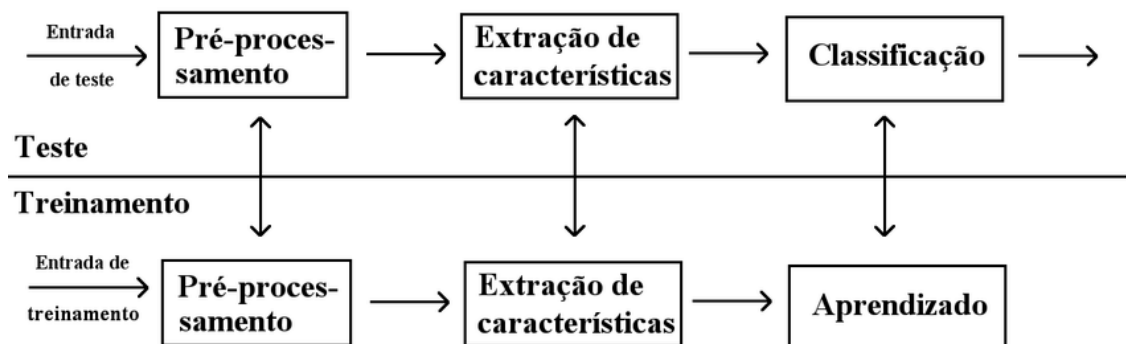


Figura 1.1. Modelo para reconhecimento de padrões. Adaptada de Passos (2019)

As técnicas de reconhecimento de padrões podem ser aplicadas em vários tipos de problemas, como: Reconhecimento de Caracteres, Reconhecimento de Assinaturas, Reconhecimento de Faces dentre outros (FAUNDEZ-ZANUY, 2005), (CAMPOS et al., 2009). Neste trabalho nosso principal objetivo é classificar documentos de texto manuscrito quanto ao seu estilo de escrita. Essa tarefa está inserida na subárea denominada análise de documentos em imagens.

1.1. Análise de Documentos em Imagens

A análise de documentos em imagens é o processo que realiza a interpretação de imagens de documentos, aplicando técnicas e algoritmos a imagens digitais de documentos (composta por *pixels*) para obter uma descrição legível para o computador (SALVI, 2014)

O reconhecimento de caracteres pode ser dividido em duas grandes áreas, texto manuscrito e texto datilografado ou impresso. Textos manuscritos normalmente implicam em uma maior dificuldade no reconhecimento de caracteres por possuírem uma maior variação (SALVI, 2014), devido a diversos fatores que vão desde o tipo da caneta e do papel até fatores como a habilidade, estilo e estado psicológico do escritor.

Segundo Obaidullah et al. (2019), a aquisição de manuscritos divide-se em duas formas: *online* a qual reconhecimento do manuscrito utiliza dispositivos especiais como mesa digitalizadoras, *tablet* entre outros equipamentos para capturar o movimento da caneta e registrar informações temporais; e o reconhecimento *offline*, que envolve apenas o texto manuscrito em papel e posteriormente a sua digitalização. O presente trabalho utiliza o reconhecimento de estilos de escrita *offline*.

Reconhecimento de caracteres são atividades desenvolvidas principalmente para sistemas *Optical Character Recognition* (OCR). Segundo Govindan e Shivaprasad (1990) sistemas OCR começaram a se popularizar nos anos de 1940 quando eram utilizados para atividades de transcrições de formulários comerciais, registros governamentais, entre outros tipos de documentos. O sucesso desses sistemas depende de diversos fatores entre eles deve-se destacar:

- Qualidade e resolução das imagens;
- Detecção do idioma/estilo de escrita do documento.

Pelo fato da identificação do estilo de escrita ser um processo que está totalmente vinculado ao sucesso dos sistemas OCR, é de extrema importância uma etapa de definição do estilo de escrita. Em ambientes que possuem uma grande quantidade de documentos é inviável a identificação de estilos de escrita/idiomas manualmente, assim, um sistema de identificação de estilos de escrita é totalmente necessário para esse tipo de atividade.

Alguns países apresentam uma variedade de estilos de escritas oficiais, desenvolver sistemas de identificação de estilos de escritas para esses países é uma tarefa desafiadora (PAL et al., 2012).

A Figura 1.2 contém alguns caracteres manuscritos de formas de escritas amplamente utilizados na Índia.

Phonation	Bangla	Oriya	Gujarati	Gurumukhi	Kannada	Telugu	Tamil	Malayalam	Phonation	Bangla	Oriya	Gujarati	Gurumukhi	Kannada	Telugu	Tamil	Malayalam
k	ক	କ	ક	क	ಕ	క	க	ക	kh	খ	ଖ	ક	ख	ಕ	క	க	ക
kh	খ	ଖ	ક	ख	ಕ	క	க	ക	g	গ	ଘ	ગ	ग	ಗ	గ	க	ക
g	গ	ଘ	ଗ	ग	ಗ	ಗ	க	ക	gh	ঘ	ଘ	ઘ	घ	ಗ	గ	க	ക
gh	ঘ	ଘ	ઘ	घ	ಗ	గ	க	ക	na	না	ନା	ના	ना	ನ	న	க	ക
na	না	ନା	ના	ना	ನ	న	க	ക	c	চ	ଢ	ઢ	ड	ಕ	చ	க	ക
c	চ	ଢ	ఢ	ड	ಕ	చ	க	ക	ch	ছ	ଢ	ఢ	ड	ಕ	చ	க	ക
ch	ছ	ଢ	ఢ	ड	ಕ	చ	க	ക	j	জ	ଞ	ઞ	ज	ಕ	చ	க	ക
j	জ	ଞ	ఞ	ज	ಕ	చ	க	ക	jh	ঝ	ଞ	ఞ	ज	ಕ	చ	க	ക
jh	ঝ	ଞ	ఞ	ज	ಕ	చ	க	ക	jn	ঞ	ଞ	ఞ	ज	ಕ	చ	க	ക
jn	ঞ	ଞ	ఞ	ज	ಕ	చ	க	ക	t	ট	ଠ	ఠ	ट	ಕ	చ	க	ക
t	ট	ଠ	ఠ	ट	ಕ	చ	க	ക	th	ঠ	ଠ	ఠ	ट	ಕ	చ	க	ക
th	ঠ	ଠ	ఠ	ट	ಕ	చ	க	ക	d	ড	ଢ	ఢ	ड	ಕ	చ	க	ക
d	ড	ଢ	ఢ	ड	ಕ	చ	க	ക	dh	ঢ	ଢ	ఢ	ड	ಕ	చ	க	ക
dh	ঢ	ଢ	ఢ	ड	ಕ	చ	க	ക	ne	নে	ନେ	ને	ने	ನ	న	க	ക
ne	নে	ନେ	ને	ने	ನ	న	க	ക	t	ত	ଠ	ఠ	ट	ಕ	చ	க	ക
t	ত	ଠ	ఠ	ट	ಕ	చ	க	ക	th	থ	ଠ	ఠ	ट	ಕ	చ	க	ക
th	থ	ଠ	ఠ	ट	ಕ	చ	க	ക	d	দ	ଢ	ఢ	ड	ಕ	చ	க	ക

Figura 1.2. Consolantes manuscritas de algumas formas de escritas regionais da Índia. Adaptada de Pal et al. (2012)

1.2. Problemática

A escrita pode ser definida como um sistema codificado de símbolos padrões baseados em um conjunto de regras, para representar ideias (PAL et al., 2012). Um estilo de escrita pode ser usada por apenas um idioma ou pode ser compartilhado por muitas línguas (SINGH et al., 2015a). A Figura 1.3 ilustra a frase do idioma inglês “One Hundred Rupees” descrito em vários idiomas diferentes. Pode-se observar que os idiomas Assamese e Bengali compartilham o mesmo estilo de escrita. Também é de se perceber que embora seja o mesmo estilo de escrita a frase formada é diferente, isso significa que o estilo de escrita Bengali é compartilhado entre diferentes idiomas (Bengali, Assamese, Manipuri).

Atualmente a Índia possui 22 idiomas oficiais os quais são descritos por 11 formas de diferentes estilos de escrita (OBAIDULLAH et al., 2019). A Tabela 1.1 mostra a relação entre os 22 idiomas oficiais da Índia com seu estilo de escrita juntamente com o número de escritores daquele idioma.

O sucesso para o reconhecimento de caracteres depende de um reconhecimento prévio do estilo de escrita/idioma em questão (HOCHBERG et al., 1999). Então, para a transcrição de documentos é de extrema importância um pré-processamento do documento ou parte dele para que seja identificado qual o estilo de escrita.

Trabalhar com estilos de escritas utilizados na Índia é um tema desafiador pelo fato da similaridade entre algumas dessas formas de escrita. Muita dessa similaridade devem-se

গুণ টকা	Assamese
প্রকণ টকা	Bengali
એકસો રૂચિયા	Gujarati
ಒಂದು ಸುರುರೂಪಿಯಾಗಳ	Kannada
നൂറ രൂപ	Malayalam
ଏକ ଶତକା	Oriya
ਇਕ ਸੈ ਰੁਪਏ	Punjabi
நூறு ரூபாய்	Tamil
నూరు రూపాయలు	Telugu
سرو روپی	Urdu

Figura 1.3. Textos manuscritos dos principais idiomas regionais da Índia. Adaptada de (PAL et al., 2012)

ao fato de que a maioria das escritas indianas serem derivadas de um único estilo de escrita como ilustrado na Figura 1.4.

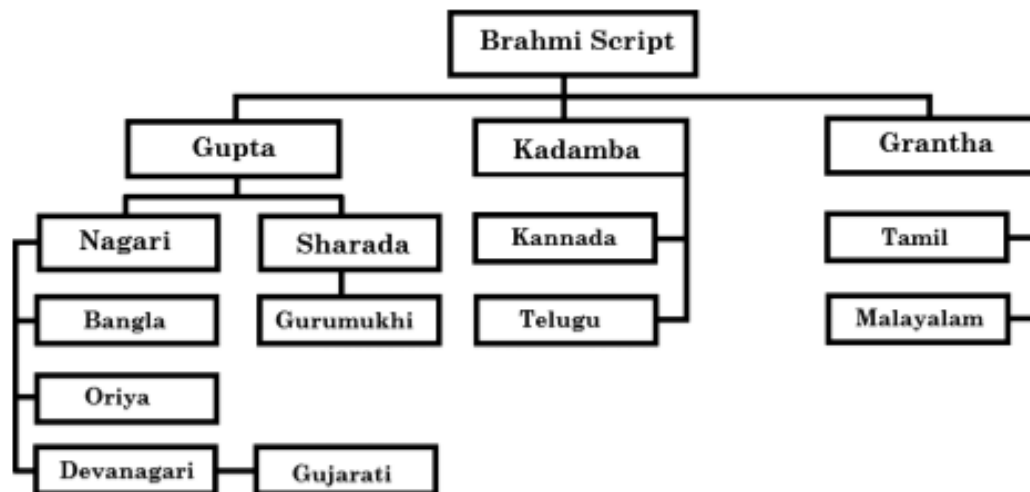


Figura 1.4. Evolução das formas de escritas indianas

Pode-se destacar ainda, que por se tratar de documentos manuscritos pode existir uma diferença muito grande entre amostras do mesmo estilo de escrita, ou seja, os caracteres ou palavras de cada estilo de escrita podem variar muito de escritor para escritor.

Tabela 1.1. Relação entre Idiomas e Estilo de Escrita. Adaptada de (OBAIDULLAH et al., 2018)

Idioma	Estilo de escrita	População (M)
Bodo	Devanágari	328,23
Konkani		
Marathi		
Maithili		
Nepali		
Sanskrit		
Sindhi		
Hindi		
Bangla	Bengali	211,5
Assamese		
Manipuri		
Kashmiri	Gujarati	52,1
Gujarati		
Telugu	Telugu	69,80
Tamil	Tamil/Perso-arabic	65,70
Urdu	Urdu	60,60
Malayalam	Malayalam	35,90
Oriya	Oriya	31,70
Kannada	Kannada	3,63
Punjabi	Gurumukhi	1,05
Dogri	Gurumukhi/Devanagari	3,80
Santali	Roman	6,20

1.3. Objetivos e Questões de Pesquisa

Esse trabalho tem como principal objetivo realizar uma análise para o reconhecimento *offline* de estilos de escrita. Para isso serão considerados onze estilos de escrita amplamente utilizados na Índia para descrever seus idiomas. Este objetivo principal está atrelado com os seguintes objetivos secundários:

- Verificar o desempenho dos descritores de textura LBP, Quantização Local de Fase (LPQ), SURF empregando o classificador SVM;
- Empregar técnicas de segmentação a fim de utilizar documentos, linhas e palavras no processo de treinamento;
- Avaliar o impacto da quantidade de amostras utilizados no processo de treinamento/-teste.
- Comparar o desempenho da abordagens proposta utilizando três diferentes níveis, documentos, linhas, palavras.

Considerando os objetivos apresentados, as questões que esse trabalho dispõe-se a responder são:

- qual a real influência da quantidade de amostras utilizadas no processo treinamento?
- Dividir a imagem em linhas e palavras para seu reconhecimento tem forte influência no resultado?
- Qual dos descritores de textura empregados tem melhores resultados?

1.4. Organização do Texto

No Capítulo 2 são apresentados alguns dos conceitos necessário para o entendimento do trabalho. No Capítulo 3 encontra-se descritos alguns trabalhos importantes para a tarefa de reconhecimento de estilo de escrita. No Capítulo 4 apresentamos o método utilizado para a realização deste trabalho. Experimentos e Resultados estão dispostos no Capítulo 5. Por fim, as conclusões encontram-se no Capítulo 6.

Referencial Teórico

Neste capítulo, são descritas algumas técnicas computacionais empregadas neste trabalho. Para um maior aprofundamento em qualquer dos temas abordados, as referências aqui citadas devem ser consultadas. A Seção 2.1 apresenta os descritores que serão utilizados para a extração de características em nosso trabalho, enquanto a Seção 2.2 apresenta a técnica de aprendizado conhecida como SVM.

2.1. Descritores de Textura

Segundo Haralick (1979), uma textura pode ser descrita pelo número e tipos de suas primitivas e pela organização espacial ou *layout* de suas primitivas. A organização espacial pode ser aleatória ou pode obedecer um padrão relativamente uniforme entre uma primitiva e suas primitivas vizinhas. Esta dependência pode ser estrutural, probabilística ou funcional. Já Gool et al. (1985) define textura como uma estrutura composta de uma área contendo um grande número de elementos relativamente ordenados ou padrões que individualmente não representam muito. Mas se observados globalmente, tais elementos podem ser usados para identificar uma determinada área. Em resumo textura é uma propriedade de um *pixel* com seus vizinhos.

Gonzalez e Woods (2002) definem as principais abordagens para extrair descritores de textura são: Estatística, Estrutural e Espectral, atualmente a literatura apresenta diversas formas de extrair textura de imagens (HARALICK et al., 1973),(WU et al., 2000). Neste trabalho propomos utilizar algumas abordagem estrutural, que estão descritas a seguir.

2.1.1. Padrões Binários Locais (LBP)

O LBP é um método estrutural invariante à rotação para descrição de textura (OJALA et al., 2002). Esse método baseia-se na premissa de que padrões binários locais e a região da

vizinhança de um pixel são características fundamentais na textura da imagem. O LBP avalia para cada pixel da imagem seus vizinhos e um valor binário é atribuído a cada vizinho v por meio da Equação 2.1

$$v = \begin{cases} 1, & \text{se } i^v > i \\ 0, & \text{Caso contrário} \end{cases} \quad (2.1)$$

Onde i^v e i são a intensidade de cinza do vizinho e do pixel central. Então o pixel central é substituído pelo resultado em decimal equivalente a concatenação dos números binários obtidos pelos vizinhos. Na Figura 2.1 está ilustrado um exemplo dessa operação.

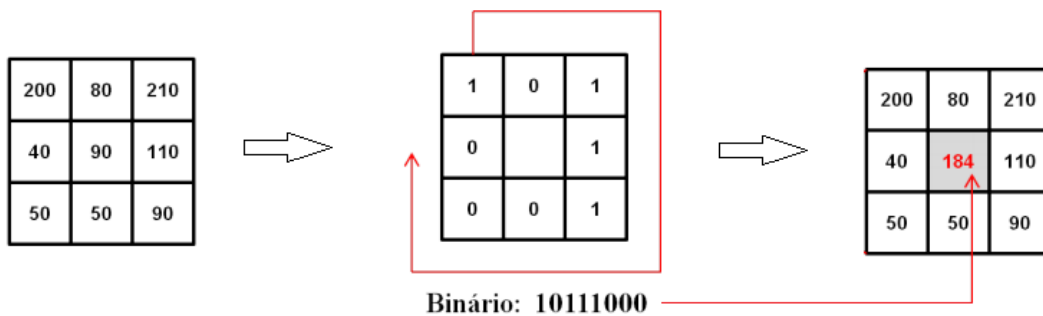


Figura 2.1. Exemplo de operação básica do LBP com valor resultante 184. adaptada de (AMARAL, 2011)

O *pixel* central é associado a um conjunto de amostras de tamanho P uniformemente espaçadas e distribuídas sobre uma circunferência de raio R , tendo como centro o *pixel* central. Assim, dois parâmetros pertencem ao método e podem assumir diversos valores; O número de vizinhos (P) e o raio (R).

A partir da imagem obtida do processo anterior, é gerado um histograma o qual mede a frequência desses valores obtidos. Usando $P = 8$ o histograma tem 256 valores. Porém, somente 58 desses valores atendem a definição de uniformidade de Ojala et al. (2002), em que a transição do *bit* zero para um e vice-versa deve ocorrer no máximo duas vezes. Os 198 valores restantes são contabilizados juntamente, resultando assim, 59 características.

2.1.2. Quantização Local de Fase (LPQ)

O LPQ é um método local para descrição de textura, proposto por Ojansivu e Heikkilä (2008). A sua característica principal é a robustez para imagens borradas ou afetadas por iluminação não uniforme. Para cada *pixel* P um código é calculado a fim de representar a textura em uma vizinhança centrada em P , de tamanho $m \times m$.

O método tem sua base nas propriedades de espectro de fases da *Short-Term Fourier Transform* (STFT) em que é computado apenas quatro coeficientes complexos dos componentes real e imaginários. São gerados, para cada *pixel* da imagem original, oito *bits*, os quais são

concatenados para formar um valor inteiro de oito *bits* que representa a textura na vizinhança. Por fim, duas outras etapas denominadas decorrelação e quantização são aplicadas. O vetor de características gerado através do LPQ é um histograma com 256 posições. outros detalhes sobre esse descritor pode ser encontrado em (OJANSIVU; HEIKKILÄ, 2008).

2.1.3. Características Robustas Aceleradas (SURF)

SURF é um descritor de textura muito similar ao *Scale-Invariant Feature Transform* (SIFT), porém, encontra os pontos de interesse com um menor custo computacional (BAY et al., 2006). O detector SURF realiza duas tarefas básicas sendo elas: detecção dos pontos de interesse, e descrição desses pontos. A detecção de pontos de interesse é baseada na matriz de Hessian, Dado um ponto $f = (x, y)$ em uma imagem I , a matriz Hessiana $H(z, \sigma)$ onde f na escala σ é definida do seguinte modo:

$$H(z, \sigma) = \begin{vmatrix} L_{x,x}(f, \sigma) & L_{x,y}(f, \sigma) \\ L_{y,x}(f, \sigma) & L_{y,y}(f, \sigma) \end{vmatrix} \quad (2.2)$$

onde $L_{x,x} = (f, \sigma)$ refere-se à derivada de segunda ordem $\frac{\partial^2}{\partial x^2}g(\sigma)$ com a imagem I no ponto x , e similarmente para $L_{x,y}(f, \sigma)$, $L_{y,y}(f, \sigma)$. Os pontos de interesse são definidos onde o determinante é máximo local.

A descrição dos pontos de interesse tem como objetivo descrever cada ponto selecionado na etapa anterior, para isso o SURF descreve como a intensidade dos *pixels* vizinhos estão distribuídos. Essa atividade consiste em fixar uma direção a partir de uma região circular definida em torno do ponto de interesse, a fim de obter um certo grau de invariância a rotação. Para a descrição, utiliza-se a Transformada de Harr aplicada nas direções horizontal e vertical. Sendo assim, é definida uma região quadrada centrada no ponto de interesse e orientada de acordo com a direção fixada. Com a finalidade de preservar as informações espaciais, esta região é dividida em sub-regiões de tamanho $n \times n$. Com $n = 4$ a quantidade de características em cada ponto é igual a 64. Esse processo está ilustrado na Figura 2.2

2.2. Máquina de Vetores de Suporte (SVM)

O classificador SVM foi proposto por Vapnik (1995) como uma técnica robusta para problemas binários. Essa técnica consiste em um método de aprendizado que busca separar duas classes por meio de um hiperplano, sendo esse hiperplano definido pelos vetores de suporte. Os vetores de suporte são instâncias mais importantes no conjunto de treinamento. Na Figura 2.3 está ilustrada a ideia da técnica do SVM.

No entanto, nem todos os problemas são separáveis linearmente, a partir disso o

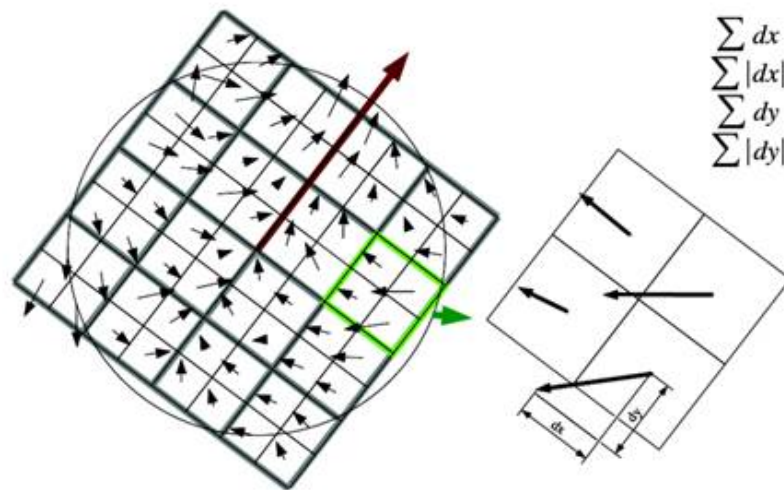


Figura 2.2. Representação gráfica do descritor SURF.

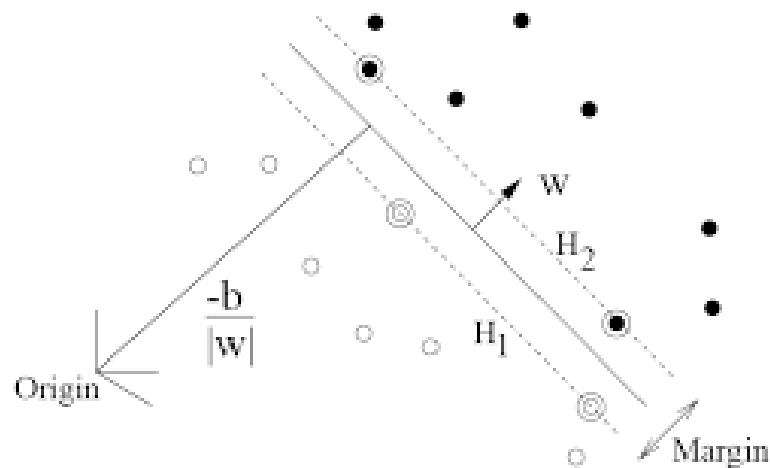


Figura 2.3. Ideia geral da técnica SVM com os vetores de suporte circulos. Retirado de (GONÇALVEZ, 2008).

SVM mapeia esses problemas para uma dimensão maior na qual passe a ser separável. Essa função denomina-se *kernel*.

Como o SVM foi concebido para classificações binárias, para contornar esse problema são utilizadas duas abordagens, uma delas consiste em considerar uma classe contra todas as outras classe (um-contra-todos) e uma classificação em pares (um-contra-um) (SOUZA, 2015).

2.3. Considerações Finais

Neste capítulo procuramos apresentar as principais técnicas computacionais a serem empregadas neste trabalho. No próximo capítulo serão apresentados alguns trabalhos relacionados com nossa pesquisa.

Trabalhos Relacionados

Neste capítulo, encontra-se descrito importantes trabalhos realizados sobre reconhecimentos de estilos de escritas em documentos manuscritos. Para cada trabalho será destacado qual a base de dados utilizada, quantas classes foram empregadas, quais os classificadores e o desempenho. A seguir apresentamos alguns trabalhos relacionados com nosso estudo.

Um estudo de identificação de estilos de escrita multinível para investigar o impacto de reconhecimento por meio de características dependentes e independentes em vários níveis, ou seja, página, bloco, linha e palavra foi conduzido por Obaidullah et al. (2019). Segundo os autores muitos estudos da época foram baseados em apenas um nível de trabalho. Para isso foi considerado onze estilos de escritas: Bangla, Devanágari, Gujarati, Gurumukhi, Kannada, Malayalam, Oriya, Roman, Tamil, Telugu e Urdu. Para Bangla e Kannada, foram utilizadas duas bases de dados existentes, CMATERdb1 para o script Bangla e KHTD para o script Kannada. Para os outros nove estilos de escritas foi requisitado a algumas pessoas transcrever em uma página em branco de acordo com uma amostra. Também foi coletado algumas escritas sem nenhuma restrição, ou seja, um texto livre. Neste projeto foi utilizado um total de 440 páginas, 2200 blocos, 3300 linhas e 6600 palavras em conjunto com os classificadores *Multi-layer Perceptron* (MLP) e *Random forest* (RF). As melhores taxas alcançadas foram de 94,32% no nível da página, 94,05% no nível do bloco, 93,73% no nível da linha e 91,04% no nível da palavra.

Uma taxa de acerto de 94,73% na identificação dos onze principais *scripts* indianos (Bangla, Devnágari, Gujarati, Gurumukhi, Kannada, Malayalam, Oriya, Roman, Tamil, Telugu e Urdu) trabalhando em reconhecimento a nível de palavra foi relatada por Ukil et al. (2018). Diferente do método proposto por Obaidullah et al. (2019) para extração de características, o trabalho menciona ter utilizado o MLP para a classificação e várias *Convolutional Neural Network* (CNN) para extração das características. Os autores relatam a escolha pela base de dados PHD_Indic_11 por ser uma base com grande quantidade de

dados. Dessa base de dados foi retirado 1.000 palavras para cada estilo de escrita, ou seja, 11.000 palavras, essas imagens foram separadas 8.800 palavras para treino e 2.200 palavras para teste.

Singh et al. (2015b) avaliam a identificação de estilos de escrita a nível de palavras. Nessa pesquisa os autores consideram sete estilos de escrita (Bangla, Devanagari, Gurumukhi, Malayalam, Oriya Telugu e Romano), e as características foram obtidas através de técnicas de aproximação elíptica e poligonal. Como seu principal resultado os autores relatam uma taxa de 95,35% de acerto utilizando o classificador MLP utilizando cerca de 7000 palavras.

Um método de combinação de classificadores para a tarefa de reconhecimento de estilos de escrita foi conduzido por Mukhopadhyay et al. (2018). Essa pesquisa também foi feita em nível de palavras, sendo sua base de dados composta por 7200 palavras, contendo 12 estilos de escrita, ou seja, 600 palavras por estilo de escrita. Na etapa de extração de características os autores relatam o uso de três técnicas (*Elliptical Features*, *Histogram of Oriented Gradients* e *Modified Log-Gabor Filter Transform*). Os autores descrevem uma taxa de acerto de 98,45% e um ganho de 7 pontos percentuais na taxa de acerto com o algoritmo de combinação de classificadores proposto.

Recentemente Souza et al. (2019) demonstraram um ganho representativo em sua pesquisa trabalhando com uma ideia de blocos de textura. Os autores consideraram 280 documentos manuscritos digitalizados a 150 *dpi* e divididos em cinco sistemas de escritas diferentes: Bengali, Oriya, Kannada, Persa e Romana. Em seus experimentos, 60% das amostras foram usadas no conjunto de treinamento e 40% no conjunto de teste. Com a extração de características por meio do descritor de textura SURF e uso do classificador SVM os autores relataram uma taxa de acerto de 98,48%.

A aplicação de um pré-processamento nas imagens de documentos foi proposta por Singhal et al. (2003), utilizando 480 documentos divididos entre quatro estilos de escrita: Romano, Devanágari, Bengali e Telugu. O pré-processamento era dividido em cinco etapas que foram executados na seguinte ordem: remoção de ruídos, esqueletização morfológica, *pruning*, conectividade-*m* e por fim, a normalização do tamanho do texto. O Filtro de Gabor Multicanal foi utilizado como característica de textura invariante a rotação. Em seguida, foi realizado uma abordagem probabilística para clusterização. A motivação para o uso dessa abordagem surgiu devido à grande diferença intraclasse das quatro classes utilizadas nesse trabalho. Assim, os autores reportam uma taxa de acerto de 91,6%.

Nos países indianos uma quantidade considerável de documentos são escritos com mais de um estilo de escrita. Singh et al. (2018) trabalham com o reconhecimento *bi-script* e *tri-script* para isso ele utiliza os dois estilos de escritas mais populares na Índia (Devanagari e Bangla) combinado com o estilo de escrita Romano. No seu trabalho Singh et al. (2018) organiza uma base de dados com 150 documentos descritos em Bangla-Romano e 150 Devanagari-Romano, essas páginas geram no total 18.931 palavras em Bangla, 15.528

palavras em Devanagari e 10.331 palavras em Romano. A extração das características é feita por meio do filtro *Modified log-Gabor*. As melhores taxas relatadas são de 92,32% para Bangla-Romano com 14.950 palavras para treino e 8.879 palavras para teste, 95,30% para Devanagari-Romano utilizando 13.366 palavras para treino e 7615 para teste e 93,78% Bangla-Devanagari-Romano empregando 29.018 palavras no conjunto de treinamento e 15.792 no teste, essas taxas foram obtidas usando o classificador MLP.

Utilizando o cenário de *bi-script* e *tri-script*, Rajput e Ummapure (2019) apresentam uma combinação de estilos de escrita utilizando o Romano juntamente com os 11 principais estilos de escritas da Índia, gerando alguns cenários *bi-script* e *tri-script*. Nesse trabalho são recortados blocos de ao menos três linhas de um documento manuscrito, e com esses blocos os autores empregam o descritor de textura LBP para a extração das características juntamente com os classificadores SVM e K Vizinhos mais Próximos (KNN). As melhores taxas de acerto neste trabalhos foram de 99,5% utilizando o classificador SVM e 98,46% com o classificador KNN.

Diferente do método proposto por Rajput e Ummapure (2019) o qual propõe combinar estilos de escritas para o reconhecimento *bi-script* e *tri-script*, Pal et al. (2003) propõe um método de identificação dos de estilo de escrita a nível de linha do documento sem nenhum processo de treinamento, ou seja, se em uma única página do documento existir todos esses doze estilos de escrita (Devnagari, Bangla, Romano, Gujarati, Kannada, Kashmiri, Malayalam, Oriya, Gurumukhi, Tamil, Telugu e Urdu). O método proposto busca identificar sem nenhum conhecimento prévio do documento. Para o reconhecimento, os autores apresentam uma árvore de decisão baseada em características estruturais presentes em cada estilos de escrita. Com essa proposta é obtida uma taxa de 97,52% de acerto.

3.1. Discussão

A Tabela 3.1 apresenta um resumo dos trabalhos aqui descritos. A partir dela é possível notar um aumento no uso de técnicas similares as empregadas neste trabalho. Entretanto, realizar uma comparação entre elas torna-se uma tarefa difícil devido a utilização de diversas base de dados, as quais possuem quantidade e estilos de escrita distintos, e também ao emprego do reconhecimento em diferentes níveis (páginas,palavra,linha e bloco). Percebemos também que alguns detalhes não são descritos em diversos trabalhos. Um exemplo é que não fica claro se os trabalhos utilizam partes do mesmo documento nos conjuntos de treinamento e teste. Em alguns casos pode haver uma super adaptação com relação a escrita de uma mesma pessoa, ou seja, o classificador por "aprender" detalhes do escritor e não do *script* escrito. Outro detalhe é que algumas bases possuem texto-dependente. Neste caso, os escritores copiam um texto base. Desta forma pode existir muita similaridade entre diferentes documentos. Dependendo do método de extração de características utilizado, usar o mesmo texto no treinamento e

teste pode gerar um viés na tarefa de reconhecimento de *scripts*.

Tabela 3.1. Resumo da Revisão Bibliográfica

Autores	Ano	Estilos de Escrita	Quant. de Dados	Nível	Características	Classificação	Desempenho
Obaidullah et al. (2019)	2019	11	440 Páginas, 2200 Blocos, 3300 Linhas e 6600 Palavras	Página,Bloco, Linha e Palavra	Características Dependentes e Independentes	Rede Neural Perceptron Multi-camada, Floresta Aleatória.	94,32% para página, 94,05% para bloco, 93,73% para linha e 91,04% para palavra
Ukil et al. (2018)	2018	11	11000 Palavras	Palavras	Rede Neural Convolutacional	Rede Neural Perceptron Multi-camada	94,73%
Singh et al. (2015b)	2015	7	7000 Palavras	Palavra	Aproximação Elíptica e Poligonal	Rede Neural Perceptron Multi-camada	95,35%
Mukhopadhyay et al. (2018)	2018	12	7200 Palavras	Palavras	<i>Elliptical Features, Histogram of Oriented Gradients e Modified Log-Gabor Filter Transform</i>	-	98,45%
Souza et al. (2019)	2019	5	280 Páginas	Bloco	Características Robustas Aceleradas	Máquina de Vetores de Suporte	98,48%
Singhal et al. (2003)	2003	4	480 Páginas	-	Filtro de Gabor Multicanal	Abordagem Probabilística	91,6%
Singh et al. (2018)	2018	3	300 Páginas	Palavra	<i>Modified log-Gabor</i>	Rede Neural Perceptron Multi-camada	92,32% Bangla-Romano, 95,30% Devanagari-Romano, 93,78% Bangla-Devanagari-Romano
Rajput e Ummapure (2019)	2019	12	250 Páginas	Bloco	Padrões Binários Locais	K Vizinhos mais Próximos, Maquina de Vetores de Suportes	99,5% SVM 98,46% k-NN
Pal et al. (2003)	2019	12	-	Linhas	Características estruturais	-	97,52%

Com isso, em nossa pesquisa trabalhamos com três níveis de reconhecimento sendo eles: páginas, linhas, palavras. utilizamos os descritores de textura apresentados nos trabalhos de Souza et al. (2019), Rajput e Ummature (2019) e um terceiro descritor que não foi citado em nossos trabalhos relacionados denominado LPQ.

Ao aplicar os descritores SURF e LBP será possível analisar se as taxas apresentadas pelos autores serão mantidas ao trabalhar com outra base de dados, novos estilos de escrita e diferentes níveis.

Método

Este capítulo apresenta o método utilizado no desenvolvimento deste trabalho. A Figura 4.1 apresenta a abordagem proposta e em seguida cada etapa será descrita separadamente.

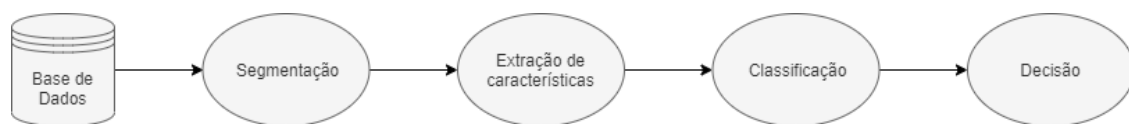


Figura 4.1. Visão geral do método proposto

Como apresentado na Figura 4.1, a partir da nossa base de dados serão aplicadas técnicas de segmentação de imagens a fim de extrair linhas e palavras dos documentos manuscritos para que seja possível trabalhar em três níveis, documento, linhas e palavras. Na etapa seguinte será empregado três abordagens para extração de características, usando os descritores LBP, SURF e LPQ. Em seguida será realizado o processo de classificação empregando o classificador SVM. Por fim, uma decisão será gerada com intuito de rotular uma determinada amostra dentre as onze possíveis classes existentes.

4.1. Base de Dados

Empregamos a base de dados proposta por Obaidullah et al. (2018) em nossa pesquisa. Essa base possui 1458 páginas manuscritas escrita por 463 pessoas em 11 estilos de escritas oficiais da Índia (Bangla, Devanagari, Roman, Urdu, Oriya, Gurumukhi, Gujarati, Tamil, Telugu, Malayalam and Kannada). A Tabela 4.1 resume como estão distribuídas as páginas, número de escritores, número de linhas e número palavras em cada uma das formas de escritas mencionadas.

Obaidullah et al. (2018) propõe a criação de uma base de dados de documentos manuscritos na qual os documentos são separados como documentos de texto-dependente e texto-independente. Segundo Obaidullah et al. (2018), texto-dependente são na verdade

Tabela 4.1. Distribuição dos dados na base PHDIndic_11

Forma de Escrita	Num. Escritores	Num. páginas	Num. linhas	Num. Palavras
Bangla	42	161	1820	12447
Devanagari	60	220	2457	23264
Urdu	45	201	1595	18422
Oriya	40	172	1422	10673
Tamil	71	120	991	5534
Telugu	46	85	826	4585
Malayalam	36	107	1028	6896
Kannada	17	46	307	1562
Roman	45	112	1521	14128
Gurumukhi	50	132	1601	12946
Gujarati	11	100	1442	13822
Total	463	1458	15010	124279

cópias manuscritas de algum texto de referência. Neste caso todos os escritores escrevem a próprio punho (copiam) um texto base. No caso de texto-independente ou livre, cada escritor redige sobre um tema qualquer, não há exigência sobre um número mínimo ou máximo de caracteres ou linhas. Assim, para a elaboração dessa base de dados cada escritor cedeu (escreveu) seis documentos, os quais cinco se tratavam de textos-dependente e um texto-independente. Esses documentos foram digitalizados usando um *scanner* HP a 300 *dpi* e foram disponibilizados em nível de cinza. Como mencionado na Seção 3.1 o uso de texto-dependente pode gerar um certo viés na pesquisa como, adaptação aos detalhes do escritor e não as características específicas do estilo de escrita. Isso pois são utilizadas amostras de texto-dependente nos conjuntos de treinamento/teste. Assim, podemos ter o mesmo texto no conjunto de treinamento e no conjunto de testes. Ao trabalhar em níveis como documento, blocos, linhas e palavras, texto-dependente podem ocultar o real desempenho do trabalho.

4.2. Segmentação

Ao iniciarmos essa pesquisa notamos que é comum em países indianos o uso de mais de um *script* no mesmo documento. Ou seja, é comum o autor escrever linhas ou palavras usando um outro *script*. Assim, rotular um documento como pertencente a uma ou outra classe (*script*) pode ser um erro. Na base a qual estamos empregando isso não ocorre. Pois cada documento é escrito utilizando um único *script*. Desta forma nos deparamos com as seguintes questões: a) Se considerarmos um documento específico em diferentes níveis ou seja, página, linha ou palavra, nossas taxas de acerto serão diferentes de um nível para outro?; b) Quais características são mais adequadas para um determinado nível?; c) Trabalhar com linhas gera um resultado melhor que trabalhar com páginas, ou seja, qual é o nível ideal de trabalho para

a atividade de reconhecimento de estilo de escrita? Para responder as questões mencionadas acima segmentamos toda nossa base de dados em dois níveis. A Figura 4.2 apresenta as etapas realizadas na fase de segmentação da base de dados. Posteriormente será descrito como essas etapas foram realizadas.

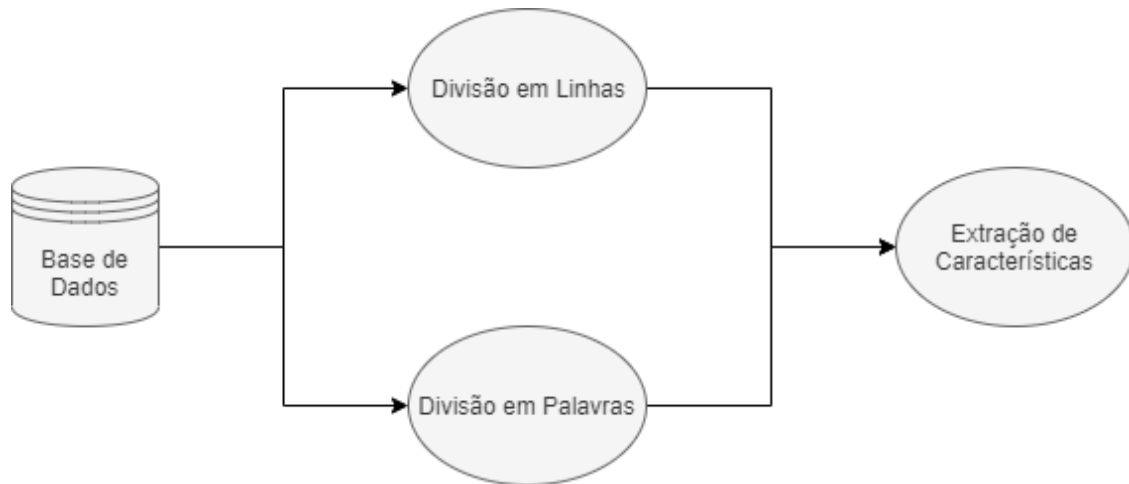


Figura 4.2. Etapa de Extração de Características

O método utilizado no processo de segmentação de nossas imagens é baseado em operações morfológicas. Em resumo, operações morfológicas são uma série de algoritmos de processamento de imagens que consiste em aplicar aos *pixels* de uma imagens alguns kernels pré definidos. Para que fosse possível utilizar essas operações morfológicas, inicialmente tivemos que converter nossas imagens em imagens binárias. Para isso, utilizamos um método de limiarização proposto por Otsu (1979). Após o processo de limiarização e conversão das imagens para binárias, tivemos que definir um elemento estruturador de medidas $M \times N$ *pixels*. Com esse elemento estruturador aplicamos uma operação de dilatação em nossas imagens binárias com o intuito de aumentar as palavras da imagem. Após esse processo, encontramos os componentes conexos nas imagens e demarcamos-as com caixas delimitadoras, ou seja, as palavras foram encontradas e delimitadas.

Com as coordenadas dessas caixas delimitadoras, foi aplicado uma função de recorte na imagem original para que uma nova imagem fosse gerada. Essa nova imagem em nosso trabalho representava as linhas/palavras. A Figura 4.3 ilustra o processo aqui descrito, onde Figura 4.3(a) representa a imagem binária e Figura 4.3(b) representa a imagem demarcada com as caixas delimitadoras.

Ainda no processo de segmentação, a base de dados passou por um seleção das amostras. Essa seleção se deu pelo motivo de trabalharmos com documentos manuscritos os quais possuem espaçamentos não regulares, sobreposições de traços/caracteres entre outros. Ou seja, caso palavras em linhas diferentes estejam conectadas de alguma forma, o método de segmentação pode entender que seja uma só palavra. Então extrair componentes conexos dessas imagens pode retornar imagens que não represente uma linha ou uma palavra. Nesse

ଏକ ଧର୍ମ ଓ ଧର୍ମ । ଶ୍ରୀମ ନାମ ଉଚ୍ଚିଷ୍ଟିତ୍ଵ କରନ୍ତର ଅତି-ସରସର
 ଅର ବ୍ୟୁତ୍ଥାନାମନେ ଉଦ୍ଘୋରନ । ନିର୍ଦ୍ଦିତା ସ୍ଵଳ ରିତେ ଶକ୍ତ ସଂସାର ଅର
 ଉଦ୍ଘେ ଉଦ୍ଘୋରନ । ଯିତା ସ୍ଵାଧୀନିତା ବିଦ୍ୟାଧା । ଦିଗ୍ଵୁ ବେଶେ ମିତାଧ୍ୟା-
 କାଳେର ଅତ୍ୟାଧିକାରେ ଉଦ୍ଘେଷ୍ଟିତ୍ଵ ସ୍ଵଳସ୍ଵାଧୀନିତା',
 'ଜ୍ଞାନାଶ୍ରମୀ ଦିଗ୍ଘୋଦିଗ୍ଘାଧା ମୁକ୍ତାଧ୍ୟାଧିକାର', 'ସେକ୍ଵିତ୍ଵକାର ବିଦ୍ଵାଧ୍ୟାଧିକାର
 ବିଦ୍ଵିଷ୍ଟାଧିକାର ଶ୍ରୀକାର୍ଯ୍ୟ ସ୍ଵାଧୀନିତା' - ନାମକ ମାଧ୍ୟମିକ ଅତ୍ୟାଧିକାର
 କରନ୍ତେ ସାରିଲେ ମୁଁ ଶକ୍ତ ଅଦ୍ଘେ ପ୍ରକୋଷାଧା କରନ୍ତେ । ଯି
 ଉଦ୍ଘେ ଶ୍ରୀକାର୍ଯ୍ୟ ଓ ଶକ୍ତାଧୀନିତା ଚକ୍ଵୁର ଅଦ୍ଘେ ଅନିତାଧ୍ୟାଧିକାର
 କି ଶ୍ରୀକ ! ଅନିତାଧ୍ୟାଧିକାର ଚକ୍ଵୁର ଓ-ଅଧ୍ୟାଧିକାର ଯିତେ ଅନିତାଧ୍ୟାଧିକାର
 ଅନିତାଧ୍ୟାଧିକାର - ଶକ୍ତ କରନ୍ତେ କରନ୍ତେ କରନ୍ତେ ।

ଏକ ଧର୍ମ ଓ ଧର୍ମ । ଶ୍ରୀମ ନାମ ଉଚ୍ଚିଷ୍ଟିତ୍ଵ କରନ୍ତର ଅତି-ସରସର
 ଅର ବ୍ୟୁତ୍ଥାନାମନେ ଉଦ୍ଘୋରନ । ନିର୍ଦ୍ଦିତା ସ୍ଵଳ ରିତେ ଶକ୍ତ ସଂସାର ଅର
 ଉଦ୍ଘେ ଉଦ୍ଘୋରନ । ଯିତା ସ୍ଵାଧୀନିତା ବିଦ୍ୟାଧା । ଦିଗ୍ଵୁ ବେଶେ ମିତାଧ୍ୟା-
 କାଳେର ଅତ୍ୟାଧିକାରେ ଉଦ୍ଘେଷ୍ଟିତ୍ଵ ସ୍ଵଳସ୍ଵାଧୀନିତା',
 'ଜ୍ଞାନାଶ୍ରମୀ ଦିଗ୍ଘୋଦିଗ୍ଘାଧା ମୁକ୍ତାଧ୍ୟାଧିକାର', 'ସେକ୍ଵିତ୍ଵକାର ବିଦ୍ଵାଧ୍ୟାଧିକାର
 ବିଦ୍ଵିଷ୍ଟାଧିକାର ଶ୍ରୀକାର୍ଯ୍ୟ ସ୍ଵାଧୀନିତା' - ନାମକ ମାଧ୍ୟମିକ ଅତ୍ୟାଧିକାର
 କରନ୍ତେ ସାରିଲେ ମୁଁ ଶକ୍ତ ଅଦ୍ଘେ ପ୍ରକୋଷାଧା କରନ୍ତେ । ଯି
 ଉଦ୍ଘେ ଶ୍ରୀକାର୍ଯ୍ୟ ଓ ଶକ୍ତାଧୀନିତା ଚକ୍ଵୁର ଅଦ୍ଘେ ଅନିତାଧ୍ୟାଧିକାର
 କି ଶ୍ରୀକ ! ଅନିତାଧ୍ୟାଧିକାର ଚକ୍ଵୁର ଓ-ଅଧ୍ୟାଧିକାର ଯିତେ ଅନିତାଧ୍ୟାଧିକାର
 ଅନିତାଧ୍ୟାଧିକାର - ଶକ୍ତ କରନ୍ତେ କରନ୍ତେ କରନ୍ତେ ।

(a)

(b)

Figura 4.3. Resumo do processo de segmentação da imagem em palavras.

processo eliminamos todas as amostras que não representavam unicamente linhas ou palavras. A partir disso, seguimos para a fase de extração de características.

4.3. Extração de Características

A Figura 4.4 apresenta o processo de extração de características, o qual consiste em extrair essas características por meio de três descritores de textura os quais iram ser utilizadas para a classificação.

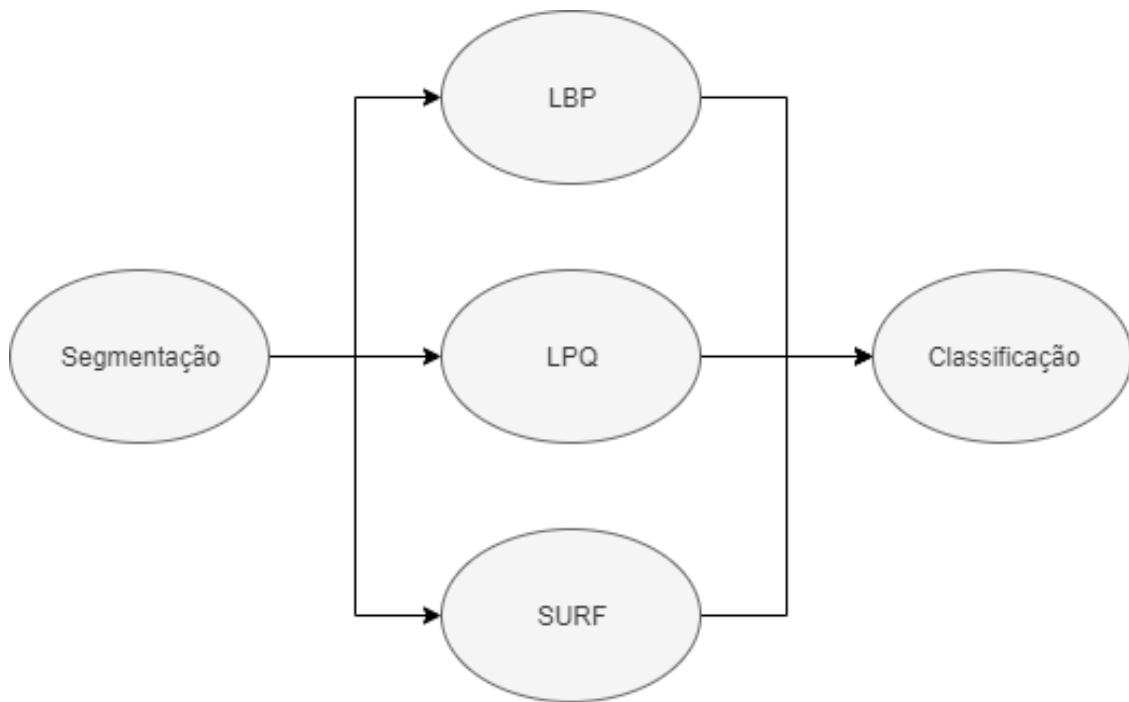


Figura 4.4. Reconhecimento do estilo de escrita (Abordagem I)

Nossa abordagem considera três descritores distintos para a extração de características (SURF, LBP, LPQ) em três níveis: documentos, linhas e palavras. Essas características

foram utilizadas para alimentar o nosso classificador SVM. Embasados nos trabalhos de Souza et al. (2019) e Rajput e Ummapure (2019), os quais utilizam os descritores SURF e LBP trabalhando com blocos respectivamente, optamos por utilizar esses dois descritores em outros níveis (linhas e palavras) acrescentando novos estilos de escrita.

Na pesquisa conduzida por Souza et al. (2019) os autores relatam o uso do descritor SURF com $SURFsize = 128$. Também foi considerado medidas como: média, desvio padrão, obliquidade e curtose, gerando assim um vetor com 513 características. Essa abordagem será investigada para relatar o comportamento desse descritor quando aplicado a um conjunto maior de classes (estilos de escritas) e diferentes níveis de trabalho.

O descritor LBP foi empregado no trabalho de Rajput e Ummapure (2019) com a problemática de reconhecimento de estilos de escritas em cenários *bi-script* e *tri-script* utilizando os mesmos 11 estilos de escritas de nossa pesquisa. Essa pesquisa é apresentada com taxas muito próximas a 100%, então aplicaremos esse descritor em nossa pesquisa a documentos com apenas um estilo de escrita. Esse descritor irá nos ajudar a investigar se essas taxas também se repetem ao trabalhar com outra base de dados e diferentes níveis de trabalho.

Propusemos a utilização de um terceiro descritor de textura denominado LPQ. Embora esse descritor não tenha aparecido em nossos trabalhos relacionados, esse descritor é amplamente utilizado em pesquisas de identificação de escritor como apresentado em (HANNAD et al., 2016), (BERTOLINI et al., 2016) entre outros. Então, aplicamos esse descritor para o problema de reconhecimento de estilos de escrita em documentos manuscritos. Assim, verificaremos a viabilidade de utilização desse descritor para o problema.

Com os descritores aqui apresentados, analisamos a viabilidade de nossa proposta para o problema de reconhecimento de estilos de escrita (reconhecimento de *script*) e também concluiremos algumas questões levantadas durante a revisão bibliográfica.

4.4. Classificação

Em nossa pesquisa, usamos o classificador SVM na etapa de classificação com o *kernel Radial Basis Function* (RBF). esse *kernel* possui dois parâmetros. O parâmetro de margem C que determina um ponto de equilíbrio razoável entre a maximização da margem e a minimização do erro da classificação, e o parâmetro γ que controla o raio do kernel RBF. O problema ao utilizar esse kernel consiste em procurar um ponto ótimo que maximize a precisão do classificador, ou seja, melhor C e γ para um dado problema.

Para melhor estimar nossos parâmetros C e γ utilizaremos a biblioteca LIBSVM. A biblioteca LIBSVM disponibiliza processos extras que auxilia na escolha dos parâmetros C e γ utilizados pelo kernel RBF. Para melhor estimar os parâmetros o LIBSVM executa o processo de *Cross-validation*. Esse processo consiste em dividir todas as instâncias que serão

utilizadas no processo de treinamento em 5 subconjuntos iguais. Logo após essa divisão é executado o processo de *Grid-search* que por sua vez tenta estimar os melhores parâmetros C e γ para cada subconjunto gerado no processo de *Cross-validation* então os parâmetros com melhores acurácia são escolhidos para treinar e gerar o modelo de treinamento.

Experimentos e Resultados

Nos experimentos realizados, empregamos a base proposta por (OBAlDULLAH et al., 2019) a qual possui onze estilos de escritas diferentes, sendo elas, Bangla, Devanagari, Gujarati, Gurumukhi, Kannada, Malayalam, Oriya, Roman, Tamil, Telugu e Urdu.

Inicialmente, foi empregada a técnica de segmentação descrita anteriormente nesse trabalho. Em seguida, a partir dessas do processo de segmentação as imagens geradas passaram por uma etapa de validação de amostras, essa verificação tinha como objetivo excluir todas amostras que não representavam palavras ou linhas, isso foi feito de forma manual. Após a validação do conjunto de dados essas amostras foram utilizadas no processo de extração de características. A extração de características se deu por meio de três descritores, LBP, LPQ e SURF..

Para a extração de características os parâmetros utilizados para cada descritor de textura foram: LBP uniforme $LBP_{P,R}^{U2}$ com 59 características; SURF com $SURFsize = 64$, além das medidas: Média, desvio padrão, obliquidade e curtose, gerando assim um vetor de 257 características; LPQ utilizando o $winSize = 7$, gerando um vetor de 256 características. A classificação foi realizada utilizando o classificador SVM com o *Kernel RBF*.

Durante os experimentos foi utilizado o esquema de validação cruzada utilizando $K folds$. As imagens de entrada foram divididos em $K folds$. No processo de treinamento foi utilizado $k - 1$ subconjuntos e um subconjunto é utilizado para teste. Esse processo foi repetido K vezes, de forma com que todos os $fold$ s façam parte do conjunto de treinamento e teste. Neste trabalho adotamos $k = 5$.

A métrica utilizada para a avaliação dos resultados é a precisão do reconhecimento ou acurácia, que é calculado da seguinte forma:

$$acurácia = \frac{Documentos_Corretamente_Classificados}{Total_Documentos} * 100\% \quad (5.1)$$

Como cada experimento foi executado cinco vezes, para diferentes combinações de amostras no conjunto de treinamento e teste, os resultados apresentados na seção a seguir referem-se a média e desvio padrão dessas cinco execuções.

Nossos experimentos foram separados em duas etapas. A primeira etapa consiste em aplicar nosso método a uma quantidade de documentos pré definido baseado no número de amostras utilizadas em alguns trabalhos em quais nos embasamos. No segundo experimento adicionamos algumas amostras nos conjuntos de treino/teste, o objetivo dessa inserção era analisar o comportamento de nossa taxa de acerto ao utilizar mais informação no processo de treino/teste em nossa aplicação. Com isso, na Seção 5.1 apresentamos os resultados alcançados nessa pesquisa trabalhando com nível de documento, na Seção 5.2 é apresentado os resultados quando trabalhamos no nível de linhas, e na Seção 5.3 está nossos resultados utilizando palavras em nossos experimentos.

5.1. Reconhecimento a Nível de Documentos

Para compararmos nossos resultados com alguns trabalhos apresentado na revisão bibliográfica, definimos que a quantidade total de documentos utilizados em nossa pesquisa será limitada pelo numero de amostrar utilizados nos trabalhos aos quais nos embasamos. O trabalho descrito por (OBAIDULLAH et al., 2019) emprega as mesmas onze classes, porém com diferentes quantidades de amostras no treinamento e teste. Neste caso, os autores utilizaram 352 amostras no conjunto de treinamento e 88 no conjunto de teste. A partir disso, executamos nossos experimentos com essa mesma quantidade, nossos resultados estão presentes na Tabela 5.2. Para avaliarmos os impactos gerados pela quantidade de amostras utilizados no processo de treinamento, realizamos outro experimento diminuindo a quantidade de documentos utilizados. O resultado desse experimento está descrito na Tabela 5.1

Tabela 5.1. Taxa de Reconhecimento (%) para nível de Páginas - Experimento 1

Descritor	Num. Amostra Teste	Num. Amostra Treino	Taxa Acerto	Desvio Padrão
LBP	55	220	87,04%	±6,27
LPQ	55	220	94,55%	±3,15
SURF	55	220	90,91%	±4,18

Tabela 5.2. Taxa de Reconhecimento (%) para nível de Páginas - Experimento 2

Descritor	Num. Amostra Teste	Num. Amostra Treino	Taxa Acerto	Desvio Padrão
LBP	88	352	92,04%	±4,13
LPQ	88	352	92,05%	±2,94
SURF	88	352	97,73%	±2,66

Dentre esses resultados, estão destacadas as maiores taxas obtidas em cada um dos experimentos. Podemos observar que o descritor SURF obteve o melhor resultado ao trabalhar em nível de documentos no experimento 2. Esse desempenho relatado pelo descritor SURF era esperado em nossa pesquisa já que Souza et al. (2019) reportaram taxas muito próximas em seu trabalho. Podemos concluir que essa abordagem proposta considerando o descritor SURF é promissora e robusta para o problema de reconhecimentos de estilo de escrita em nível de documentos. Entretanto, na Seção 4.1 levantamos alguns problemas sobre a base de dados utilizada. Devido ao fato do descritor SURF ser utilizado em um processo de *Feature Matching* esse descritor pode ter se beneficiado do texto-dependente presente na base de dados.

Para investigar se a quantidade de documentos utilizado no processo de treinamento/teste está diretamente relacionada com a taxa de acerto comparamos os resultados dos experimentos realizados. Então, através do gráfico presente na Figura 5.1 podemos notar que essa relação varia conforme o método de extração de características empregado. A fim de investigar se essa variação realmente era mantida ao adicionarmos cada vez mais amostras nos conjuntos de treino/teste, conduzimos um terceiro experimento. Nesse experimento, consideramos utilizar toda base de dados nos conjuntos de treino/teste. Os resultados desse experimento estão apresentados na Tabela 5.3. Através deste experimento é possível notar que a quantidade de documentos utilizados no processo de treinamento/teste não é o único fator que influencia no desempenho da aplicação. A partir disso, concluímos que considerar cada vez mais amostras no conjunto de treinamento não implicará em um melhor desempenho. Acreditamos que, um processo de seleção de instâncias também possa influenciar no desempenho da aplicação, ou seja, se acrescentarmos as instâncias mais difíceis de ser rotuladas no conjunto de treinamento, o classificador aprenderá melhor sobre aquela determinada classe.

Tabela 5.3. Taxa de Reconhecimento (%) para nível de Páginas - Experimento 3

Descritor	Num. Amostra Treino	Num. Amostra Teste	Taxa Acerto	Desvio Padrão
LBP	1169	289	99,31%	$\pm 0,29$
LPQ	1169	289	99,12%	$\pm 0,12$
SURF	1169	289	99,77%	$\pm 0,15$

5.2. Reconhecimento a Nível de Linha

Ao trabalhar com o reconhecimento a nível de linha, percebemos que nossos resultados são muito similares aos que foram alcançados quando trabalhamos com documentos. Os resultados dos experimentos realizados estão descritos nas Tabelas 5.4 e 5.5.

A partir dos resultados descritos nas Tabelas 5.4 e 5.5 é possível notar que o descritor LPQ obteve o melhor desempenho dentre os descritores empregados no reconhecimento a

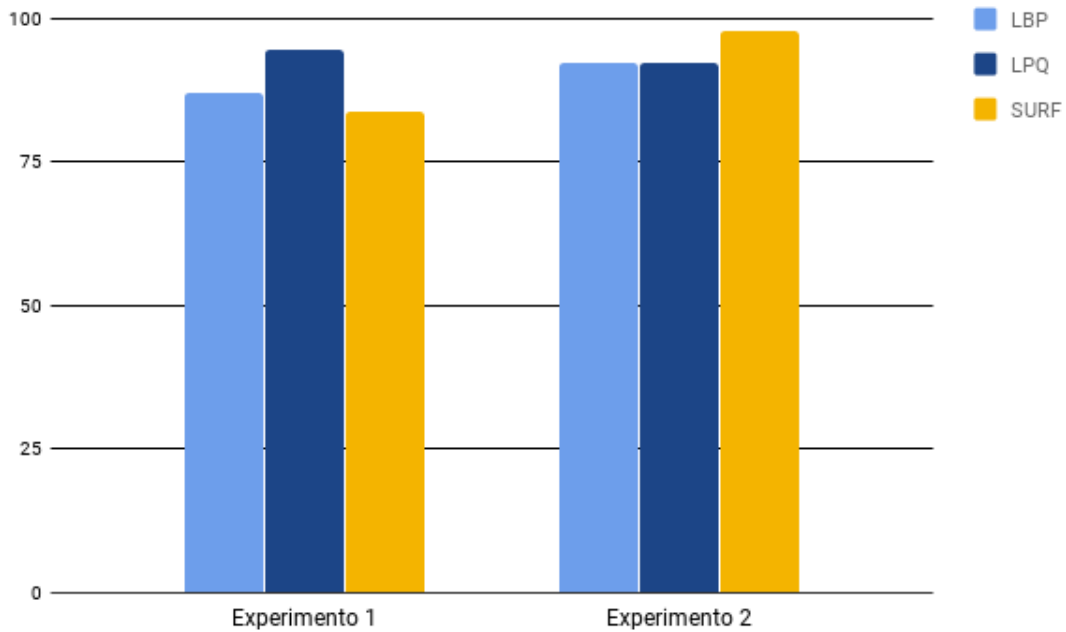


Figura 5.1. Resultado dos experimentos em nível de Páginas

Tabela 5.4. Taxa de Reconhecimento (%) para nível de Linhas - Experimento 1

Descritor	Num. Amostra Treino	Num. Amostra Teste	Taxa Acerto	Desvio Padrão
LBP	1760	440	89,49%	$\pm 3,87$
LPQ	1760	440	91,19%	$\pm 1,81$
SURF	1760	440	88,78%	$\pm 4,17$

Tabela 5.5. Taxa de Reconhecimento (%) para nível de Linhas - Experimento 2

Descritor	Num. Amostra Treino	Num. Amostra Teste	Taxa Acerto	Desvio Padrão
LBP	2200	550	92,55%	$\pm 2,2$
LPQ	2200	550	96%	$\pm 2,77$
SURF	2200	550	95,09%	$\pm 4,11$

nível de linhas. Seguido pelo descritor SURF que em algumas execuções foi superior ao descritor LPQ, porém as taxas apresentadas por esse descritor possui uma variação muito grande o que interferiu muito na sua taxa média de acerto comparado com o descritor LPQ.

Embasado nesses resultados, podemos concluir que trabalhar a nível de linhas podemos alcançar resultados próximos aos apresentados a nível de documentos.

5.3. Reconhecimento a Nível de Palavra

Dentre todos os nossos experimentos, o reconhecimento a nível de palavras obteve o pior desempenho. Esse desempenho está descrito nas Tabelas 5.6 e 5.7.

Tabela 5.6. Taxa de Reconhecimento (%) para nível de Palavras - Experimento 1

Descritor	Num. Amostra Treino	Num. Amostra Teste	Taxa Acerto	Desvio Padrão
LBP	8800	2200	86,37%	$\pm 4,84$
LPQ	8800	2200	90,91%	$\pm 1,53$
SURF	8800	2200	68,23%	$\pm 3,5$

Tabela 5.7. Taxa de Reconhecimento (%) para nível de Palavras - Experimento 2

Descritor	Num. Amostra Treino	Num. Amostra Teste	Taxa Acerto	Desvio Padrão
LBP	17600	4400	90,47%	$\pm 2,88$
LPQ	17600	4400	94,37%	$\pm 2,12$
SURF	17600	4400	75,59%	$\pm 1,50$

Como esperado e também descrito no trabalho de (OBAIDULLAH et al., 2018), o reconhecimento em nível de palavras tem o desempenho inferior se considerarmos os níveis de páginas e linhas. Esse desempenho inferior deve-se a quantidade de informação presente nos documentos.

Assim como no reconhecimento a nível de linhas, o descritor LPQ obteve a maior taxa de acerto. O descritor SURF obteve o pior desempenho ao trabalhar nesse nível. Esse desempenho foi muito afetado devido à presença de pouca informação nesses documentos trabalhando com nível de palavras. Com isso, o descritor SURF encontra poucos pontos de interesse para descrever essas imagens, ou seja, essas imagens são descritas de uma forma muito superficial.

5.4. Considerações Finais

A partir dos resultados obtidos, é possível verificar que a utilização de ambos descritores apresentam ótimas taxas de acerto no problema de reconhecimento de estilos de escrita. Na Tabela 5.8, comparamos os resultados alcançados nesse trabalho com alguns trabalhos da literatura que utilizam a mesma base de dados. Podemos notar que o desempenho do método proposto em alguns casos é superior aos descritos na literatura.

Tabela 5.8. Estudo comparativo

Métodos	Taxas		
	Documentos	Linhas	Palavras
Obaidullah et al. (2019) (<i>Hand-crafted Features</i>)	94,32%	93,73%	91,04%
Ukil et al. (2018) (<i>Deep learning</i>)	-	-	94,73%
Nosso Método (LBP, LPQ, SURF)	97,73%	96%	94,37%

Conclusões

O principal objetivo desse trabalho foi avaliar a aplicação de um método robusto na tarefa de reconhecimento de estilos de escrita indiana. Para isso foi feito um levantamento bibliográfico a fim de avaliar desempenho de características e classificadores além de detalhes do problema de reconhecimento de estilos de escrita. Foi identificado detalhes como: características, bases de dados, desempenho, dentre outras. Esses trabalhos são destacados na Seção 3.

Na Seção 4 apresentamos o método utilizado na realização do presente trabalho. Consideramos uma abordagem para realizarmos nossos objetivos de pesquisa, na qual, consiste em aplicar três descritor de textura com o classificador SVM.

A partir dos resultados obtidos em nossa pesquisa podemos assumir que a proposta apresentada pode ser descrita como uma abordagem robusta e promissora, uma vez que consideramos uma base de dados com uma vasta quantidade de classes e uma grande variedade de estilos de escritas.

As Seções 6.1, 6.2 apresentam respectivamente as contribuições deste trabalho e os trabalhos a serem realizados futuramente.

6.1. Contribuição

Dentre as contribuições do presente trabalho, podemos destacar:

- Estudo da influência de um esquema de divisão em páginas,linhas e palavras.
- Estudo do impacto do uso de diferentes descritores de textura;
- Verificação da relação entre o número de amostras no processo de treinamento e a dificuldade do problema;
- Análise de uma abordagem de segmentação de textos em imagens;

6.2. Trabalhos Futuros

Durante a realização deste trabalho, observamos algumas questões relacionadas, que mesmo não fazendo parte deste trabalho, são relevantes e podem ser investigadas futuramente

- **Aplicação do método em outra base de dados:** Observamos que a base utilizada nessa pesquisa era composta por documentos de texto-dependente, então descrevemos alguns possíveis problemas com esse tipo de documentos. Entretanto a realização de novos experimento com outras bases de dados pode contribuir demonstrando a eficiência da abordagem aqui proposta.
 - **Extração de características *Non-Handcrafted* :** Acreditamos que utilizar o processo de aprendizagem profunda na etapa de extração de características pode prover bons resultados se aplicada a este problema.
 - **Seleção de instâncias para o conjunto de treinamento:** Ao utilizarmos o esquema de segmentação dos dados, notamos uma maior dependência do conjunto de treinamento para o sucesso do reconhecimento. Isto sustenta uma avaliação de métodos para a seleção de instâncias a serem utilizadas no conjunto de treinamento.
- Data Augmentation* :** Investigar aplicação de processos de *Data Augmentation*, acreditamos que esse processo possa melhorar os resultados alcançados por essa pesquisa.

Referências

- AMARAL, Vagner do. *Extração e Comparação de Características Locais e Globais para o Reconhecimento Automático de Imagens de Faces*. Dissertação (Mestrado) — Centro Universitário da FEI, 2011.
- BAY, Herbert; TUYTELAARS, Tinne; GOOL, Luc Van. Surf: Speeded up robust features. In: LEONARDIS, Aleš; BISCHOF, Horst; PINZ, Axel (Ed.). *Computer Vision – ECCV 2006*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. p. 404–417. ISBN 978-3-540-33833-8.
- BERTOLINI, D.; OLIVEIRA, L. S.; SABOURIN, R. Multi-script writer identification using dissimilarity. In: *2016 23rd International Conference on Pattern Recognition (ICPR)*. Cancún, México: IEEE, 2016. p. 3025–3030.
- CAMPOS, Teófilo Emídio de; BABU, Bodla Rakesh; VARMA, Manik. Character recognition in natural images. *Proceedings of the 4th International Conference on Computer Vision Theory and Applications*, v. 2, p. 273–280, 01 2009.
- DUDA, Richard O.; HART, Peter E.; STORK, David G. *Pattern Classification (2Nd Edition)*. New York, NY, USA: Wiley-Interscience, 2000. ISBN 0471056693.
- FAUNDEZ-ZANUY, M. Signature recognition state-of-the-art. *IEEE Aerospace and Electronic Systems Magazine*, v. 20, n. 7, p. 28–32, July 2005. ISSN 0885-8985.
- GONZALEZ, Rafael C; WOODS, Richard E. *Digital Image Processing*. 2. ed. Upper Saddle River, NJ, USA: Prentice Hall, 2002. 793 p. ISBN 0201180758.
- GONÇALVEZ, Diego Bertolini. *Agrupamento de classificadores na Verificação de Assinaturas*. 2008. Dissertação de Mestrado, Pontifícia Universidade Católica do Paraná, Curitiba, Brazil.
- GOOL, L. Van; DEWAELE, P.; OOSTERLINCK, A. Texture analysis anno 1983. *Computer Vision, Graphics, and Image Processing*, v. 29, n. 3, p. 336 – 357, 1985. ISSN 0734-189X. Disponível em: <<http://www.sciencedirect.com/science/article/pii/0734189X85901306>>.
- GOVINDAN, V.K; SHIVAPRASAD, A.P. Character recognition — a review. *Pattern Recognition*, v. 23, n. 7, p. 671 – 683, 1990. ISSN 0031-3203. Disponível em: <<http://www.sciencedirect.com/science/article/pii/003132039090091X>>.
- HANNAD, Yaacoub; SIDDIQI, Imran; KETTANI, Mohamed El Youssfi El. Writer identification using texture descriptors of handwritten fragments. *Expert Systems with Applications*, v. 47, p. 14 – 22, 2016. ISSN 0957-4174. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0957417415007551>>.
- HARALICK, R. M. Statistical and structural approaches to texture. *Proceedings of the IEEE*, v. 67, n. 5, p. 786–804, May 1979. ISSN 0018-9219.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3, n. 6, p. 610–621, Nov 1973. ISSN 0018-9472.

HOCHBERG, Judith; BOWERS, Kevin; CANNON, Michael; KELLY, Patrick. Script and language identification for handwritten document images. *International Journal on Document Analysis and Recognition*, v. 2, n. 2, p. 45–52, Dec 1999. ISSN 1433-2833. Disponível em: <<https://doi.org/10.1007/s100320050036>>.

MUKHOPADHYAY, Anirban; SINGH, Pawan Kumar; SARKAR, Ram; NASIPURI, Mita. A study of different classifier combination approaches for handwritten indic script recognition. *Journal of Imaging*, v. 4, n. 2, 2018. ISSN 2313-433X. Disponível em: <<http://www.mdpi.com/2313-433X/4/2/39>>.

OBAIDULLAH, Sk Md; HALDER, Chayan; SANTOSH, K. C.; DAS, Nibaran; ROY, Kaushik. Phdindic_11: page-level handwritten document image dataset of 11 official indic scripts for script identification. *Multimedia Tools and Applications*, v. 77, n. 2, p. 1643–1678, Jan 2018. ISSN 1573-7721. Disponível em: <<https://doi.org/10.1007/s11042-017-4373-y>>.

OBAIDULLAH, Sk Md; SANTOSH, K. C.; HALDER, Chayan; DAS, Nibaran; ROY, Kaushik. Automatic indic script identification from handwritten documents: page, block, line and word-level approach. *International Journal of Machine Learning and Cybernetics*, v. 10, n. 1, p. 87–106, Jan 2019. ISSN 1868-808X. Disponível em: <<https://doi.org/10.1007/s13042-017-0702-8>>.

OJALA, T.; PIETIKAINEN, M.; MAENPAA, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 24, n. 7, p. 971–987, July 2002.

OJANSIVU, Ville; HEIKKILÄ, Janne. Blur insensitive texture classification using local phase quantization. In: ELMOATAZ, Abderrahim; LEZORAY, Olivier; NOUBOUD, Fathallah; MAMMASS, Driss (Ed.). *Image and Signal Processing*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. p. 236–243. ISBN 978-3-540-69905-7.

OTSU, N. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, v. 9, n. 1, p. 62–66, Jan 1979.

PAL, Umapada; JAYADEVAN, Ramachandran; SHARMA, Nabin. Handwriting recognition in indian regional scripts: A survey of offline techniques. ACM, New York, NY, USA, v. 11, n. 1, p. 1:1–1:35, mar. 2012. ISSN 1530-0226. Disponível em: <<http://doi.acm.org/10.1145/2090176.2090177>>.

PAL, U.; SINHA, S.; CHAUDHURI, B. B. Multi-script line identification from indian documents. In: *Seventh International Conference on Document Analysis and Recognition, 2003. Proceedings*. Edinburgh, UK: IEEE, 2003. p. 880–884.

PASSOS, Yuri. Uso de técnicas de reconhecimento de padrões para a detecção de logotipos em imagens. 06 2019.

RAJPUT, G. G.; UMMAPURE, S. B. Script identification from handwritten document images using lbp technique at block level. In: *2019 International Conference on Data Science and Communication (IconDSC)*. Bangalore, India: IEEE, 2019. p. 1–6.

SALVI, Dhaval. *Document Image Analysis Techniques for Handwritten Text Segmentation, Document Image Rectification and Digital Collation*. Tese (Doutorado) — University of South Carolina - Columbia, 2014.

SINGH, Pawan Kumar; SARKAR, Ram; DAS, Nibaran; BASU, Subhadip; KUNDU, Mahantapas; NASIPURI, Mita. Benchmark databases of handwritten bangla-roman and devanagari-roman mixed-script document images. *Multimedia Tools and Applications*, v. 77, n. 7, p. 8441–8473, Apr 2018. ISSN 1573-7721. Disponível em: <<https://doi.org/10.1007/s11042-017-4745-3>>.

SINGH, Pawan Kumar; SARKAR, Ram; NASIPURI, Mita. Offline script identification from multilingual indic-script documents: A state-of-the-art. *Computer Science Review*, v. 15-16, p. 1 – 28, 2015. ISSN 1574-0137. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1574013714000203>>.

SINGH, P. K.; SARKAR, R.; NASIPURI, M.; DOERMANN, D. Word-level script identification for handwritten indic scripts. In: *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*. Tunis, Tunisia: IEEE, 2015. p. 1106–1110.

SINGHAL, V.; NAVIN, N.; GHOSH, D. Script-based classification of hand-written text documents in a multilingual environment. In: *Proceedings. Seventeenth Workshop on Parallel and Distributed Simulation*. Hyderabad, India: IEEE, 2003. p. 47–54.

SOUZA, Marcos Roberto e. *Uma análise de diferentes abordagens para o reconhecimento off-line de estilos de escrita em documentos manuscritos*. 2015. Monografia, Universidade Tecnológica Federal do Paraná, Campo Mourão, Brazil.

SOUZA, Marcos Roberto e; BERTOLINI, Diego; ; PEDRINI, Helio; COSTA, Yandre M. G. Offline handwritten script recognition based on texture descriptors. TO APPEAR. 2019.

THEODORIDIS, S; KOUTROUMBAS, Konstantinos. Pattern recognition (theodoridis, s. and koutroumbas, k.; 2006). *Neural Networks, IEEE Transactions on*, v. 19, p. 376–376, 03 2008.

UKIL, Soumya; GHOSH, Swarnendu; OBAIDULLAH, Sk Md; SANTOSH, K. C.; ROY, Kaushik; DAS, Nibaran. Deep learning for word-level handwritten indic script identification. *CoRR*, abs/1801.01627, 2018. Disponível em: <<http://arxiv.org/abs/1801.01627>>.

VAPNIK, Vladimir N. *The Nature of Statistical Learning Theory*. Berlin, Heidelberg: Springer-Verlag, 1995. ISBN 0-387-94559-8.

WU, P.; MANJUNATH, B.S.; NEWSAM, S.; SHIN, H.D. A texture descriptor for browsing and similarity retrieval. *Signal Processing: Image Communication*, v. 16, n. 1, p. 33 – 43, 2000. ISSN 0923-5965. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0923596500000163>>.