

UNIVERSIDADE FEDERAL DE SANTA CATARINA
Centro Tecnológico
Departamento de Engenharia Elétrica e Eletrônica



PROJETO NÍVEL I EM CONT E PROC DE SINAIS I

ALEXANDRE SOLI
VINICIUS CIM

FLORIANÓPOLIS

2021

ALEXANDRE SOLI
VINICIUS CIM

**RECONHECIMENTO DE CARACTERES (OCR) EM FORMULÁRIO PADRÃO DA
UFSC**

Relatório de projeto proposto (OCR) para a
disciplina Projeto Nível I em Cont e Proc de
Sinais I.

Universidade Federal de Santa Catarina –
Campus Florianópolis.

Professor: Dr. Joceli Mayer

Florianópolis, 22 de setembro de 2021.

FLORIANÓPOLIS
2021

RESUMO

Neste trabalho é descrito um procedimento para a detecção e separação de caracteres em um padrão de formulário para validação de disciplinas da UFSC e sua classificação em número de 0 a 9 ou letras maiúsculas de A a Z por meio de redes neurais profundas, que utiliza como entrada descritores extraídos de forma clássica para cada caractere. Também será vista uma breve introdução da teoria utilizada. Por fim serão mostrados os resultados e as dificuldades enfrentadas, tão bem quanto sugestões e propostas de melhorias que podem ser feitas no projeto e em continuações do trabalho.

ABSTRACT

This work describes a procedure for detection and separation of characters in a default formulary for validation of UFSC disciplines and their classification in number from 0 to 9 or capital letters from A to Z through deep neural networks, which uses as input, classical descriptors extracted for each character. A brief introduction to the theory used will also be seen. Finally, the results and difficulties faced will be shown, as well as suggestions and proposals for improvements that can be made in the project and in continuations of the work.

SUMÁRIO

INTRODUÇÃO	4
METAS E METODOLOGIA	4
TEORIA ABORDADA	5
FORMULÁRIO	6
OBTENÇÃO DOS CARACTERES	8
TREINAMENTO DE REDES	9
CRIAÇÃO DE DATASETS	12
CRIAÇÃO DE FORMULÁRIO DIGITAL	13
DIFICULDADES	14
MELHORIAS PROPOSTAS E SUGESTÕES PARA CONTINUAÇÃO DO PROJETO	14
CONCLUSÃO	15
REFERÊNCIAS BIBLIOGRÁFICAS	16

1. INTRODUÇÃO

Os aplicativos digitalizadores são práticos pois implementam a funcionalidade de um scanner de alta qualidade na palma de nossas mãos. O problema é quando se deseja obter o conteúdo escrito de uma região específica, seja o nome do autor do documento, a data ou qualquer frase de interesse. Nessas situações não é possível fazer a edição do documento pois o mesmo é uma foto, uma imagem estática sem a informação digital dos caracteres, ou dígitos, presentes. Sem a interpretação dessas imagens, também é inviável para as máquinas interpretarem seu conteúdo.

Esse problema é parcialmente contornável com a aplicação de técnicas de Reconhecimento Ótico de Caracteres (OCR - em inglês). Essas técnicas variam de aplicação para aplicação, podendo ser utilizadas tanto técnicas clássicas no domínio de processamento de imagens no que diz respeito a detecção de bordas, descritores e operações morfológicas tanto quanto técnicas mais modernas como redes neurais convolucionais. Muitos sistemas atuais aplicam OCR em seus processos, como os Correios e os Bancos para fins de detecção de caracteres, muitas vezes impressos e escritos a mão.

2. METAS E METODOLOGIA

Metas

As metas que foram propostas para este trabalho são de realizar o reconhecimento de caracteres e símbolos extraídos de um formulário pré definido, dos quais os caracteres e símbolos são escritos a mão em formato caixa alta, com um espaçamento mínimo (escrita não cursiva) e símbolos bem definidos (@, -, . , _), transcrever textos no formato de imagem para formato de caracteres digitais, onde é possível realizar alterações e métodos para obtenção de informações referente ao texto digitalizado, como extração de campos de assunto, título, data, tópicos e aspectos do texto.

Metodologias

A metodologia utilizada é a construção de uma pipeline para separação, extração de descritores e classificação de caracteres utilizando-se de conceitos de processamento de imagens como binarização, aplicação de filtros, operações morfológicas, detecção de bordas, histograma de gradientes orientados (HOG), Zoning, histogramas, outras técnicas e conceitos de machine learning como redes profundas, redes convolucionais, transferência de aprendizado, *fine tuning*, técnicas de regularização, dentre outras. Para obter um bom aprendizado durante a execução do projeto, foi evitado o uso de ferramentas prontas para extração ou detecção de caracteres, como a ferramenta Tesseract, que não foi utilizada neste projeto.

Inicialmente foi realizada uma pesquisa bibliográfica para tentar entender os problemas e possíveis abordagens, os trabalhos consultados estão referenciados no final deste trabalho. A

programação foi realizada na linguagem de Python 3, onde os autores trabalharam na criação de *packages* contendo classes e funções, desta forma proporcionando o compartilhamento dos códigos já desenvolvidos entre os membros da equipe. De forma a manter o trabalho bem organizado foi montado um servidor, onde a comunicação entre a equipe foi dada e foi possível o compartilhamento de dúvidas, soluções e resultados.

Ferramentas

Bibliotecas (Python 3): Numpy (Fornece uma manipulação conveniente e rápida de arrays N-dimensionais.), OpenCV (Consiste de ferramentas para trabalhar com processamento de imagens), Imutils (Oferece uma série de funções convenientes para processamento de imagem, como translação, rotação, redimensionamento etc), Tensorflow/Keras (Uma das bibliotecas mais completas e utilizadas para aprendizado de máquina, aplicável a uma ampla variedade de tarefas), Scipy (Feita para lidar com trabalhos matemáticos, sendo muito útil para cientistas e engenheiros), Scikit-image (Para processamento de imagens) e Pillow (Adiciona suporte à abertura e gravação de vários formatos de imagem, dentre outras ferramentas).

Datasets: MNIST (É um grande banco de dados de dígitos manuscritos comumente usado para treinar vários sistemas de processamento de imagem, contém 60.000 imagens de treinamento e 10.000 imagens de teste), EMNIST (É um conjunto de dígitos de caracteres manuscritos derivados do NIST Special Database 19 e convertidos para um formato de imagem de 28x28 pixels, consiste de 145600 imagens de letras.), Dataset próprio (Foram criados 6776 caracteres, onde 3058 são números e 3718 letras com as variações mais comuns).

Plataformas: Jupyter lab (É uma interface estilo notebook para programação em Python, muito versátil para realização de testes e organização de código.), Google Colab (Similar ao Jupyter lab, porém roda na nuvem e oferece GPU para trabalhos com machine learning).

3. TEORIA ABORDADA

Neste tópico, um rápido resumo da teoria é realizado, iniciando com a operação de binarização, que consiste em aplicar um limiar para o valor de uma imagem grayscale, com finalidade de binarizar seus valores, por exemplo 0 e 1.

Histograma, pode ser feito tanto no eixo horizontal como no vertical, consiste de obter a frequência de cada valor na imagem, quando aplicado a uma imagem binária, tem-se como resultado o número de pixels não nulos em cada coluna ou linha.

Operações morfológicas, dilatação, erosão, abertura, fechamento:

O processo de dilatação adiciona pixels aos limites dos objetos em uma imagem, de acordo com o elemento estruturante, que é o elemento que irá percorrer a imagem alvo, quando uma intersecção

ocorre entre o elemento estruturante e a imagem, o pixel do centro do elemento estruturante é adicionado ao conjunto solução. Já o processo de erosão é análogo à dilatação, porém o elemento estruturante deve estar contido na imagem, isto resulta na exclusão dos pixels das bordas deixando a imagem mais “fina”.

A abertura é composta de uma erosão seguida por uma dilatação resultando em uma imagem mais suave, o processo elimina pontas fazendo um alisamento, desta forma a imagem fica menor. Já o fechamento é a sequência inversa, uma dilatação seguida de uma erosão o que resulta no fechamento de buracos da imagem, neste processo a região da imagem fica maior.

O histograma de gradientes orientados (HOG) é um histograma que conta as ocorrências de certas orientações de gradientes em fatias localizadas de uma imagem, funciona de forma semelhante a detecção de bordas, a diferença é que o cálculo é realizado em uma grade de pixels uniformemente espaçados.


Zoning, é um método no qual é possível reduzir informações de uma imagem, porém manter a informação espacial, isso é feito através da passagem de um kernel pela imagem, que retorna a média dos pixels da imagem cujo está sobreposta pelo kernel no momento, assim uma nova imagem menor é formada.

Os conceitos relacionados a machine learning serão abordados em mais detalhes no tópico treinamento de redes.

4. FORMULÁRIO

O formulário utilizado possui as dimensões de uma folha A4 padrão, 21 x 29,7 cm, 2480 x 3508 pixels, consiste dos campos: Curso, Nome, E-Mail, CPF, Matricula, Telefone 1, Telefone 2, Ex-Alunos da UFSC - Matrícula anterior, 26 campos para preenchimento do código da disciplina cursada e código da disciplina a validar na UFSC, Data e um campo para assinatura, este formulário é fornecido pelo departamento de administração escolar (DAE) e é padrão para validação de disciplinas na UFSC.

Figura 4.1: Formulário de validação de disciplinas da UFSC.

 SERVIÇO PÚBLICO FEDERAL MINISTÉRIO DA EDUCAÇÃO UNIVERSIDADE FEDERAL DE SANTA CATARINA PRÓ-REITORIA DE GRADUAÇÃO DEPARTAMENTO DE ADMINISTRAÇÃO ESCOLAR																																																																																																																										
VALIDAÇÃO DE DISCIPLINAS																																																																																																																										
CURSO:																																																																																																																										
NOME:																																																																																																																										
E-MAIL:	CPF:	MATRÍCULA:																																																																																																																								
TELEFONE 1:	TELEFONE 2:																																																																																																																									
EX-ALUNO DA UFSC - MATRÍCULA ANTERIOR:																																																																																																																										
<p>ANEXAR:</p> <ul style="list-style-type: none"> - Histórico escolar onde constam as disciplinas cursadas; - Programa/Plano de ensino das disciplinas, se cursadas em outra Instituição; - Certificados de seminários ou cursos, quando for solicitação de validação de atividades complementares. <p>ATENÇÃO:</p> <ul style="list-style-type: none"> - Ao preencher este formulário, indicar a disciplina da UFSC a validar com o código da disciplina, conforme consta no currículo, composto por três letras e quatro dígitos. Exemplo: ECZ7031. O currículo com os códigos das disciplinas está disponível na página do seu Curso ou poderá ser solicitado através de e-mail. 																																																																																																																										
<p style="text-align: center; font-weight: bold; margin-bottom: 10px;"><u>REQUER VALIDAÇÃO DAS DISCIPLINAS</u></p> <table style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 25%; text-align: center; font-size: small;">COD. DISCIPLINAS CURSADAS</th> <th style="width: 5%; text-align: center; font-size: small;">—</th> <th style="width: 25%; text-align: center; font-size: small;">COD. DISCIPLINAS DA UFSC A VALIDAR</th> <th style="width: 5%; text-align: center; font-size: small;">—</th> <th style="width: 25%; text-align: center; font-size: small;">COD. DISCIPLINAS CURSADAS</th> <th style="width: 5%; text-align: center; font-size: small;">—</th> <th style="width: 25%; text-align: center; font-size: small;">COD. DISCIPLINAS DA UFSC A VALIDAR</th> </tr> </thead> <tbody> <tr><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td></tr> <tr><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td></tr> <tr><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td></tr> <tr><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td></tr> <tr><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td></tr> <tr><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td></tr> <tr><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td></tr> <tr><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td></tr> <tr><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td></tr> <tr><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td></tr> <tr><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td></tr> <tr><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td></tr> <tr><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td></tr> <tr><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td></tr> <tr><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td></tr> <tr><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td><td>—</td><td><div style="border: 1px solid black; height: 20px; width: 100%;"></div></td></tr> </tbody> </table> <p style="font-size: x-small; margin-top: 10px;">* Declaro que os documentos que anexarei junto ao formulário são cópias fiéis dos documentos autênticos, de acordo com os termos e responsabilizações cabíveis do Decreto Federal nº 8.539/2015</p> <div style="display: flex; justify-content: space-between; margin-top: 20px;"> <div style="width: 45%; text-align: center;"> <div style="border: 1px solid black; height: 20px; width: 100%; margin-bottom: 5px;"></div> <div style="text-align: left; font-size: x-small;">DATA</div> </div> <div style="width: 45%; text-align: center;"> <div style="border: 1px solid black; height: 20px; width: 100%; margin-bottom: 5px;"></div> <div style="text-align: left; font-size: x-small;">ASSINATURA</div> </div> </div>				COD. DISCIPLINAS CURSADAS	—	COD. DISCIPLINAS DA UFSC A VALIDAR	—	COD. DISCIPLINAS CURSADAS	—	COD. DISCIPLINAS DA UFSC A VALIDAR	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>
COD. DISCIPLINAS CURSADAS	—	COD. DISCIPLINAS DA UFSC A VALIDAR	—	COD. DISCIPLINAS CURSADAS	—	COD. DISCIPLINAS DA UFSC A VALIDAR																																																																																																																				
<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>																																																																																																																				
<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>																																																																																																																				
<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>																																																																																																																				
<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>																																																																																																																				
<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>																																																																																																																				
<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>																																																																																																																				
<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>																																																																																																																				
<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>																																																																																																																				
<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>																																																																																																																				
<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>																																																																																																																				
<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>																																																																																																																				
<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>																																																																																																																				
<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>																																																																																																																				
<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>																																																																																																																				
<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>																																																																																																																				
<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>	—	<div style="border: 1px solid black; height: 20px; width: 100%;"></div>																																																																																																																				

Fonte: Departamento de administração escolar (DAE) UFSC.

5. OBTENÇÃO DOS CARACTERES

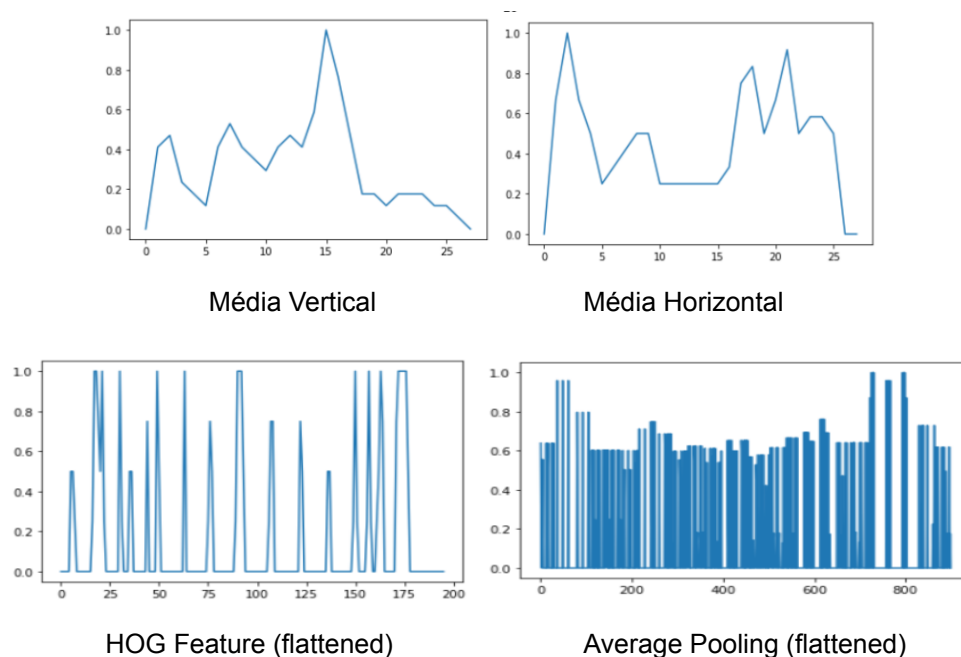
A obtenção dos caracteres foi realizada utilizando vários dos conceitos vistos na disciplina e citados na seção de metas e metodologia, inicialmente, com a imagem do formulário foi feito um mapeamento de coordenadas dos pixels dos quatro cantos de cada campo, de forma a recortar a imagem nestas coordenadas, obtendo assim uma imagem para cada um dos 61 campos, os quais são guardados em uma lista, a seguir, para cada uma das 61 imagens é realizado um pipeline para separação dos caracteres. Inicialmente a imagem RGB é convertida para *grayscale* e binário utilizando um *threshold* adaptativo médio, onde o valor é a média da área da vizinhança, este método foi utilizado pois apresentou melhores resultados do que o método Gaussiano, antes da binarização em si, é feita a aplicação de um filtro passa baixas (Blur) com *kernel* 17x17 de modo a diminuir a variação de cores e obter melhores resultados. Com a imagem binarizada, é realizada a operação de close, com um elemento estruturante retangular 300x12, o que faz com que ocorra a junção dos caracteres na horizontal, através dessa imagem gerada são retirados os contornos com a função *findcontours* da biblioteca OpenCV e então a imagem de entrada é recortada com base nestas coordenadas, assim obtemos cada linha separadamente.

Note que este processo de separação de linhas ocorre quando a imagem de entrada tem mais de uma linha, esta implementação foi realizada para utilização em formulários com esta característica, porém não é utilizado no formulário proposto neste trabalho. A seguir é realizada a separação de palavras utilizando a operação de close novamente, porém agora com um elemento estruturante de um tamanho otimizado para juntar os caracteres próximos porém não juntar palavras/clusters de caracteres, novamente são retirados os contornos e a imagem original é recortada, fazendo assim que se obtenha uma lista com as palavras de cada linha, por fim para obter cada caractere de cada palavra/cluster de caracteres, é realizado um close na vertical, mas com um elemento estruturante “fino”, o que não junta caracteres, desta forma se recorta as imagens nos contornos retornados e obtém-se uma lista que contém linhas, palavras/clusters para cada linha e caracteres de cada palavra/cluster. Vale ressaltar que após cada utilização da função *findcontours*, os contornos retornados passam por outra função, onde são organizados de forma a manter a estrutura do texto, esta função então implementa a organização nas direções esquerda para direita (quando são palavras/clusters e caracteres) e do topo para baixo (quando são linhas).

Com os caracteres separados, a lista é iterada por mais três funções, onde ocorrem os seguintes processos:

- Remoção de espaços em branco, esta etapa é feita para evitar que a imagem fique muito no canto quando um *resize* é realizado, a função consiste de operações de detecção de bordas e fechamento.
- Reshape, as imagens obtidas dos caracteres escritos a mão não possuem as mesmas dimensões, e a simples aplicação direta de um *resize* acarreta na distorção de muitos dos caracteres o que dificulta muito a sua identificação posteriormente, para contornar este problema foi realizado um “reshape”, onde ocorre a identificação da maior dimensão da

Figura 6.1: Descritores



A abordagem inicial desse projeto foi clássica, mas logo se percebeu acurácias muito baixas devido a grande diversidade de traços e tamanhos que um único caractere pode assumir. Passou-se então para o uso de RNPs, o motivo, além do descrito acima é que esse tipo de rede ainda possibilita o uso de descritores, que é a base da proposta desse projeto.

Ensinar um computador na área de redes neurais é o mesmo que treinar um modelo. Esse modelo pode ser construído e treinado pela primeira vez ou pode ser adaptado de uma outra aplicação, o que é conhecido como *transfer learning*, ou transferência de aprendizado em português, essas duas técnicas dependem do objetivo, complexidade e tempo de execução de um projeto.

Um modelo novo é construído com base em referências de outros autores e experimentações diversas, já um modelo vindo de um processo de transfer learning precisa seguir alguns padrões que a literatura aconselha. Primeiro, transfer learning só faz sentido quando as aplicações apresentam semelhanças, como por exemplo, uma rede conceituada em classificar cães e gatos pode ser utilizada em transfer learning para classificar lobos e lince. Essa técnica é muito poderosa pelo fato da nova rede se adequar aos novos dados introduzidos muito rapidamente, pois como intuitivamente se sabe, cães e gatos têm certa similaridade com lobos e lince respectivamente. Em segundo lugar, provou-se eficaz primeiro “congelar” todas as camadas do modelo importado excluindo-se a última camada, o que significa que os pesos dos neurônios dessas camadas anteriores não se alteram, isto é, o modelo retém sua estrutura base e apenas a nova camada se adequa às mudanças introduzidas pelo novo problema. Em terceiro lugar, caso a rede não esteja convergindo bem, não esteja tendo bons resultados, é aconselhável o “descongelamento” gradativo de camada por camada, do final para

o início, com uma taxa de aprendizado muito baixa, isso garante que os pesos não se alterem drasticamente preservando o que já foi aprendido

Três redes foram utilizadas, uma para classificação de números (0 a 9), uma para classificação de letras maiúsculas (0 a 25) e uma para classificação binária referente à diferenciação de números e letras (0 e 1),

As duas redes utilizadas respectivamente para classificação de números e classificação de letras são RNPs, chamemos de REDE 1 e REDE 2 cada uma delas. Ambas usam como descritores médias verticais, médias horizontais, *Zoning* e *HOG Feature* que são concatenados em forma de um vetor de descritores achatados (Flatten). Esses vetores são passados na entrada das redes com tamanhos respectivos de 1005 (para números) e 2181 (para letras) unidades cada. A experimentação fez-se concluir que para a Rede 2 eram necessárias mais unidades, justamente pelo fato de ser uma classificação com maior quantidade de possíveis resultados (0 a 25).

Para a classificação binária de números e letras foi necessário a aplicação de uma rede que conseguisse escolher os melhores descritores automaticamente, particularmente uma rede convolucional, RNC, chamemos ela de REDE 3. Isso foi preciso pois existem caracteres desses dois grupos manuscritos que se confundem muito, como os pares ["0", "O"], ["1", "l"], ["2", "Z"], ["5", "S"], ["6", "G"] e ["8", "B"]. As redes convolucionais apresentam a capacidade de reter muito mais informações espaciais do que redes neurais profundas convencionais pelo fato de iterativamente aprenderem quais os melhores descritores para se observar no modelo. Como a proposta desse trabalho é usar features escolhidas manualmente, apenas foi implementado redes convolucionais nessa etapa mais crítica. Abaixo algumas dessas semelhanças para o leitor observar.

Figura 6.2: Exemplos



Fonte: Autores

O primeiro método de construção de modelos seguiu a seguinte ordem:

1. Criação de um modelo para os datasets referência, MNIST e EMNIST. Tanto as duas RNPs como a RNC foram construídas e refinadas pelos autores.
2. Uso dos modelos para prever (testar) os caracteres vindos do processamento do formulário.

Porém essa tentativa foi falha, pois apesar desses datasets serem bem conceituados e terem muitos dados para serem treinados eles carecem de uma diversidade caligráfica como por exemplo a falta de traços horizontais usuais em caracteres como o número "1", "7", e as letras "G" e "Z", quando são escritos manualmente. Para isso, foi pensado no uso de transfer learning, o que necessitaria da

criação de um dataset próprio (descrição no próximo tópico). A criação de modelo para esse uso seguiu o seguinte fluxo:

1. Criação de um modelo para os datasets referência, MNIST e EMNIST, que tenha bons resultados nos seus próprios conjuntos de testes.
2. Importação desses modelos e consequente transfer learning para dataset próprio. O que inclui descongelamento gradativo, adição/remoção de novas camadas, como camadas de regularização para evitar *overfitting*, ou simplesmente camadas com mais neurônios. Se inclui também o uso de uma taxa de aprendizado muito baixa para evitar que a rede perca sua estrutura previamente aprendida. Cabe destacar que essas decisões são as mesmas tomadas quando se importa uma rede conceituada em outra área, por isso a denominação de transfer learning.
3. Uso dos modelos modificados para prever (testar) os caracteres vindos do processamento do formulário.

Com essa técnica, foi obtido acurácias muito maiores, beirando a faixa de 99% tanto para classificação de números quanto para letras. O uso de transfer learning se seguiu da mesma forma para a rede convolucional de classificação binária. Abaixo uma tabela que relaciona os modelos finais.

Tabela 6.1: Acurácias obtidas

	REDE 1 (Números)	REDE 2 (Letras)	REDE 3 (Números x Letras)
Acurácia no conjunto de teste	99%	98%	98%

7. CRIAÇÃO DE DATASETS

Foram criadas 3058 imagens de números, 3718 de letras maiúsculas . Essas imagens foram criadas com o auxílio de mesas digitalizadoras e também obtidas com a participação de familiares e colegas (caracteres escritos em papel com caneta) para então conseguir uma boa diversidade de traços.

Figura 7.1: Caracteres escritos em mesa digitalizadora



Fonte: Autores

Usando as funções disponíveis no package de pré processamento desenvolvido pela equipe, foram passadas imagens de cada caractere. Além disso, as imagens geradas são caracteres de dimensão 28x28, são binarizadas e com tonalidade invertida (fundo preto e traço branco). Abaixo alguns exemplos desses caracteres.

Figura 7.2: Caracteres próprios gerados



Fonte: Autores

As dimensões de 28x28 foram escolhidas pois é a dimensão utilizada pelos datasets de referência, MNIST (para números) e EMNIST (para letras), assim possibilitando tanto a técnica de transfer learning como a comparação com o treinamento do zero. Além disso, as imagens são binarizadas e com as cores invertidas, entre valores de 0 a 255 porque o pré-processamento do formulário inclui a binarização das imagens dessa forma para melhor segmentação das mesmas.

8. CRIAÇÃO DE FORMULÁRIO DIGITAL

Nesta etapa, após ter sido realizada a predição de cada caractere, uma lista é gerada contendo o resultado para cada campo, através de uma função de escrita, as respostas são correlacionadas com seus campos e um arquivo txt é gerado, o qual poderá futuramente ser utilizado para uma integração com um formulário digital, através da utilização de php e linguagens de banco de dados por exemplo, desta forma é possível economizar tempo na validação de disciplinas e implementar um sistema automatizado.

Figura 8.1: Arquivo .txt resultante com 92,8% de acurácia.

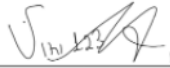
Curso: ENGENHARIA ELETRONICA
 Nome: VINICIUS CIM
 Email: V1N1123
 CPF: 12345678
 Matrícula: 19250114
 Telefone 1: 40028922
 Telefone 2: 999666123456789
 Matrícula Anterior: 2021 1 202
 Cursada: FSC02 Validar: MTM5
 Cursada: FSC028 Validar: EE7511
 Cursada: EEL9876 Validar: INE123
 Cursada: 1MC22 Validar: QMC93
 Cursada: XYZ013 Validar: ABC57
 Cursada: EEL01 Validar: EEL0Q
 Cursada: KLM612 Validar: OICR27
 Data 21 09 21

CURSO: ENGENHARIA ELETROICA			
NOME: VINICIUS CIM			
E-MAIL: V1N1123		CPF: 12345678	MATRÍCULA: 19250114
TELEFONE 1: 40028922		TELEFONE 2: 999666123456789	
EX-ALUNO DA UFSC - MATRÍCULA ANTERIOR: 2021 1 202			

COD. DISCIPLINAS CURSADAS		COD. DISCIPLINAS DA UFSC A VALIDAR		COD. DISCIPLINAS CURSADAS		COD. DISCIPLINAS DA UFSC A VALIDAR
FSC02	—	MTM5		FSC028	—	EE7511
EEL9876	—	INE123		INE346	—	INE92
QMC22	—	QMC93		XYZ03	—	ABC57
EEL01	—	EEL02		KLM612	—	OCR22
	—				—	
	—				—	

21 09 21

DATA



ASSINATURA

Fonte: Autores

9. DIFICULDADES

Neste tópico citamos um resumo das dificuldades encontradas em forma de lista.

1. Separação de caracteres, nesta etapa percebeu-se que realizar a separação de caracteres escritos a mão de forma livre não é uma tarefa trivial, e que seria necessário mais tempo de projeto e técnicas mais avançadas para abarcar estilos de escrita com letras mais próximas, desta forma foi necessária a limitação do escopo do projeto para números e letras maiúsculas em caixa alta e com um espaçamento mínimo determinado.
2. Classificação entre números e letras, percebeu-se neste projeto que existe uma dificuldade na diferenciação de algumas letras e números, e também entre algumas letras, como por exemplo a letra {O e 0}, {I e 1}, {S e 5}, {B, 8}, {Z, 2}, {Q, 0}, {V, U}. Na próxima seção são comentadas algumas possibilidades para estes casos.

10. MELHORIAS PROPOSTAS E SUGESTÕES PARA CONTINUAÇÃO DO PROJETO

Começando pela extração dos caracteres, é proposto que sejam estudadas técnicas mais avançadas para separação de caracteres, de modo a deixar as restrições de escrita mais livres e que

seja utilizado algum método para corrigir distorções, o que facilitaria na parte de classificação.

Outra proposta é a realização de um estudo aprofundado para entender e escolher melhores descritores, que tenham mais robustez na diferenciação de caracteres, alguns dos descritores vistos na literatura que não foram implementados neste projeto são: Descritor de direção de traço, de perfil do caractere, de distâncias de cruzamento de traços, descritores estruturais que contabilizam pontos de cruzamento, pontos finais e loops e descritores que utilizam a transformada de Fourier, como sua aplicação nos contornos dos caracteres. Mais detalhes destes e outros descritores podem ser visto nos trabalhos *Optical Handwritten Character/Numerical Recognition, An Overview of Feature Extraction Techniques in OCR for Indian Scripts Focused on Offline Handwriting* e *Feature extraction methods for character recognition | a survey*, citados nas referências bibliográficas.

A utilização de redes com memórias, como as redes recorrentes, que levam em consideração a sequência temporal das entradas, o que permitiria prever com mais precisão as próximas letras em uma palavra, como exemplo pode-se dizer que ao ter sido prevista a letra L, e tendo um nível de certeza de 50% de que o próximo caractere é a letra O e 50% de que é o número 0, tendo a informação da letra anterior (L) é mais provável que o próximo caractere seja a letra O.

Uma aplicação prática seria a criação de um aplicativo para celular que possa fazer a digitalização da imagem do formulário, sugere-se também que se altere a etapa de extração dos campos para poder receber formulários com dimensões diferentes, desta forma não é necessário que a folha seja escaneada e os formulários podem ser digitalizados com a câmera do celular do próprio estudante e enviado para um servidor onde o processamento será realizado.

11. CONCLUSÃO

Com os resultados desse projeto reitera-se a importante participação de descritores clássicos no âmbito de reconhecimento de caracteres e como a escolha dos mesmos é uma questão de revisão de literatura e experimentação. A obtenção dos descritores se baseia em técnicas e métodos vistos em teoria que ao serem implementados em uma linguagem de programação massificam o conteúdo exposto em sala de aula.

Quando aplicados na entrada de Redes Neurais, descritores como médias axiais de pixel, *Zoning* e *HOG Feature* são poderosos instrumentos para classificação de números e letras escritos à mão. O uso de um dataset relativamente pequeno permite o ajuste fino de um modelo treinado com maior capacidade, conclui-se também que é válido um estudo mais aprofundado para realizar comparações entre redes que utilizam descritores escolhidos manualmente e redes convolucionais, que aprendem seus próprios descritores.

12. REFERÊNCIAS BIBLIOGRÁFICAS

Pulak Purkait, **Optical Handwritten Character/Numeral Recognition**, ECSU, Indian Statistical Institute, Disponível em:
<https://www.isical.ac.in/~vlrg/sites/default/files/Pulak/Off-Line%20Handwritten%20OCR.pdf>
 Acesso em: 02 set. 2021.

Gaurav Y. Tawde , Mrs. Jayashree M. Kundargi, **An Overview of Feature Extraction Techniques in OCR for Indian Scripts Focused on Offline Handwriting**, K.J. Somaiya College of Engineering, Mumbai University, India, Disponível em:
<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.415.3676&rep=rep1&type=pdf>
 Acesso em: 02 set. 2021.

Joceli Mayer, **Conjunto de slides da disciplina EEL7815 - Projeto Nível I em Cont e Proc de Sinais I**, Universidade Federal de Santa Catarina - Campus Florianópolis.

Ivind Due Trier, Anil k. Jain and Torfinn Taxt, **Feature extraction methods for character recognition | a survey**, University of Oslo, Michigan State University, Disponível em:
<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.51.7439&rep=rep1&type=pdf>
 Acesso em: 09 set. 2021.

Gregory Cohen, Saeed Afshar, Jonathan Tapson, and Andre van Schaik, **EMNIST: an extension of MNIST to handwritten letters**, Western Sydney University Penrith, Australia, Disponível em: <https://arxiv.org/pdf/1702.05373v1.pdf> Acesso em: 09 set. 2021.

Satya Mallick, **Histogram of Oriented Gradients explained using OpenCV**, Ph.D. at the University of California, San Diego, Disponível em:
<https://learnopencv.com/histogram-of-oriented-gradients/> Acesso em: 19 Ago. 2021.