

Otimização de um reconhecedor óptico de caracteres baseado em redes neurais usando algoritmos genéticos

João Carlos Batista*

Instituto de Matemática e Computação
São Carlos, SP, Brazil
djoneuspiano@usp.br

ABSTRACT

O reconhecedor óptico de caracteres (OCR) surge da necessidade de economizar tempo na digitalização dos caracteres impressos nos documentos. O OCR captura a imagem como entrada, processa essa imagem e a converte em texto digitalizado. Esse sistema pode ser útil em vários aplicativos, como bancos, setor jurídico, outras indústrias e automação residencial e de escritório. Em geral, os OCRs usam o potencial das redes neurais e, para o qual usam como entradas os vetores de características das imagens previamente filtradas (separação e processamento). Hoje em dia continua uma busca constante para melhorar a eficiência dos OCRs, este trabalho propõe um modelo híbrido para reconhecimento óptico de caracteres que baseia em duas técnicas de inteligência artificial, sendo redes neurais e algoritmos genéticos. A rede neural foi utilizada para o reconhecimento dos caracteres e os algoritmos genéticos para otimizar o treinamento.

CCS CONCEPTS

• **Computer systems organization** → **Embedded systems**; *Redundancy*; Robotics; • **Networks** → Network reliability;

KEYWORDS

OCR, rede neural, algoritmo genético.

ACM Reference Format:

João Carlos Batista. 2018. Otimização de um reconhecedor óptico de caracteres baseado em redes neurais usando algoritmos genéticos. In *Proceedings of the Genetic and Evolutionary Computation Conference 2018 (GECCO '18)*. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

1 INTRODUÇÃO

Atualmente no mundo da informação, diversos documentos como são os relatórios, contratos e cartas que são gerados a cada dia, e portanto, há necessidade de arquivar, recuperar, atualizar, replicar e distribuir documentos impressos, cada vez mais torna-se importante e por este motivo, este tipo de sistemas surge na raiz da necessidade de economizar tempo na digitalização dos caracteres impressos nos documentos.

O Reconhecedor Óptico de Caracteres (OCR), onde os documentos impressos (após a digitalização) são transformados em arquivos ASCII, é um dos mais típicos mais antigo dentro do Reconhecimento de Padrões, e está ganhando popularidade por suas aplicações potenciais em escritórios, automação de bibliotecas[8], bancos, serviços postais, entre outros.

Atualmente, existem muitos softwares livres e aplicativos proprietários que fornecem solução para o problema de reconhecimento, dada a imagem digital, o texto correspondente a cada um dos caracteres, por isso a eficiência esta em constante melhoria:

- OmniPage (http://spain.scansoft.com/news/20011018_omnipage.asp)
- Proyecto Clara OCR (<http://www.claraocr.org/>)
- GNU Ocrad (<http://www.gnu.org/software/ocrad/ocrad.html>)
- GOCR 0.41 (<http://jocr.sourceforge.net/>)
- SimpleOCR 3.1 (<http://www.simpleocr.com/>)
- Cuneiform OCR 6.0 (<http://www.ocr.com/>)

Esses sistemas geralmente, para esse reconhecimento, usam o potencial das redes neurais e para isso precisam de entradas para os vetores característicos das imagens que antes precisavam ser filtrados, separados e processados de tal forma que as letras tivessem o mesmo tamanho e sem ruídos

Devido ao fato de que ainda há uma busca constante para melhorar a eficiência do OCR, o presente trabalho propõe um modelo híbrido para o reconhecimento óptico dos caracteres baseados em duas técnicas de inteligência artificial, como redes neurais e algoritmos genéticos.

No entanto, todos eles têm uma alta margem de erro quando a entrada são caracteres manuscritos ou quando a qualidade do documento e sua digitalização estão deteriorados, por essa razão que o reconhecimento de caracteres, continua em uma constante procura para melhorar sua eficiência. A rede neural será utilizada para o reconhecimento dos caracteres e os algoritmos genéticos para otimizar o treinamento.

O principal objetivo deste trabalho é desenvolver um modelo híbrido para o reconhecimento óptico de caracteres em documentos digitalizados. Dado que os OCRs ainda são uma fonte de pesquisa, por esse motivo o reconhecimento de caracteres continua em uma busca constante para melhorar sua eficiência. Destina-se a destacar os conhecimentos adquiridos durante a realização dos mesmos e irá lançar as bases para novos estudos que surjam a partir do problema acima mencionado.

O presente trabalho está organizado da seguinte forma, as técnicas, métodos e trabalhos relacionados com o modelo proposto são apresentadas na Seção 2. Na Seções 3, 4 e 5 são mencionadas algumas técnicas básicas que são utilizadas para o processamento

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
GECCO '18, July 15–19, 2018, Kyoto, Japan
© 2018 Copyright held by the owner/author(s).
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00
<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

de imagem, assim como é iniciado a proposta é dada para implementar um modelo híbrido. Na Seção 6 apresenta uma avaliação experimental mostrando os testes e os respectivos resultados e na Seção 7 as conclusões do trabalho, bem como algumas linhas futuras de trabalho. Na última parte são apresentadas as referências bibliográficas.

2 RECONHECIMENTO ÓPTICO DE CARACTERES

O reconhecimento de caracteres ópticos começou em 1951 com a invenção de M. Sheppard, de um robô leitor-escritor. Várias empresas, incluindo a IBM, Recognition Equipment, Inc., Farrington, Control Data e Optical Scanning Corporation, Eles comercializaram o Software OCR nos anos de 1967.

O reconhecimento óptico de caracteres (OCR) é um tipo de análise aplicado em imagem de texto, em que uma imagem digitalizada, contendo impressões de máquina ou manuscrito, que serve de entrada para o OCR software que a transforma em um formato de texto digitalmente legível e editável (texto ASCII) [27].

Para [19], o objetivo final de um sistema OCR é o geração de um arquivo com os códigos ASCII correspondentes os caracteres de um texto em papel cuja a imagem foi previamente capturado pelo scanner, portanto, consiste em uma etapa de isolamento dos objetos a serem reconhecidos (neste caso caracteres impressos), outro de extração de suas características individual e, finalmente, a aplicação de uma função discriminante que permite estabelecer a classe a que pertence cada objeto (neste caso, a classe de caracteres, que será representado pelo seu código ASCII correspondente).

A complexidade do sistema varia dependendo do tipo de caracteres que aparecem no documento. Nesse sentido, o seguintes variedades de menor a maior dificuldade [14]:

- (1) Reconhecimento de um ou mais tipos de letra de impressão (“fixed-font OCR” y “multifont OCR”)
- (2) De qualquer tipo de letra impressa (“omnifont OCR”)
- (3) De caracteres manuscritas isoladas (“handwritten OCR”) y
- (4) Escrita discursiva (“script recognition”)

2.1 Trabalhos relacionados

O reconhecimento óptico de caracteres é um dos tópicos mais antigos no reconhecimento de padrões que tem como um problema essencial o reconhecimento de caracteres manuscritos. São várias as aplicações que vem sendo estudados nos últimos anos, como por exemplo o trabalho de Rodriguez e colegas [10], da mesma forma [26] propôs o modelo conceitual para a automação de bibliotecas no contexto digital. Por outro lado, Santos-Torres [8] propôs uma análise de algoritmos para reconhecimento de manuscritos caracteres.

Hoje existem vários algoritmos que realizam essa tarefa, cada um com suas vantagens e desvantagens. Muitos desses métodos enfrentam o problema de reconhecer caracteres de diferentes orientações, como é através da heurística (*Active Character Recognition*) [17], que se concentra em sub-imagens geradas a partir de um pedaço de caráter, que de forma recursiva refina para alcançar os critérios de aceitação, o trabalho de Park et al. [17] apresenta o uso de uma função heurística ativa, com base em Algoritmo de busca A*, que determina de forma adaptativa o tamanho do vetor

de característica ao mesmo tempo em que é usado para classificar o padrão de entrada.

Nesta mesma área, na obra [20], com base no reconhecimento geométrica (estatística ou com base na teoria da decisão) e reconhecimento estrutural, um novo método de representação, chamada LLF é proposto (*Linha local de montagem* Local Line Fitting), para o reconhecimento de caracteres manuscritos.

Outra abordagem é colocar o problema de reconhecimento de padrões de texto por um sistema inteligente, do campo da inteligência artificial, redes neurais especificamente [4], em que a referida rede é treinada com número de dados de treinamento e possíveis resultados, abordagem que foi considerada neste trabalho.

Além disso, [19] sugere um sistema de OCR com o recurso de ‘aprender’, usando uma rede neural, padrões de caracteres que representam as possíveis variações (tamanho, tipo de letra) forma diferente de caracteres impressos podem aparecer nos documentos no futuro e com a mesma rede, para “reconhecer” e converter o texto escrito no papel armazenado em um arquivo de texto ASCII. Definindo um pouco mais tecnicamente, consiste em ter qualquer padrão T, a partir do qual espera-se que o sistema não o identifique como padrão X, porque T é mais similar ou tem mais características de X que qualquer outro padrão conhecido[13].

O uso de redes neurais permitiu obter resultados muito bons no reconhecimento de caracteres. A maioria dos trabalhos publicados menciona o uso de métodos clássicos de reconhecimento, tais como redes neurais do tipo Hopfield ou redes de *feed-forward* de múltiplas camadas (perceptron) treinadas com retro-propagação[23].

Essa arquitetura tem sido reconhecida como uma poderosa ferramenta para o reconhecimento de padrões, dado seu poder discriminativo e sua capacidade de aprender e representar o conhecimento implícito. Mas ante certos problemas reais este modelo pode apresentar dificuldades, como são determinar o tamanho da estrutura de rede apropriada para aproximar limites de decisão complexo e uma pobre generalização causada pela dimensão de entrada elevada[1, 25]. Outra dificuldade é a velocidade de convergência e a existência de mínimos locais quando a técnica de descida de gradiente é usada para ajustar os pesos [24].

Outra questão a considerar ao projetar uma arquitetura de rede é definir se a extração de características dos padrões de entrada é feita em um estágio de pré-aprendizagem (pré-processamento) [12], ou se será incluído como uma função da rede[24].

Com relação à rede neural do tipo Hopfield, essa rede não realizará um reconhecimento completo de padrões para uma possível extensão ao reconhecimento da caligrafia humana, devido à sua limitação de generalização [13].

Por outro lado, a Rede Neural Probabilística (RNP), que se baseia no cálculo da probabilidade a posteriori, é um classificador que utiliza o método da máxima probabilidade que executa uma modelagem Gaussiana das características de entrada sendo centradas nas amostras de treinamento. [3] A RNP não possui um estágio real de treinamento, mas percorre todo o conjunto classificado em cada execução, o qual inferioriza sua performance.

É possível obter resultados altamente competitivos no reconhecimento de caracteres impressos ou manuscritos, usando técnicas de aprendizado não supervisionadas, como os mapas auto-organizados de Kohonen [21, 22], e mesmo combinados com outras técnicas [6].

Métodos baseados em aprendizado não supervisionado e na estrutura adaptativa de redes [7, 11] também foram propostos.

Finalmente, existem outras técnicas que buscam alguma forma de otimizar a porcentagem de acertos. Um trabalho recente com base em técnicas de hibridização da Inteligência Artificial em [2] pode-se observar uma hibridização na amostra aplicada a um ambiente semelhante ao OCR, em outros trabalhos, eles são mostrados sobre como avaliar essas redes [18].

3 REDE NEURAL KOHONEN

O modelo Kohonen [21], também conhecido como mapas de Kohonen, apresenta um arquitetura simples, que consiste em uma malha de neurônios conectados a outros (ver Figura 1).

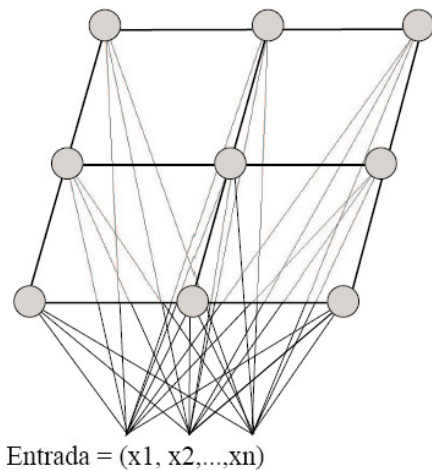


Figure 1: Arquitetura básica de um mapa auto-organizável [9].

Este tipo de rede, em geral, é um marco do paradigma não supervisionado, o algoritmo utilizado para o processo de aprendizagem é baseado na competitividade da rede é baseada no Hebbiano [16].

Esquema geral do algoritmo de aprendizagem em SOM

- (1) inicializar os pesos
 - Atribuir pesos a pequenos valores aleatórios
- (2) Envie uma entrada
 - O conjunto de aprendizagem é apresentado repetidamente até atingir a convergência da rede
 - Atualizar α (reduzir seu valor).
- (3) Propagar o padrão de entrada para a camada de competição
 - Obter os valores de saída (distâncias) dos neurônios da referida camada
- (4) Selecione o neurônio vencedor G
 - A menor distância para o padrão
- (5) atualizar conexões entre a camada de entrada e o neurônio C
 - Também atualizar os pesos de seus vizinhos de acordo com o grau de vizinhança
- (6) Se α se mantiver acima do limite de parada, volte para 2, caso contrário.

3.1 Aprendizagem competitiva em mapas kohonen

A aprendizagem competitiva é usada nas redes Kohonen [22] e ART [5]; Deve-se notar que durante esta “competição”, há apenas **um neurônio** que pode estar ativo em um determinado momento, [15], que é a última característica que torna o aprendizado competitivo mais apropriado para descobrir características estatisticamente salientes que pode ser usado para classificar um conjunto de padrões de entrada. [15], da mesma forma a desvantagem de usar este tipo de aprendizado é o tempo decorrido para o treinamento da dita rede

Também é importante notar que este tipo de redes não possui uma função de ativação, como as redes mencionadas acima (Perceptron, MLP, entre outras), porque a **competição** é baseado na aprendizagem competitiva [15], na Equação (1).

$$y_k = \begin{cases} 1, & \text{se } v_k > v_j \text{ para todo } j, j \neq k \\ 0, & \text{caso contrário} \end{cases} \quad (1)$$

O que implica que, para um neurônio k ser o neurônio vencedor, seu campo local induzido por v_k , para um padrão de entrada X especificado, deve ser o mais alto dentre todos os neurônios da rede, para o qual o vencedor recebe um **1** e o restante **0**. Por outro lado, o campo local induzido v_k , é representado pela combinação de todas as entradas diretas e de *feedback* do neurônio k .

3.2 A taxa de aprendizado η

De acordo com a regra de aprendizagem competitiva, uma variação em Δw_{kj} , aplicado ao peso sináptico w_{kj} , é definido pela regra de aprendizagem competitiva [15], na Equação (2).

$$\Delta w_{kj} = \begin{cases} \eta(x_j - w_{kj}), & \text{se o neurônio } k \text{ ganha a competição} \\ 0, & \text{se o neurônio } k \text{ perde a competição} \end{cases} \quad (2)$$

Esta regra tem um efeito global de **mover** o vetor de pesos sinápticos w_k do neurônio vencedor na direção do padrão de entrada X [15].

3.3 Mapas auto-organizáveis

Um mapa auto-organizado é caracterizado pela formação de um *mapa topográfico* dos padrões de entrada nos quais as localizações espaciais (coordenadas) dos neurônios são indicadores das características contidas nos padrões de entrada. Consequentemente, os neurônios transformam os sinais de entrada em uma distribuição de probabilidade codificada por sua localização, que representa os valores calculados dos parâmetros por posições de atividade relativa máxima dentro do mapa, isto é, a vizinhança do neurônio vencedor [15].

O objetivo dos mapas de auto-organização é transformar um padrão de sinal incidente de dimensão arbitrária em um mapa uni ou bidimensional discreto e realizar de forma adaptativa essa transformação de forma topologicamente ordenada. [15]

Em geral, o que uma rede de Kohonen faz é executar uma tarefa de classificação, já que o neurônio de saída ativado antes de uma entrada representa a classe à qual a informação de entrada pertence e, além de outra entrada similar, o mesmo neurônio é ativado, saída, ou outro próximo do anterior devido à similaridade entre as classes,

assegurando assim que os neurônios topologicamente vizinhos sejam sensíveis a entradas fisicamente similares. Por esse motivo, a rede é especialmente útil para estabelecer relacionamentos, até então desconhecidos, entre conjuntos de dados

4 PROPOSTA

No presente trabalho, um modelo híbrido inteligente é proposto para reconhecimento óptico de caracteres. É assim que o uso de algoritmos genéticos tem sido visto como conveniente para a **evolução do vetor de pesos**.

Como visto na literatura, muitos dos modelos propostos tem uma margem de erro alta quando na entrada são caracteres manuscritos ou quando a qualidade do documento e sua digitalização estão deteriorados, para razão o reconhecimento de caracteres, continua em uma constante busca para melhorar a sua eficiência, um ponto que se destina a ser abordada no presente trabalho. é por isso necessário o método para a aquisição e processamento de imagens, documentos digitalizados e o próprio modelo híbrido, para o qual o uso de uma rede neural do tipo Kohonen, de acordo com muitos dos estudos avaliados durante o levantamento da literatura, que mostram alguns das dificuldades que o algoritmo SOM tem, como o cálculo dos pesos, e a escolha do neurônio vencedor, por Este motivo viu diferentes abordagens para alcançar um melhor topologia que possivelmente melhore os resultados, onde se procura melhorar o número de acertos de letras reconhecidas corretamente.

4.1 Limitações do método proposto

Abaixo estão algumas das limitações básicas para o que o trabalho está sujeito:

- (1) As imagens contendo **texto** serão usadas, então aquelas que contiverem desenhos ou similares não serão aceitas. Se não forem necessariamente páginas brancas, elas podem conter texturas ou cores, então o tipo de folha: bond, volumoso, cartões são admitidos, da mesma forma o formato que irá trabalhar com os seguintes formatos: jpg, jpeg, bmp, png.
- (2) O conjunto de treinamento, ou seja, as amostras que serão utilizadas, serão do alfabeto latino (constituído por 26 letras), somente com base em suas letras maiúsculas.
- (3) Outra limitação é o tipo de escrita, que pode ser tanto **type printing** (script), quanto sua extensão para manuscritos de maneira similar.

4.2 Etapas do modelo proposto

As fases envolvidas em um sistema de OCR baseado em redes neurais é o seguinte [19]:

- (1) Obtenção da imagem digitalizada do documento impresso na forma de um bitmap do scanner ou de outro dispositivo.
- (2) Executa o pré-processamento da imagem.
- (3) Isolamento dos caracteres.
- (4) Extração de características capazes de definir cada caractere isolado (vetor de características).
- (5) Aprendizagem e reconhecimento de caracteres com a rede neural.
- (6) Durante este processo, os parâmetros internos da rede são ajustados repetidamente até que seja possível gerar o código ASCII como saída.

Na Figura 2 é mostrado com mais detalhes as fases envolvidas no desenvolvimento do modelo proposto, a imagem é obtida primeiro digitalizado do documento impresso na forma de um bitmap, simplesmente uma imagem que contém texto.

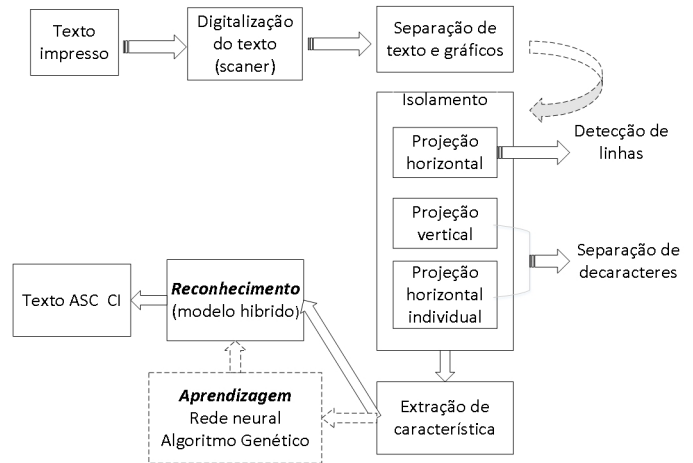


Figure 2: Estágios para o modelo híbrido. Imagem adaptada de [19]

Esta imagem tem uma estrutura matricial bidimensional, cuja a magnitude é dada em função do tamanho e resolução que você tem. O primeiro passo é reconhecer dentro da imagem, os componentes correspondentes ao caracteres, uma vez que geralmente a imagem de entrada tem algum ruído, ou seja porque está presente no sinal original ou porque ocorre durante a etapa de aquisição do mesmo. Esse módulo de processamento de imagem digital é usado para filtrar a imagem, neste caso foi usado o filtro passa-baixo em seguida a imagem é binarizada.

Após esse processamento anterior da imagem, ocorre o isolamento de cada um dos caractere que compõem o texto que está sendo tratado e, posteriormente, para a extração de um conjunto de características capazes de definir cada caractere isolado.

Uma vez obtido o vetor de características de cada um dos caracteres incluídos no documento, ativa o estágio de reconhecimento que consiste em uma rede neural capaz de gerar como saída o código ASCII de 8 bits do indivíduo cujo vetor de características recebidos como entrada. Para isso ser possível, anteriormente foi submetido à rede neural de Kohonen levantamento da matriz de peso ideal, que é obtida usando um algoritmo genético, note que o formação, uma vez que esta já foi obtida para este problema específico do OCR, deve também ser tido em conta que para a "formação" da rede de Kohonen durante as fases de seleção e mutação do algoritmo genético, um conjunto de treinamento com vários conjuntos de caracteres ou alfabeto (todas as letras maiúsculas A-Z, 26 letras) com tamanhos e fontes diferentes (Arial, Times New Roman e Comic Sans).

Quando os dados de entrada e vetores de características tenham sido obtidas e processados os parâmetros internos da rede repetidamente até que seja capaz de gerar como saída o código ASCII do caractere cuja a imagem é apresentada na sua entrada, para todas

as formas e tamanhos considerado para tal caractere, se o classificador puder determinar que letra é, a tarefa é terminada para essa possível entrada e continuar com o próximo vetor de entrada. Nos seguintes pontos serão detalhadamente sobre os passos necessários envolvidos em cada etapa.

4.3 Pré-processamento

Para a representação do vetor de características, ele deve ser representado numericamente na forma vetorial, e isso nos força a fazer a equivalência entre uma letra do alfabeto e um vetor n -dimensional. Para isso, foi usado o processo de pixelização no qual cada letra é representada em uma matriz de F linhas por C colunas, que depois de criadas, será processada não como uma matriz, mas como o vetor resultante da junção de cada um das F linhas da matriz representativa em uma única, e assim tem-se um vetor composto de $F \times C$ de elementos, ou seja um vetor n -dimensional. Para exemplificar, a matriz de dados apresentado na Figura 3 foi transformada num vetor na Figura 4.

0	0	0	1	1	1	0	0	0
0	0	0	1	1	1	1	0	0
0	0	1	1	1	1	1	0	0
0	0	1	1	1	1	1	0	0
0	0	1	1	1	1	1	0	0
0	1	1	1	0	1	1	1	0
0	1	1	1	0	1	1	1	0
0	1	1	1	1	1	1	1	0
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
1	1	1	0	0	0	1	1	1
1	1	1	0	0	0	1	1	1

Figure 3: Matriz de 9×12 para uma letra amostrada da fonte Arial.

Para determinar quais valores devem ter F e C , depende-se da quantidade ou qualidade do pixel que é aplicado a essa letra, porque se n for muito grande, a letra estará em um nível de pixelização no qual ela pode ser definida como boa como as curvas e os picos que a letra possui, mas se n for pequena, a letra terá um estilo não tão real, mas um pouco mais digital.

Neste trabalho, em particular, foi definido uma matriz de 9 linhas por 12 colunas com $n = 108$ atributos que consegue representar uma letra escrita como pode ser visto na Figura 3. A matriz representativa será binária e o valor 0 indicará um espaço branco ou vazio, e o valor de 1 representará essa parte do traço da letra e depois de transformá-lo em um vetor binário, finalmente com isso é obtido um padrão de treinamento que servirá de entrada para a rede Kohonen.

0	0	0	1	1	1	0	0	0	.	.	.	1	1	1	0	0	0	1	1	1
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Figure 4: Vetor de características de 108 entradas.

4.4 Processamento

Um módulo foi implementado que executa separadamente o processamento de imagens, ou seja, obter o vetor de características mencionado anteriormente. Deve-se notar que isto tem sido usado em biblioteca o que muita ajuda para o tratamento de Imagens JAI, referentes às técnicas utilizadas para este caso eram os seguintes:

Filtro passa-baixo, com uma máscara do tipo KMedia3x3.

DownSampling, encarregado de redimensionar a imagem para 9×12 ao cortar a imagem.

Projeção Vertical e Horizontal para a segmentação das letras, primeiro é feita uma projeção horizontal para detectar a linhas e, em seguida feita uma projeção vertical para detectar a presença de Preto na imagem que implica o começo de uma letra.

PixelGrabber extração de características, permite binarização e normalização para 1 e 0.

Estes são alguns dos resultados que foram obtidos a partir do processamento da imagem:

REFRESCA TEU
CORACAO. SOFRE,
SOFRE, DEPRESSA,
QUE E PARA AS
ALEGRIAS NOVAS
PODEREM VIR
GUIMARÃES ROSA

Figure 5: Imagem digitalizada do texto de Guimarães Rosa

Realizando o processamento, para a segmentação das imagens:

REFRESCA TEU
CORACAO. SOFRE,
SOFRE, DEPRESSA,
QUE E PARA AS
ALEGRIAS NOVAS
PODEREM VIR
GUIMARÃES ROSA

Figure 6: Imagem processada

4.5 Configuração da rede neural Kohonen

A topologia da rede proposta é a seguinte:

- Número de entradas: 108
- Número de saídas: 26, este valor pode ser aumentado como um conjunto de nível superior
- Taxa de aprendizado 0,5, um valor intermediário entre 0,1 e 1, que é a convergência mais rápida
- Coeficiente de redução: 0,99, valor pelo qual a rede Kohonen reduzirá o valor da taxas de aprendizagem.

4.6 Conjunto de treinamentos

Fontes usadas no conjunto de treinamento:

- Arial
- Times New Roman
- Comic Sans

Como pode ser visto na figura 7

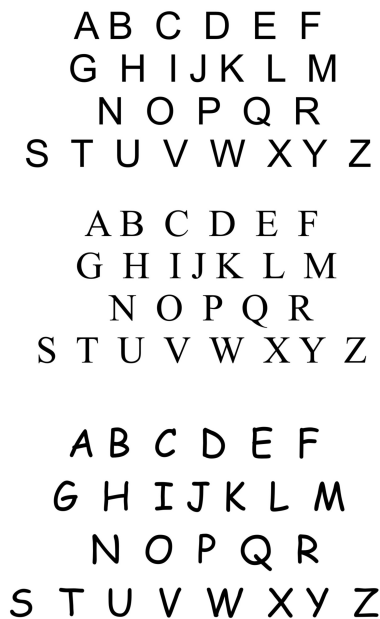


Figure 7: Conjunto de treino: letras do tipo Arial, Times e Comic

5 PROPOSTA DE OTIMIZAÇÃO USANDO ALGORITMO GENÉTICO

5.1 Algoritmos genéticos

Algoritmos genéticos (AGs) são métodos adaptativos que podem ser usados para resolver problemas de busca e otimização, tentando encontrar a melhor solução para um dado problema entre um conjunto de possíveis soluções. Os mecanismos que os GAs usam para realizar essa busca podem ser vistos como uma metáfora para os processos de evolução biológica.

Genética: É a ciência que estuda os mecanismos de transmissão de características de uma espécie de uma geração para outra.

População: Conjunto de indivíduos que corresponde a um conjunto de soluções iniciais.

Indivíduos: Também é chamado de cromossomo e corresponde a uma possível solução, que está associada a um valor chamado de aptidão.

Função de aptidão: Grau de aptidão de um indivíduo e medir a capacidade da solução (individual) para a solução de um determinado problema.

As características desta proposta são:

- O AG trabalha em um conjunto de parâmetros codificados, ou seja, realiza as operações nos cromossomos e não no problema.
- Eles realizam a busca sob um conjunto de pontos, paralelismo implícito
- Usa as informações da função de aptidão, não requer outro conhecimento.
- As regras de transição são probabilísticas não determinísticas.

Os passos que seguem o algoritmo genético são:

- (1) Inicialize a população
- (2) Seleção
 - (a) Aplicar a função objetivo
 - (b) Ordenar de acordo com o valor retido
 - (c) Selecione indivíduos para o cruzamento
- (3) Cruze
- (4) Mutação
- (5) Inserção
- (6) Se cumpre o critério de parada então pare, caso contrário volta ao passo (2)

5.2 Operadores Genéticos

- **Seleção ou reprodução:** Este operador escolhe os cromossomos entre a população para realizar a reprodução. Quanto mais capaz for o cromossomo, mais vezes ele será selecionado para se reproduzir.
- **Cruzamento:** Este é um operador cujo trabalho é escolher um lugar, e alterar as sequências antes e depois daquela posição entre dois cromossomos, para criar novos descendentes (por exemplo, as cadeias 10010011 e 11110101 podem ser cruzadas após o terceiro lugar para produzir descendentes 10011010 e 11110011). Ele imita a recombinação biológica entre dois organismos haploides.
- **Mutação:** Este operador produz variações de maneira aleatória em um cromossomo (por exemplo, a string 00011100 pode sofrer mutação em sua segunda posição para dar origem à string 01011100). A mutação pode ocorrer em cada posição de um bit em uma cadeia, com uma probabilidade, geralmente muito pequena (por exemplo, 0,001).

Este algoritmo genético tem as seguintes características:

- **Individual:** matriz de peso da rede de Kohonen, uma vez que se destina a otimizar os pesos da rede.
- **Tamanho da população:** 7, este valor deve ser aumentado para mais de 100, devido a limitações de tempo, era conveniente usar essa quantidade
- **Porcentagem de mutação:** 0,10
- **Tamanho da Combinação:** População / 2
- **Error inicial global :** 1
- **fitness:** .3, foi considerado conveniente usar este fator, porque a rede Kohonen simples, para o seu treinamento atinge um valor de 0.96, que otimiza o Algoritmo Genético

6 TESTES E RESULTADOS

Nesta seção serão apresentados dois experimentos, na primeira parte será feita uma análise com os dados sintéticos e na segunda parte com os dados escritos à mão. Para isso o primeiro conjunto de dataset sintético foi construído utilizando a ferramenta do word com 11 diferentes fontes de letras. Observe que cada uma dessas letras contém dificuldades para ser reconhecida pelos algoritmos, essas letras são bem variadas com respeito ao conjunto de treinamento.

6.1 Dataset teste sintético

Esta base de dados pode ser encontrada na Tabela 1. As fontes usadas para formar este dataset foram: AGENCY FB, BELL MT, BOOK ANTIQUA, COURIER, SEGOE UI, TERMINAL, CONSOLAS, FORTE, COLONNA MT, Chiller, CALIBRI. Nesta mesma Tabela são apresentados as taxas de acerto para o reconhecimento de caracteres sintéticos e é feita uma comparação ao usar algoritmos genéticos (esquerda) e a ausência deste (direita).

Table 1: Taxa de acerto (%) usando o dataset de letras sintéticas. As últimas colunas amostram os resultados usando algoritmos genéticos e sem ele.

										Com G.	Sem G.
A	À	Á	A	Â	Ã	Ä	Å	Ả	Ǻ	81%	27%
B	Ĳ	Ĳ	B	Ĳ	Ĳ	Ĳ	Ĳ	Ĳ	Ĳ	36%	27%
C	Ĳ	C	C	C	C	C	C	C	Ĳ	81%	72%
D	Ĳ	D	D	D	D	D	D	D	Ĳ	72%	18%
E	Ĳ	E	E	E	E	E	E	E	E	36%	0%
F	F	F	F	F	F	F	F	F	F	72%	27%
G	Ĳ	G	G	G	G	G	G	G	Ĳ	90%	63%
H	Ĳ	H	H	H	H	H	H	H	H	72%	36%
I	I	I	I	I	I	I	I	I	I	72%	36%
J	J	J	J	J	J	J	J	J	J	81%	63%
K	K	K	K	K	K	K	K	K	K	72%	36%
L	L	L	L	L	L	L	L	L	L	81%	36%
M	M	M	M	M	M	M	M	M	M	63%	18%
N	N	N	N	N	N	N	N	N	N	45%	27%
O	O	O	O	O	O	O	O	O	Ĳ	90%	63%
P	P	P	P	P	P	P	P	P	P	90%	72%
Q	Q	Q	Q	Q	Q	Q	Q	Q	Ĳ	45%	18%
R	R	R	R	R	R	R	R	R	R	72%	27%
S	S	S	S	S	S	S	S	S	S	81%	63%
T	T	T	T	T	T	T	T	T	T	72%	54%
U	U	U	U	U	U	U	U	U	U	72%	36%
V	V	V	V	V	V	V	V	V	V	90%	36%
W	W	W	W	W	W	W	W	W	W	72%	45%
X	X	X	X	X	X	X	X	X	X	90%	72%
Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	81%	72%
Z	Z	Z	Z	Z	Z	Z	Z	Z	Z	90%	90%

6.2 Dataset teste a mão

Da mesma forma que foi realizado o experimento dataset sintético , agora sera aplicado o reconhecimento de carácter escrito a mão .O resultado é apresentado na figura 2 onde estão as respectiva taxa de acerto para cada letra .Da mesma forma que contabiliza um total de 26 letras com 11 exemplos para cada dataset

Table 2: Taxa de acerto (%) usando o dataset de letras escritas a mão. As últimas colunas mostram os resultados usando algoritmos genéticos e sem ele.

[illegible]

6.3 Comparações de performance

A Figura 8 mostra o desempenho do algoritmo da rede neural Kohonen comparado com o algoritmo híbrido de rede neural Kohonen usando o algoritmo genético.

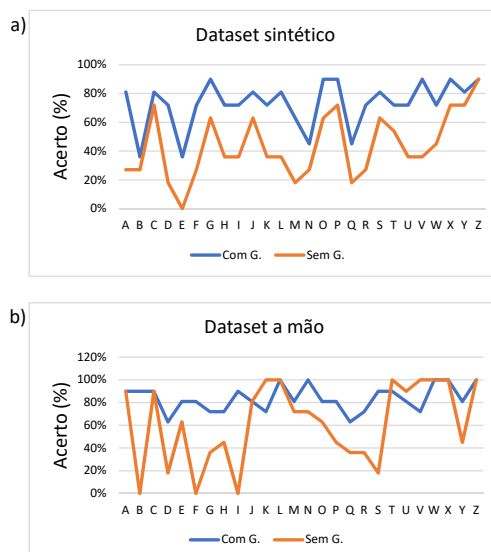


Figure 8: Taxa de acerto usando rede Kohonen versus a rede Kohonen treinada com algoritmo genético para o (a) dataset sintético e (b) dataset a mão.

7 DISCUSSÃO E CONCLUSÕES

- O modelo híbrido proposto permitiu o reconhecimento óptico de caracteres em documentos digitalizados de forma eficiente.
- Foi observado que o algoritmo proposto, da combinação de rede Kohonen e algoritmos genéticos, conseguiu uma performance muito maior quando comparada ao resultados apenas da rede Kohonen.
- Ainda existem algumas desvantagens no que diz respeito ao tratamento de imagens, que deve ser levado em conta, uma vez que é uma parte importante de um sistema de OCR deve ter mais atenção para as fases de segmentação e cortar em letras.
- Dentro da área de computação evolucionária, existe uma outra abordagem diferente dos algoritmos genéticos, cuja vantagem é o tempo de convergência.
- Nesse papel, de acordo com os passos do algoritmo genético, que toma como pesos iniciais aleatoriamente, seria usado como uma rede doméstica simples treinado desta forma para melhorar o tempo de convergência e precisão.

Sobre trabalhos futuros:

- Sobre o processamento de imagem, o uso de outros tipos de filtros, ou mesmo filtro passa-alta para reforçar as bordas das letras, poderia otimizar o processamento de imagem também foram avaliadas possíveis máscaras novo filtro, entre eles, o **Kirsh mask**;
- Outro problema a ser enfrentado é a rotação das imagens, um assunto que deve ser visto mais profundamente em um trabalho posterior;
- Aumentar o conjunto de treinamento, este trabalho deixa em aberto a possibilidade de estender o número de amostras por carta e aumentou para minúsculas, números e sinais específicos, considerando que algumas dificuldades têm, no caso da classificação de um O (maiúsculo) de um, ou (minúscula), um 0 (zero);
- Da mesma forma, a extensão do trabalho do manuscrito é possível considerando que um conjunto de treinamento com caracteres manuscritos ou em itálico de algum formato deve ser adicionado;
- Deve-se notar que houve um conjunto de amostras pobres, para trabalhos futuros, pretende-se usar um banco de dados fornecido, como é o caso do NIST.

REFERENCES

- [1] Eric B. Baum and David Haussler. 1989. What size net gives valid generalization? *Neural Comput.* 1, 1 (1989), 151–160. <https://doi.org/10.1162/neco.1989.1.1.151>
- [2] Luis Federico Bertona. 2005. *Entrenamiento de redes neuronales basado en algoritmos evolutivos*. Master's thesis.
- [3] J.L. Blue, G.T. Candela, P.J. Grother, R. Chellappa, and C.L. Wilson. 1994. Evaluation of Pattern Classifiers for Fingerprint and OCR Applications. *PR* 27, 4 (April 1994), 485–501.
- [4] Eric W. Brown. 1992. Applying Neural Networks to Character Recognition. (1992).
- [5] S. Grossberg Carpenter and J. H. Reynolds. 1991. A self-organizing ARTMAP neural architecture for supervised learning and pattern recognition. (1991), 43–80.
- [6] Z.R. Chi and H. Yan. 1995. Handwritten Numeral Recognition Using a Small Number of Fuzzy Rules with Optimized Defuzzification Parameters. *NeurNet* 8, 5 (1995), 821–827.
- [7] Sung-Bae Cho. 1997. Self-Organizing Map with Dynamical Node Splitting: Application to Handwritten Digit Recognition. *Neural Comput.* 9 (1997), 1345–1355.
- [8] Guillermo de-los Santos-Torres. 2004. *Análisis de algoritmos para reconocimiento de caracteres manuscritos de un solo escritor: caso de Porfirio Díaz*. Master's thesis. Maestría en Ciencias con Especialidad en Ingeniería en Sistemas Computacionales. http://catarina.udlap.mx/u_dl_a/tales/documentos/msp/de_l_g/portada.html
- [9] Ernesto-Cuadros. 2004. *Recuperación de información por similitud utilizando técnicas inteligentes*. Ph.D. Dissertation. Instituto de Ciências Matemáticas e de Computação | ICMC/USP USP - Sao Carlos.
- [10] Xulio Fernández H. Fernando M. Rodríguez. 1998. RAMA: Reconocedor Automático de Matriculas de Automóviles.
- [11] Bernd Fritzke. 1994. Growing cell structures - a self-organizing network for unsupervised and supervised learning. *Neural Networks* 7 (1994), 1441–1460.
- [12] K. Fukushima and N. Wake. 1991. Handwritten Alphanumeric Character Recognition by the Neocognitron. *TNN* 2, 3 (1991), 355–365.
- [13] Juan Gomez. 2005. Optimización de un reconocedor de texto manuscrito por medio de mapas auto-organizados de Kohonen(RNA's). *Scientia et Technica Año XI* (2005).
- [14] V.K. Govindan and A.P. Shivaprasad. 1990. Character Recognition: A Review. *PR* 23, 7 (1990), 671–683.
- [15] Simon Haykin. 1999. *Neural Networks - A Comprehensive Foundation* (second edition ed.).
- [16] D.O. Hebb. 1949. *The Organization of Behavior*. New York: John Wiley. (1949).
- [17] Venu-Govindaraju Jaehwa-Park. 2000. Active Character Recognition Using A*-like Algorithm. (2000). http://www.cedar.buffalo.edu/papers/articles/Active_Character_2000.pdf
- [18] Arnaldo-Lara e Fabricio-Martins-Lopes Jesús Mena Chalco, Alexandre-Noma. 2006. Análise e Reconhecimento de Formas. (2006), 12.
- [19] José A. Gutiérrez de Mesa José R. Hilera González, Juan P. Romero Villaverde. 2002. Sistema de Reconocimiento Óptico de Caracteres (OCR) con Redes Neuronales. (2002).
- [20] Enrique Vidal y Lourdes Sánchez. Juan C. Pérez. 1993. Un sistema geométrico de reconocimiento óptico de caracteres. (1993).
- [21] Teuvo Kohonen. 1984. *Self-organization and associative memory*. Springer Series of Information Science.
- [22] Teuvo Kohonen. 1997. *Self-organizing maps*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.
- [23] E-C-Segura L.-Seijas. 2005. Modelos Conexionistas Auto-organizados y su Aplicación en Reconocimiento de Patrones. (2005).
- [24] Seong-Whan Lee. 1995. Multilayer cluster neural network for totally unconstrained handwritten numeral recognition. *Neural Networks* 8, 5 (1995), 783–792. [https://doi.org/10.1016/0893-6080\(95\)00020-Z](https://doi.org/10.1016/0893-6080(95)00020-Z)
- [25] A.-K.-Jain. S.-Raudys. 1991. Small sample size effects in statistical pattern recognition: recommendations for practitioner. *IEEE Trans Pattern Analysis and Machine Intelligence* 31 (1991), 252–264.
- [26] Tomás Saorín-Pérez. 2002. *Modelo Conceptual para la automatización de bibliotecas en el contexto digital*. Master's thesis. Universidad de Murcia, Departamento de Información y Documentación.
- [27] Simon Tanner. 2004. Deciding whether Optical Character Recognition is feasible. (2004).