

União de métodos de processamentos de imagens digitais, que encadeado em processo de reconhecimento dígitos escritos a mão livre

Iury Adones Xavier dos Santos^{1*}, Dr. Adenilton José da Silva¹, Dr. Péricles Barbosa Cunha de Miranda¹, Dr. Danilo Ricardo Barbosa de Araújo¹,

Resumo

Neste presente trabalho, têm experimentos voltados na área de processamento de imagens digitais e heurística voltada ao reconhecimento de números escritos, no entanto, aborda-se alguns passos antes mesmo de chegar-se ao reconhecimento dos dígitos. Nesses passos que nos revela um fluxo encadeado, são atualmente produzidos nas literaturas acadêmicas as seções do fluxo de processamento de imagens, que inicia desde aquisição da imagem bruta, limpeza de ruídos e binarização dos objetos, pode compor até mesmo com reconhecimento de objetos. Tais trechos da trajetória do encadeamento, necessita de métodos de otimização, pois melhoram os resultados, tanto na visualização e classificação. Nós construímos experimentos, que testa algumas das possíveis combinações dos fluxos de encadeamento, estão contidas em um espaço de busca restrito, e nesta construção abordamos com métodos baseado em algoritmos genéticos, que nos informa quais as melhores combinações do fluxo encadeado, passadas por avaliação de minimização do quantitativo de erros na detecção dos dígitos, sem a necessidade de otimizar cada seção por vez.

Palavras-chaves

Algoritmos genéticos — Processamento encadeado — Reconhecimento de dígitos

¹ Programa de Pós-Graduação em Informática Aplicada (PPGIA), Universidade Federal Rural de Pernambuco (UFRPE), Recife-PE, Brasil

*Autor correspondente: iuryadones@gmail.com

Sumário		2	Resultados e Análise	3
	Introdução	2		
1	Metodologia	2		
1.1	Aquisição das imagens de treino	2	2.1 Experimento - 1	3
1.2	Elaboração das imagens testes	2	2.2 Experimento - 2	4
1.3	Escolha do classificador	2	2.3 Experimento - 3	4
1.4	Descritor de objeto	2		
1.5	Algoritmos de processamento de imagens	2		
1.6	Algoritmo genético	3		
1.7	Restrição do algoritmo genético	3		
			Agradecimentos	4
			Referência	4

Introdução

Nos recentes trabalhos têm heurísticas que focam em melhorar as imagens para depois ter uma segmentação, classificação e reconhecimento de objetos, mas devido ao progresso nas pesquisas, agora podemos usar uma imagem de entrada e dizer que espera se na saída como resposta, tal como uma rede neural supervisionada, no entanto existe uma longa jornada com experimentos e avaliações, ou seja, até chegarmos nos excelentes resultados. Temos em [3], a contribuição na seleção de métodos de pre-processamento para melhoras nas classificações, e com isso temos entendimento sobre a identificação de caracteres, dígitos, letras e números em uma imagem, porém se tivermos uma grande quantidade de imagens identificarmos, tais elementos, isso será um gasto de tempo e energia mental exorbitante. Foram desenvolvidos algoritmos que ajudam na redução desses gastos, porém necessita de recursos como computadores, para tal gasto, portanto, percebe se que sempre existira um gasto para tal feito, mas também podemos otimizar tais gastos, no reconhecimento de dígitos, podemos usar métodos e ou heurísticas dos algoritmos genéticos e ou evolucionários, logo existem grandes desafios em otimizarmos o nosso gasto computacional e encontrar a melhor solução para todas as imagens ou grupos de imagens.

1. Metodologia

Experimentos baseados no processamento de imagem digital, que iremos abordar desde aquisição da base, escolhas dos algoritmos de processamento de imagens, escolha dos algoritmos de binarização até a escolha do método de classificação.

1.1 Aquisição das imagens de treino

Usamos a base da MNIST [4], usamos no classificador para treinar, e também as imagens da MNIST, já são bem conhecida, para pequenos testes e já estão em forma de binarização, têm umas imagens de escrita a mão livre, com dígitos e de diferentes tipos de escritas.

1.2 Elaboração das imagens testes

Imagens testes foram elaboradas as escritas dos dígitos de forma aleatória, mas que fossem próximas as que estão na base da MNIST, visto que nosso objetivo é em preprocessar as imagens e usar o classificador para os dígitos elaborados e ter uma classificação satisfatória, mas não é como objetivo principal tal classificação. Nesta fase de teste de hipótese, de tal, será possível construir um único fluxo encadeado para o processamento de imagens de diferentes tamanhos e fontes diferentes, também será possível classificar todos os dígitos contidos nas imagens só com um único encadeamento? Está é a pergunta que nos move em tal pesquisa.

1.3 Escolha do classificador

A escolha do classificador, foi usado dois critério que são a velocidade de treinamento e classificação, e também sua frequência nos periódicos de classificação de textos, dígitos e caracteres. Escolhemos o máquina de suporte vetorial, também conhecida como SVM que em inglês é “suporte vector machine” [1].

1.4 Descritor de objeto

O descritor usado neste trabalho, foi o histograma de gradientes orientados, conhecido como HOG em inglês é “Histogram of Oriented Gradients” [5]. Buscamos usar este descritor, pelo que foi visto sendo usado em algumas literaturas, mas também nosso objetivo é só classificar objetos nas imagens testes e de forma mais flexível.

1.5 Algoritmos de processamento de imagens

Foram usados um conjuntos de algoritmos, mas que já foram passados por um crivo de uns periódicos. Tais algoritmos foram de limiarização gaussiana, filtro de média, filtro gaussiano, algoritmo de Otsu para binarização, também foi usado binarização invertida, e método adaptativo de binarização por média e com distribuição gaussiana de acordo com [2], algoritmo de dilatação e de erosão, algoritmos de transformação de imagens de três camadas para uma camada, também conhecida como transformação de tons de cinza.

1.6 Algoritmo genético

{Reference} Utilizamos nesta pesquisa o algoritmo genético, seguindo um fluxo, primeiro geração da população de forma aleatória com distribuição uniforme, para os alelos, temos que alelos são todos os algoritmo de processamento de imagens e também consideramos seus variantes, no entanto deixamos o primeiro índice do alelo como sendo nulo, ou seja, não existe um operador de processamento de imagem, como fixamos o tamanho da solução em 8 genes e 12 genes, isso faz com que não necessariamente usaremos os operadores em todos os genes, objetivo é encontrar a solução ótima, e tal solução é nosso encadeamento.

1.7 Restrição do algoritmo genético

{Reference} Usamos restrição no gene, que é um dos possíveis operadores de processamento de imagem, porém existe certos operadores que são usados em tensores e outros que só se aplicam em matrizes, mas na formação da solução, o algoritmo que gera não sabe qual é a ordem dos operadores. Para que seja executados todos corretamente, no entanto, abordamos dois tipos de restrições, uma que se tiver no gene um operador que não consiga ser executado, então passe para o prosseguimos e a informação daquela solução será passada para a próximas geração, já a outra como somos especialista, colocamos as imagens já no formato de matrizes para não tenha uma passagem forçada, pois sabemos que todos outros operadores, conseguem processar em matrizes, mas só alguns que conseguem em tensores.

2.1 Experimento - 1

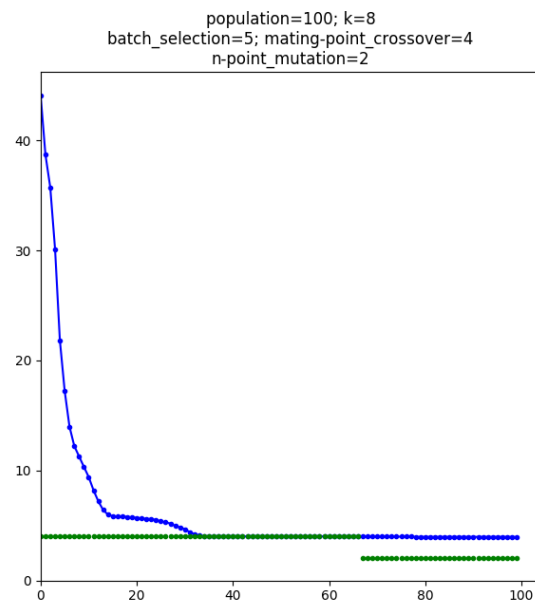


Figura 1. Solução com 8 passos de encadeados de processamento e tendo cada passo 8 operadores possíveis de escolha, sem restrição da solução.

2. Resultados e Análise

Conseguimos alguns indicadores que poderemos contribuir com a comunidade acadêmica, tais indicadores são no pré processamento das imagens antes mesmo de classificação dos objetos contidas, ou seja, uma preparação da imagem ou um conjunto de imagens para único tipo de encadeamento.

2.2 Experimento - 2

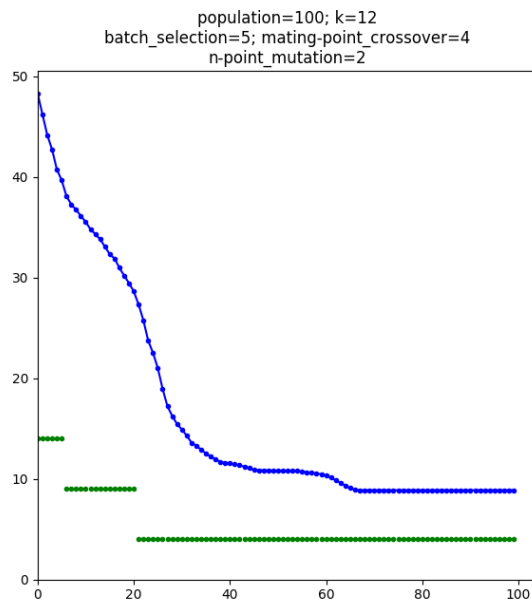


Figura 2. Solução com 12 passos de encadeados de processamento e tendo cada passo 20 operadores possíveis de escolha, com restrição de passar para o próximo operador do encadeamento.

2.3 Experimento - 3

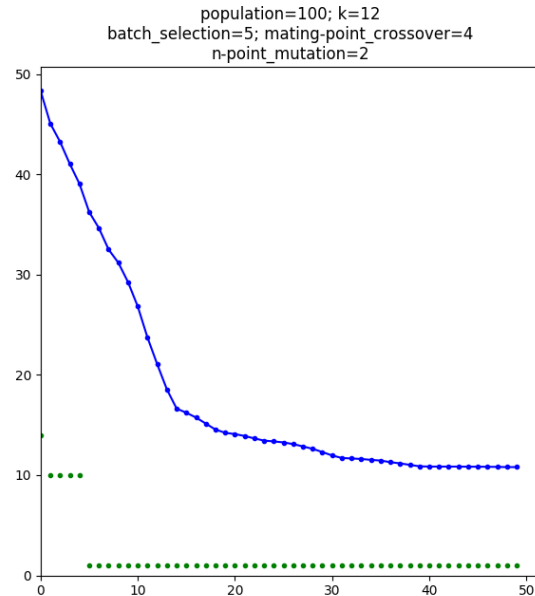


Figura 3. Solução com 12 passos de encadeados de processamento e tendo cada passo 20 operadores possíveis de escolha, com restrição do especialista que transforma inicialmente a imagem tensorial para matriz na escala de cinza.

Observamos que podemos usar uma heurística para pre-processamento de imagens de forma a servir para diversas imagens da mesma classe de problema, de acordo com os resultados apresentados na figura 3, que chega apenas errar um dígito na imagem, sabendo que o conjunto é de 40 dígitos.

Agradecimentos

Agradeço a todos que contribuíram com o projeto.

Referências

- [1] R. Anil et al. "Convolutional Neural Networks for the Recognition of Malayalam Characters". Em: *Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014*. Ed. por Suresh Chandra Satapathy et al. Cham: Springer International Publishing, 2015, pp. 493–500. ISBN: 978-3-319-12012-6.

- [2] Maya R. Gupta, Nathaniel P. Jacobson e Eric K. Garcia. “OCR binarization and image pre-processing for searching historical documents”. Em: *Pattern Recognition* 40.2 (2007), pp. 389–397. ISSN: 0031-3203. DOI: <https://doi.org/10.1016/j.patcog.2006.04.043>. URL: <http://www.sciencedirect.com/science/article/pii/S0031320306002202>.
- [3] D. V. Katkar, A. K. Almas e J. Kajal. “Experiments on Selection of Pre-processing Method for Performance Augmentation of Classifier”. Em: *2015 International Conference on Computing Communication Control and Automation*. Fev. de 2015, pp. 451–456. DOI: 10.1109/ICCUBEA.2015.94.
- [4] Yann LeCun e Corinna Cortes. “MNIST handwritten digit database”. Em: (2010). URL: <http://yann.lecun.com/exdb/mnist/>.
- [5] A. J. Newell e L. D. Griffin. “Multiscale Histogram of Oriented Gradient Descriptors for Robust Character Recognition”. Em: *2011 International Conference on Document Analysis and Recognition*. Set. de 2011, pp. 1085–1089. DOI: 10.1109/ICDAR.2011.219.