



Modelos matemáticos para processamento de imagens, uma revisão

Igor Rossi Fermo¹ – igor_fermo@hotmail.com

Cid Marcos Gonçalves Andrade¹ – cidmga@yahoo.com.br

1 – Universidade Estadual de Maringá – Maringá - Paraná

Resumo: No mundo moderno a necessidade de simplificação e obtenção de melhores resultados com o mínimo desprendimento de trabalho e recursos é cada vez mais procurado. Uma das formas mais desenvolvidas nos últimos tempos para atingir esses objetivos é o processamento digital de imagens. O processamento digital de imagens é uma técnica computacional dedicada a extrair informações significativas de um objeto de estudo por meio de imagens e processar as informações apresentando um resultado. O processamento das imagens digitais são realizados por meio de modelos matemáticos descritos como rotinas computacionais, que objetivam o pré tratamento das imagens e a extração de informações relevantes. O objetivo deste trabalho é apresentar os principais modelos matemáticos utilizados para o processamento de imagens, juntamente com suas principais características, vantagens e desvantagens. Dentre os modelos apresentados os mais populares são MLP e CNN, que tem seu comportamento baseado em organismos humanos, e apresentam resultados relevantes com uma pequena complexidade. No entanto existem algoritmos de simples concepção que possuem resultados muito similares aos demais e empregam uma metodologia matemática muito mais simplificada. O processamento digital de imagens tem sido largamente empregado em todos os setores produtivos, visando sempre a automação e simplificação dos processos, diante deste fato o desenvolvimento de novas tecnologias e métodos de análise de imagem e desenvolvimento de novas tecnologias se fazem necessários.

Palavras-chave: Processamento digital de imagens; Visão computacional; Algoritmos para processamento de imagens; Modelos matemáticos.

1. Introdução

Na era digital, a ênfase na simplificação de atividades é muito importante, juntamente com a necessidade de encontrar soluções ótimas com o desprendimento do mínimo de esforço. A necessidade de processamento de dados em alta velocidade com a eliminação de defeitos durante o processo produtivo ou de

manufatura associados a baixos custos. Todos estes aspectos são necessários para fomentar operações rápidas e lucrativas. (SÓLUS; TURÁN, 2016)

Uma alternativa para evitar altos custos é o método de classificação automática, o qual utiliza processamento digital de imagens associado com equipamentos computacionais no processo de classificação. O sistema de visão computacional tem sido largamente utilizado na indústria de alimentos e agricultura. Os propósitos são de inspeção e avaliação, uma vez que são fornecidos resultados rápidos, econômicos, consistentes e objetivos.

Análise de imagens é o campo dedicado a extrair informações significativas de imagens e processar a informação extraída. Um problema limitante do processamento digital de imagens é a falta de informação contida nas imagens, que é comumente conhecida como imagem de baixa resolução. (RIGUI; D'ALCUNTO; SALVETTI, 2016)

Normalmente as técnicas estudadas de processamento de imagem são baseadas em trabalhos acadêmicos, os quais tem como parâmetros principais cor, tamanho, forma entre outros. (ZOU, ZHAO, 2015)

As técnicas desenvolvidas e métodos conhecidos utilizados em processamento de imagens utilizam diferentes parâmetros aplicados a aprendizagem de máquina para criar sistemas capazes de classificar um determinado tipo de produto, sendo os parâmetros determinados de acordo com o tipo de produto a ser classificado. A maioria dos sistemas faz uso de componentes como luz, câmera, software computacional.

Grande parte das aplicações de processamento de imagem são aplicáveis a

indústria, porém existem muitas outras áreas nas quais o processamento de imagem pode ser utilizado, como na área médica, alimentícia, agricultura, indústria de madeira entre outras.

Os sistemas de processamento de imagens são a essência da visão computacional e envolvem operações nas imagens para a remoção de defeitos e melhora da qualidade, os defeitos podem ser distorções geométricas, ruído, foco inadequado, iluminação não uniforme entre outros problemas. (LEFEBRE E COLABORADORES, 1993)

O foco do processamento de imagens é a separação da região de interesse do fundo da imagem (distinguir o objeto), este processo é feito pelo método de treinamento com a utilização de aprendizagem de máquina e tratamento de imagens. Os métodos de treinamento e tratamento de imagem são propostos de diferentes formas de acordo com os autores e seus métodos.

Diante deste cenário, novos métodos de processamento de imagem, com a utilização de modelos matemáticos mais precisos e eficientes são de grande importância para todos os setores produtivos da indústria.

2. Materiais e métodos

O processamento digital de imagens é composto de várias etapas, desde a captura dos níveis de intensidades luminosas de um objeto ou local por um dispositivo foto sensível, quantização e transformação dos níveis de intensidade luminosa em grandezas elétricas e posterior conversão das grandezas elétricas em informação computacional.

Todo o processo envolvendo o tratamento de imagens é realizado por rotinas matemáticas computacionais que visam a extração de informação de um conjunto de dados. O pré-tratamento e processamento das imagens é realizado por softwares de processamento de imagens ou rotinas computacionais de cunho matemático que visam a extração de informações do conjunto de dados de entrada, sendo feito das mais diversas formas.

A seguir são apresentados os principais algoritmos de processamento de imagens utilizados e suas principais vantagens e desvantagens.

2.1. KNN

O algoritmo KNN (k-nearest neighbor) ou simplesmente k-vizinhos mais próximos é um algoritmo de simples concepção e que faz o uso das distâncias métricas entre dois exemplos para realizar classificações. O algoritmo tem se mostrado eficiente e competitivo e muitas vezes superior a demais algoritmos classificadores no que diz respeito a reconhecimento facial e categorização textual.

A classificação é feita de modo que os vizinhos mais próximos pertencem a uma mesma classe e amostras de diferentes classes são separadas por uma grande margem e a proximidade dos dados revela uma relação entre eles.

O algoritmo KNN também tem seu princípio utilizado por outros algoritmos como o k-means e SVM (LUO, LIU, TAO, 2014)

2.2. NN

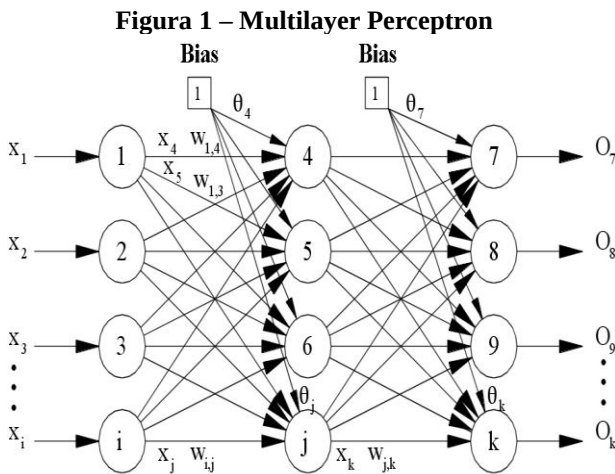
Redes neurais (*neural network*) são frequentemente utilizadas em aprendizagem de máquina e são inspiradas em comportamentos do sistema nervoso biológico. A rede neural é constituída da sobreposição de diferentes tipos de camadas. Camadas adicionais são alocadas entre as camadas de entrada e saída para que o algoritmo seja capaz de resolver problemas não lineares.

O aprendizado nestes algoritmos é realizado por exemplos.

2.3 MLP

O *Multilayer Perceptron* (MLP) é bem-sucedido na classificação dos dados. O MLP é um dos mais importantes algoritmos de tomada de decisões utilizado, com ele pode-se obter os melhores resultados com conjunto de dados incompletos e limitados. A

Figura 1 mostra a estrutura da rede MLP.



Fonte: <https://www.cse.unsw.edu.au/~cs9417/ml/MLP2/BackPropagation.html>

A grande vantagem da utilização do MLP é a capacidade do algoritmo de propagar o erro da saída para as camadas antecedentes, o que permite um melhor ajuste dos parâmetros do algoritmo.

Juntamente com o algoritmo MLP pode ser utilizado o algoritmos delta-barradelta, o que proporciona um aprendizado mais rápido ao MLP.

(AVUÇULA E, BASÇIFTÇI, 2018)

2.4. FPCNN

A FPCNN é a rede neural de par de pulsos (*The pulse couple neural network*), inspirada biologicamente, a qual teve origem nas observações de pulsos no córtex cerebral de gatos. A PCNN é uma rede adequada para localizar recursos regionais em uma imagem, caso seja necessário um estudo mais aprofundado recomenda-se sua associação com demais ferramentas de processamento de imagem afim de se obter uma segmentação mais apurada.

A arquitetura de um algoritmo FPCNN é uma matriz 2D a qual é capaz de mapear cada célula para o pixel correspondente na imagem original, basicamente cada pixel da imagem é controlado por um FPCNN. Cada célula é composta de 3 partes: Campos receptivos, produto de modulação e gerador de pulsos.

2.5. WST

Um dos métodos mais utilizados para a segmentação de imagem é o algoritmo de transformação de watershed (*watershed transform algorithm*), o algoritmo foi desenvolvido por Meyer e Beucher (1990). O algoritmo WST tem sido muito utilizado no estudo de segmentação de imagens.

O método consiste na divisão sucessiva de regiões de acordo com seu gradiente, ou seja, regiões de maior variação são separadas das regiões de menor variação do gradiente. Porém o algoritmo apresenta uma segmentação excessiva quando aplicado em grãos por exemplo, portanto deve-se ter cautela para a escolha do algoritmo de segmentação que será utilizado para cada estudo de caso.

2.6. DTDML

O algoritmo proposto por Luo, Liu e Tao 2014 para a decomposição de imagens denominado *transfer distance metric learning* (aprendizagem métrica de distância de transferência). O algoritmo desenvolve uma métrica própria tomando como base outras métricas básicas. Em sumo, o algoritmo “aprende” uma combinação de métricas básicas forçando a métrica de segmentação a ser próxima a uma integração das métricas básicas de origem.

2.7. CNN

Redes neurais celulares (*convolutional neural network*) foram desenvolvidas por Chua (1988), a estrutura é baseada em estruturas neurais e foram desenvolvidas para aplicações envolvendo processamento de imagens.

A maioria das CNN's propostas é baseada em uma estrutura hierárquica que tem início com múltiplos estágios de módulos convolucionais e finaliza com um módulo de classificação. (Ye e colaboradores, 2018)

2.8. MSER

O algoritmo MSER (*maximally stable*
ISSN 2525-94585

extremal regions) ou máximo de regiões extremas estáveis é um algoritmo muito utilizado para a identificação de caracteres em imagens, sua rotina identifica regiões que são intrinsicamente mais claras ou escuras do seu entorno. O algoritmo MSER possui resultados satisfatórios quando aplicados em mudanças de iluminação, operações de mudança de escala, rotação e transformações afim. No entanto nem todas as imagens nas quais se deseja identificar uma sequência de caracteres possui as características citadas, o que inviabiliza a aplicação do algoritmo em alguns casos. (WEI e colaboradores, 2018).

2.9. SWT

Com o objetivo de corrigir as falhas apresentadas pelo algoritmo MSER, Ephstein e colaboradores (2011). O algoritmo chamado SWT (*stroke width transform*) ou transformação de largura de traçado.

O algoritmo parte do princípio de que caracteres dispostos em uma imagem possuem uma largura de traçado praticamente constante, a qual permite sua identificação. Trabalhos anteriores realizaram por Ephstein e colaboradores (2011) capturaram bordas em uma imagem em escala de cinza e em seguida, identificaram pares de arestas paralelas para calcular a largura do traçado em cada pixel, em seguida, pixels com larguras de traçado semelhantes foram agrupados em caracteres. Porém o trabalho proposto exige uma série de processos e parâmetros precisam ser definidos antes da aplicação do algoritmo, além do fato do algoritmo ser sensível a ruídos e borrões nas imagens devido ao fato de serem baseados na detecção de bordas (componentes de alta frequência).

2.10. RLB

O RLB (*Run Length Based*) é um método considerado híbrido apresentado por Basavanna e colaboradores (2013), o algoritmo utiliza como princípio os métodos MSER e SWT. O algoritmo é aplicado após a aplicação dos filtros de Sobel para realce de

bordas em uma imagem, o algoritmo calcula os comprimentos horizontais contando o número de pixels pretos consecutivos entre os pixels brancos e propõe um agrupamento para a classificação das frequências dos comprimentos de execução representando pixels de texto. WEI e colaboradores (2018).

3. Conclusão

Este trabalho apresenta os métodos matemáticos mais utilizados no processamento digital de imagens. O processamento digital de imagens tem sido largamente utilizado na indústria, medicina e agricultura com o propósito de agregar rapidez, economia e confiabilidade com o mínimo de contato humano no decorrer dos processos.

Os algoritmos utilizados para o processamento de imagens partem dos métodos mais básicos, os quais utilizam a distância métrica entre elementos para estimar a similaridade entre itens, um exemplo é o KNN e podendo ser aperfeiçoados para o DTDML como uma junção de métricas de avaliação.

Algoritmos inspirados em organismos humanos como redes neurais (NN), MLP e CNN são algoritmos mais complexos que envolvem a sobreposição de camadas, em alguns casos como o MLP a camada o erro da camada de saída pode ser propagado para camadas inferiores e correções podem ser feitas nos parâmetros do algoritmo, características não existentes nas redes NN.

Outra metodologia empregada nos algoritmos de processamento de imagens é a aplicação de filtros para um pré-tratamento da imagem, realçando características as quais o algoritmo demonstra maior sensibilidade, como no caso do algoritmo RLB.

O processamento digital de imagens é uma união da aprendizagem de máquina e visão computacional, que aliados a métodos matemáticos proporcionam novas técnicas. Sua flexibilidade e características de análise não destrutivas compõe a principal atratividade para sua aplicação nos mais diversos ramos e atividades.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001

REFERÊNCIAS

As referências são alinhadas somente à margem esquerda do texto e de forma a identificar o documento, separadas por espaços simples.

Sólus, D., Turán, L. O. J., “**Optical correlator in industrial image pattern recognition system**” Carpathian Journal of Electronic and Computer Engineering, vol. 9 ED-2, pp. 27-31, Jan. 2016

Righi, M., D’Acunto, M., Salvetti, O. “**An image enhancement tool: Patter Recognition image augmented resolution**”, Pleiades Publishing, vol. 26, pp. 518-523, May 2016.

X. ZOU and J. ZHAO, **Nondestructive Measurement in Food and Agro-Products**. Beijing, CN: Springer, 2015.

LEFEBVRE M, GIL DS, BRUNET E, NATONEK C, BAUR P, GUGERLI, PT, **Computer vision and agricultural robotics for disease control: the potato operation**. Computer Electron Agriculture vol 9, pp. 85–102, 1993.

AVUÇULA E, BASÇIFTÇI F, **New approaches to determine age and gender in image processing techniques using multilayer perceptron neural network**. Applied Software Computing, Elsevier, vol. 80, pp. 157-168, may 2018.

Meyer, F., Beucher, S., 1990. **Morphological segmentation**. J. Vis. Commun. Image Represent. 1 (1), 21–46.

Yong Luo, Tongliang Liu, Dacheng Tao,., **Decomposition-Based Transfer Distance Metric Learning for Image Classification**. iee transactions on image processing, vol. 23, no. 9, september 2014

L.O. Chua, L. Yang, IEEE Trans. Circuits Systems 35 (1988) 1257

B. Epshtein, E. Ofek, Y. Wexler, **Detecting text in natural scenes with stroke width transform**, Computer Vision and Pattern Recognition, 2011, pp. 2963–2970.

J.F. Canny, **A computational approach to edge detection**, IEEE Trans. Pattern Anal. Mach. Intell. 8 (6) (1986) 679–698.

M. Basavanna, P. Shivakumara, S.K. Srivatsa, G. Kumar, **Multi-oriented text detection in scene images**, Int. J. Pattern Recognit. Artif. Intell. 26 (07) (2013) 1255010– 1255010-19.

Wei Y., Wei S., Zeng D., Ye L., Zhang Z., **Multi-oriented text detection from natural scene images based on a CNN and pruning non-adjacent graph edges**, Signal Processing:Image Communication, Elsevier, Ed 64, pp. 89-98, march 2018.

Ye J., Shen Z., Behrani P., Ding F., Shi Y., **Detecting USM image sharpening by using CNN**, Signal Processing: Image Communication, Elsevier, Ed 64, pp. 258-264, may 2018.