

RECONHECIMENTO DE FORMAS GEOMÉTRICAS COM REDE NEURAL MULTILAYER PERCEPTRON

Dyego Alekssander Maas
dyego.maas@gmail.com
CPM-FURB, Inteligência Artificial

Henrique Oecksler Bertoldi
henriquebertoldi10@gmail.com
CPM-FURB, Inteligência Artificial

Resumo – O reconhecimento de formas geométricas num espaço bidimensional é um problema clássico cuja solução pode ter múltiplas aplicações. Este problema se torna mais difícil com a necessidade de obter invariância na rotação, escala e translação das formas apresentadas. Ao se extrair da imagem de entrada características invariantes a rotação, escala e translação, é possível utilizar lógica fuzzy e uma rede neural para classificar as formas geométricas. Este trabalho utiliza projeções radiais para encontrar pontos de amostra ao longo do contorno dos objetos, treinando uma rede neural para aprender a relação dos ângulos internos da forma geométrica. A rede neural utilizada é do tipo Feedforward (Multilayer Perceptron) multicamada, sendo treinada com o algoritmo Resilient Propagation (RPROP), através do uso do Encog Machine Learning Framework.

Palavras-Chaves – Lógica Fuzzy, Redes Neurais Feedforward, Multilayer Perceptron, RPROP, Encog Machine Learning Framework.

Abstract – The recognition of geometric shapes in a bidimensional space is a classical problem whose solution can have multiple applications. This problem becomes even more difficult with the necessity in obtaining invariant in rotation, scale and translation of the presented shapes. When extracting from the input image characteristics invariant to rotation, scale and translation, it is possible to use fuzzy logic and a neural network to classify the geometric shapes. This work uses radial projections to find sample points along the contours of the objects, training the neural network to learn the relations between the geometric shapes' internal angles. The neural network used is a Multilayer Feedforward Neural Network (Multilayer Perceptron), being trained with Resilient Propagation (RPROP) algorithm, through the use of Encog Machine Learning Framework.

Keywords – Fuzzy Logic, Feedforward Neural Networks, Multilayer Perceptron, RPROP, Encog Machine Learning Framework.

INTRODUÇÃO

O reconhecimento de figuras geométricas pode servir para uma série de aplicações. Já não é algo trivial, porém levando em consideração a invariância na rotação, escala e translação da imagem/figura, este reconhecimento pode se tornar mais trabalhoso e não tão trivial.

O objetivo deste trabalho é fazer com que a rede neural aprenda a relação entre os ângulos internos das formas geométricas apresentadas. Foi desenvolvido um algoritmo que irá reconhecer Triângulos, Elipses e Triângulos, através da capção das bordas e com relação dos ângulos internos da figura. Feito este reconhecimento, é feito o treinamento da rede neural do tipo Feedforward multicamada, através dos dados captados.

REDES NEURAIIS FEEDFORWARD

As redes neurais feedforward são o tipo mais simples e foram as primeiras criadas. Sua principal característica é que as conexões entre os neurododos não formam um ciclo dirigido. Desta forma, cada camada da rede neural alimenta a camada seguinte, não havendo ciclos internos dentro da rede neural.

O tipo mais simples de rede feedforward é o Single Layer Perceptron, onde existe apenas uma camada de saída. Os neurônios de entrada são diretamente conectados aos de saída através de uma série de pesos.

As redes do tipo feedforward são também comumente conhecidas por Multi Layer Perceptron, onde há ao menos uma camada oculta entre as camadas de entrada e saída. Cada neurônio das camadas ocultas efetua o somatório das entradas aplicando seu peso, que então passa por uma função de ativação, geralmente não linear, que também pode ser chamada de função do neurônio. Já os neurônios de saída não possuem qualquer função de ativação não linear além da aplicação dos próprios pesos.

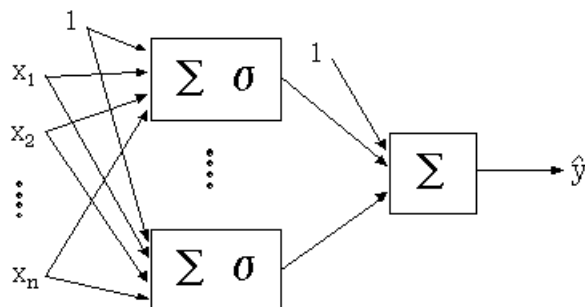


Figura 1: Rede neural feedforward com uma camada oculta e um neurônio de saída
Fonte: Nobre e Kim (2008, p. 31)

A rede neural utilizada foi do tipo Multi Layer Perceptron com uma camada oculta e função de ativação Sigmoidal. A função Sigmoidal é comumente utilizada em problemas de logística e produz números positivos entre 0 e 1. Normalmente é utilizada para dados de treinamento também na faixa de 0 e 1.

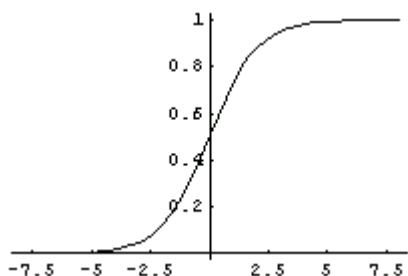


Figura 2: Função de ativação Sigmoid[x], {x, -8, 8}.

ENCOG MACHINE LEARNING FRAMEWORK

Encog é um framework avançado destinado ao aprendizado de máquina, que dá suporte há uma variedade de algoritmos específicos para tal fim. Algoritmos de Rede Neural, de Genética e Redes Bayesianas são alguns dos algoritmos suportados. A maioria dos treinamentos feito pelo Encog envolve multi-threads e podem ser escaláveis num ambiente multiprocessado.

Utilizamos o Encog para criação, treinamento e resultado da rede neural Feedforward.

EXTRAÇÃO DAS CARACTERÍSTICAS

O objetivo deste trabalho é fazer com que a rede neural aprenda a relação entre os ângulos internos das formas geométricas apresentadas. Para extrair estes ângulos alguns passos se fazem necessários. O primeiro passo é encontrar o centro da forma geométrica. Para isso podemos calcular o ponto médio através das seguintes fórmulas:

$$x_C = \frac{\sum_{i=0}^{n-1} x_i}{n}, \quad y_C = \frac{\sum_{i=0}^{n-1} y_i}{n},$$

A partir daí serão escolhidos pontos de amostra. Neste ponto foram realizadas projeções radiais a partir do centro para encontrar os pontos de amostra ao longo da borda, conforme visto na **figura 3**.

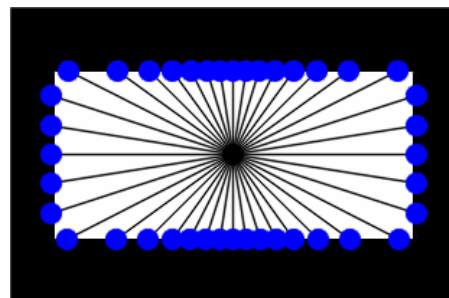


Figura 3: Projeção radial para encontrar os pontos de amostra (intersecção entre as linhas radiais e a a borda do objeto).

Para encontrar os ângulos internos, os pontos de amostra são agrupados em trios, formando dois segmentos de reta consecutivos. Para cada segmento, calculamos o ângulo interno entre os dois pontos e o centro da imagem, conforme a fórmula $\cos \theta = \frac{-(c_2)^2 + a^2 + b^2}{2 \cdot a \cdot b}$. O ângulo entre os dois segmentos é obtido somando-se os dois ângulos, conforme pode ser visto na imagem **figura 4**.

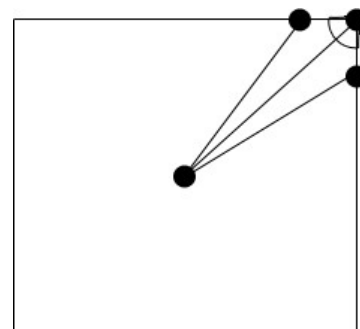


Figura 4: Identificando o ângulo entre os dois segmentos de reta.

Em seguida cada um dos ângulos calculados é classificado de acordo com uma lógica fuzzy (definida no próximo capítulo), obtendo assim um valor classificatório.

A entrada da rede neural será o valor do somatório dos valores classificatórios de todos os ângulos, codificado em binário. Deste modo, a camada de entrada da rede deverá ter neurônios de entrada suficientes para acomodar o número de bits ativos no pior caso, onde todos os ângulos estejam na faixa 3, entre 75 e 110 graus.

CLASSIFICAÇÃO FUZZY

Aos ângulos internos é aplicada uma lógica fuzzy para classificá-los conforme sua abertura:

- 2 para ângulos com menos de 75 graus;
- 3 para ângulos entre 75 graus e 110 graus;
- 1 para ângulos entre 110 e 145 graus;
- 0 para ângulos maiores que 145 graus.

Ao atribuir valor 0 (zero) aos ângulos maiores que 145 graus, fazemos com que grandes segmentos de reta sejam ignorados, e apenas os ângulos significantes sejam contabilizados na soma e influenciando na combinação binária que servirá como entrada para a rede neural. Um exemplo onde este comportamento pode ser facilmente observado são os retângulos, que possuem quatro ângulos na categoria 3, com todos os outros valores zerados.

ESTRUTURA DA REDE NEURAL

Dado o pior caso, onde todos os ângulos sejam da categoria três, o valor da soma será $3n$, sendo n é o número de pontos de amostra utilizados, o número de neurônios na camada de entrada será equivalente ao número de bits necessários para armazenar o resultado da soma dos valores classificatórios. Para 32 pontos de amostra, por exemplo, temos somatório de 96, que pode ser representado por 7 bits, e portanto 7 neurônios na camada de entrada. Neste trabalho foram utilizados 40 pontos de amostra, e portanto 7 neurônios na camada de entrada.

Como o objetivo é classificar os objetos em três classes geométricas (elipses, retângulos e triângulos), temos três neurônios na camada de saída.

Já na camada oculta, determinamos que os melhores resultados de aprendizado são obtidos com 10 neurônios. Em nossos testes foi constatado que abaixo de 8 neurônios nessa camada a taxa de erro no aprendizado sobe consideravelmente. Já a utilização de mais de 10 neurônios não melhora o resultado da aprendizagem. Outra estratégia testada foi a utilização de duas camadas com 5 neurônios cada, obtendo taxa de erros ligeiramente maior que uma única camada com 10 neurônios.

O algoritmo de treinamento utilizado foi o Resilient Propagation (RPROP), também conhecido como resilient backpropagation. Este algoritmo utiliza uma heurística de aprendizado supervisionado em redes neurais feedforward, cujo algoritmo é uma otimização do algoritmo de backpropagation. Assim como o algoritmo de correção em cascata, este também é um dos mais rápidos mecanismos de atualização de pesos.

RESULTADOS EXPERIMENTAIS

O software implementado é capaz fazer o reconhecimento de uma forma geométrica, dada uma imagem para análise. No entanto, nenhum pré-processamento é realizado na imagem, de modo que as imagens analisadas devem ter sido previamente transformadas. O software requer uma imagem monocromática, com a forma geométrica na cor branca com contornos na cor preta. Caso alguma imagem contenha múltiplos objetos, apenas o primeiro será analisado.

O treinamento da rede foi realizado com 66 amostras de elipses, 83 amostras de retângulos e 87 amostras de triângulos, cada qual com dimensão e rotação distintas.

O treinamento realizado pela rede neural resultou em uma taxa de erro de 0,075%, dadas as amostras citadas acima e os valores classificatórios gerados através da aplicação de lógica fuzzy. Diferentes configurações de valores teriam gerariam somatórios diferentes e, conseqüentemente, combinações binárias diferentes para alimentar o treinamento da rede neural.

Foram realizados testes com 10 imagens de cada forma geométrica, diferentes daquelas utilizadas para o treinamento, de diferentes dimensões e rotações. Os resultados podem ser vistos na **tabela 1**.

	Acerto (percentual)
Elipses	100%
Retângulos	70%
Triângulos	90%

Tabela 1: Percentual de acertos obtidos na classificação de 10 objetos de cada classe geométrica.

Todas as amostras classificadas incorretamente têm uma característica em comum: dependendo da rotação ou dimensões do objeto (um retângulo muito estreito, por exemplo), os pontos obtidos através da projeção radial podem não incluir os cantos dos triângulos ou retângulos, que formam os ângulos mais significantes para a classificação.

Uma solução para este problema é implementar um algoritmo para identificação dos cantos, para inclui-los em substituição ao ponto amostral mais próximo. Dessa forma, as características predominantes de cada forma geométrica serão conservadas e a performance da rede neural pode ser consideravelmente melhorada.

CONCLUSÕES E PERSPECTIVAS FUTURAS

Apresentamos neste trabalho uma técnica de reconhecimento de formas geométricas através do uso de projeções radiais, lógica fuzzy e uma rede neural forward (Multilayer Perceptron). A rede implementada é capaz de reconhecer três classes geométricas: triângulos, elipses e retângulos. De forma invariante em relação à escala, rotação e translação, este algoritmo busca as borda da figura e extrai características que irão servir para o treinamento da rede. Após treinada, ela é capaz de reconhecer elipses com perfeição. Triângulos atingem 90% de acertos e retângulos 70%.

Conforme citado anteriormente, a implementação de um algoritmo de detecção de cantos pode melhorar significativamente a performance da rede ao conservar as características principais de cada forma geométrica. Além disso, outras alterações podem trazer eventuais ganhos, como o ajuste dos valores de classificação fuzzy, a utilização de diferentes tipos de redes neurais, funções de ativação e algoritmos de treinamento.

REFERÊNCIAS BIBLIOGRÁFICAS

MIHU I. Z., GELLERT A., SUCIU C. N., “GEOMETRIC SHAPE RECOGNITION USING FUZZY AND NEURAL TECHNIQUES”, 1999

MIHU I. Z., GELLERT A., SUCIU C. N., “IMPROVED METHODS OF GEOMETRIC SHAPE RECOGNITION USING FUZZY AND NEURAL TECHNIQUES”,

MIHU I. Z., GELLERT A., SUCIU C. N., “HIERARCHICAL ARCHITECTURE IMPLYING FUZZY AND NEURAL TECHNIQUES FOR ON-LINE GEOMETRIC SHAPE RECOGNITION”

ENCOG MACHINE LEARNING FRAMEWORK,
<http://www.heatonresearch.com/encog>

NOBRE, KIM, “PROJEÇÕES EM LINHAS RADIAIS”,
(2008, p. 31)