



RECONHECIMENTO DE SÍMBOLOS EM EXPRESSÕES MATEMÁTICAS MANUSCRITAS

Cristiano P. Garcia, Nina S. T. Hirata (Orientadora)

Instituto de Matemática e Estatística, Universidade de São Paulo, Brasil



Introdução

Uma das etapas do reconhecimento de expressões matemáticas manuscritas é o reconhecimento individual dos caracteres e símbolos [2]. Em contraste aos problemas bem estudados de reconhecimento de dígitos ou caracteres do alfabeto [1], no caso de expressões matemáticas a quantidade de caracteres e símbolos a serem reconhecidos é bem maior.

Como consequência, reconhecedores tradicionalmente usados como as redes neurais artificiais são difíceis de treinar e tendem a apresentar desempenho insatisfatório.

O objetivo deste trabalho é decompor o problema de reconhecimento inicial em subproblemas utilizando características *online* e *offline* extraídas dos símbolos.

Métodos

Uma característica particular do reconhecimento de expressões matemáticas é a grande quantidade de possíveis símbolos a serem reconhecidos. Nesses casos, em geral, um único classificador não é capaz de fazer um reconhecimento satisfatório.

Na literatura encontramos diferentes abordagens para tratar este problema [3]. A seguir consideramos duas possíveis abordagens de utilização de redes neurais para esse tipo de classificação.

Uma possível abordagem é a utilização de uma única rede neural como classificador. Entretanto, já é fato conhecido através da literatura que as redes neurais sofrem piora no desempenho à medida que aumenta o número de classes.

Uma segunda abordagem proposta é a utilização de características extraídas dos próprios símbolos para dividir sucessivamente o conjunto total de forma que cada rede neural reconheça sempre um número de símbolos não superior a um determinado limiar (por exemplo, 20). Algumas características que podem ser extraídas dos símbolos são ilustradas abaixo [4]:

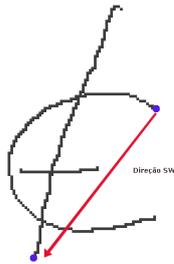


FIGURE 1: Direção do primeiro ponto ao último ponto.

Para a característica ilustrada acima, os símbolos são divididos em oito categorias: N, NE, E, SE, S, SW, W, NW. Esta classificação é feita de acordo com a direção, em linha reta, partindo do primeiro ponto do símbolo em direção ao último ponto.

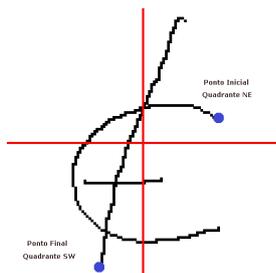


FIGURE 2: Posição do primeiro e do último ponto.

A figura acima ilustra a classificação dos símbolos de acordo com a posição do ponto inicial e do último ponto do símbolo. Os símbolos são classificados em quatro categorias: NE, SE, SW, NW.

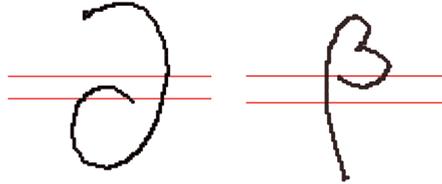


FIGURE 3: Região de maior densidade de pontos.

Também é possível classificar os símbolos de acordo com a região onde há uma maior densidade de pontos. Na figura acima vemos um exemplo no qual ocorre uma maior concentração de pontos na região inferior da imagem. Vemos também um outro exemplo onde ocorre uma concentração maior na região superior da imagem.

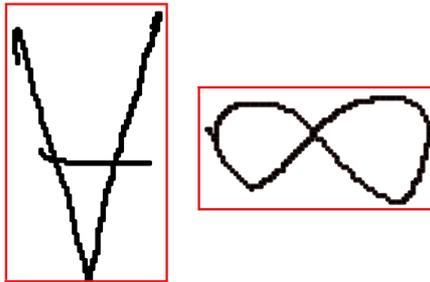


FIGURE 4: Relação entre altura e largura do símbolo.

Uma característica que é facilmente computada é a relação entre a altura e a largura do *bounding box* de um símbolo. A partir desta característica podemos dividir os símbolos em três categorias: altura maior que a largura, altura menor que a largura e altura aproximadamente igual à largura.

Além das características descritas acima, também foi utilizada nos testes o número de traços que compõem um símbolo. Neste caso é contado o número de vezes que a caneta toca o papel para escrever o símbolo.

De acordo com o conjunto de símbolos a ser reconhecido, é escolhido um subconjunto entre as características descritas acima. A partir deste conjunto de características é montada uma árvore na qual todos os nós internos no mesmo nível analisam os símbolos de acordo com a mesma característica. As folhas desta árvore são compostas por redes neurais artificiais conhecidas como *multi-layer perceptrons*.

A figura abaixo ilustra a árvore utilizada para o reconhecimento dos símbolos:

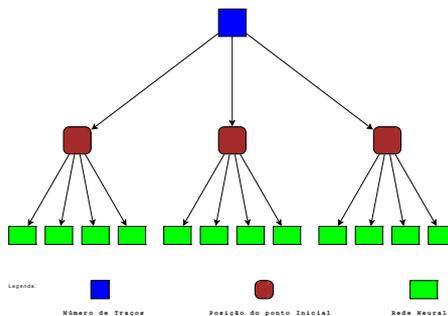


FIGURE 5: Árvore de reconhecimento.

Resultados

Para verificar se a abordagem hierárquica pode efetivamente melhorar a taxa de reconhecimento de um classificador, foram realizados testes nos quais foi necessário reconhecer um número crescente de símbolos comumente presentes em expressões matemáticas.

O gráfico abaixo resume os resultados obtidos:

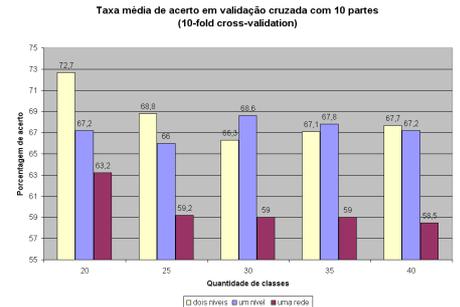


FIGURE 6: Reconhecimento de um grande número de classes.

Pode-se observar que para quase todas as quantidades de classes, o melhor desempenho foi obtido com o classificador hierárquico com dois níveis. O segundo melhor desempenho foi obtido com o classificador hierárquico com apenas um nível. Como era esperado, o classificador composto apenas por uma rede neural teve o pior desempenho nos testes realizados.

Conclusões

Neste pôster descrevemos um trabalho que encontra-se em andamento. A proposta deste trabalho é a utilização de uma abordagem hierárquica para o problema de classificação de símbolos em expressões matemáticas.

Nesta abordagem, os símbolos a serem classificados são divididos sucessivamente em subgrupos de acordo com alguma característica facilmente computável, até que cada um dos grupos possua um número de caracteres não superior a um dado limiar. Em seguida, para cada grupo é treinada uma rede neural para reconhecer os caracteres naquele grupo.

Implementamos e testamos redes neurais do tipo *multilayer perceptrons* com diferentes conjuntos de dados de caracteres e dígitos manuscritos. Esses testes mostraram resultados já conhecidos na literatura da área sobre o desempenho de redes neurais para classificação de dados: o desempenho das redes degrada quando o número de classes torna-se muito grande.

Experimentos preliminares mostram que a taxa de acerto de classificação pode ser melhorada com a abordagem hierárquica proposta. O atual desafio é a escolha das características que resultem numa hierarquia com a melhor taxa de acerto.

Agradecimentos

C. P. Garcia é bolsista ITI do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq. N. S. T. Hirata é bolsista Produtividade em Pesquisa do CNPq. Este trabalho está sendo desenvolvido no âmbito de projetos apoiados pelo CNPq.

Referências

[1] Christopher M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press Inc., New York, 1995.
[2] Kam-Fai Chan and Dit-Yan Yeung, *Mathematical Expression Recognition: A Survey*, International Journal on Document Analysis and Recognition **3** (2000), 3-15.
[3] Guohui Ou and Yi Lu Murphy, *Multi-class pattern classification using neural networks*, Pattern Recognition **40** (2007), 4-18.
[4] Stephen M. Watt and Xiaofang Xie, *Recognition for Large Sets of Handwritten Mathematical Symbols*, 2005, Eighth International Conference on Document Analysis and Recognition (ICDAR 2005), pp. 740-744.