

1º Congresso Brasileiro de Redes Neurais

Escola Federal de Engenharia de Itajubá
Itajubá, 24 a 27 de outubro de 1994

Reconhecimento Off-line de Assinaturas Utilizando MLP-Backpropagation e Momentum: um Estudo Comparativo¹

Herman Martins Gomes

Edson Costa de Barros Carvalho Filho

UFPE - Universidade Federal de Pernambuco
Departamento de Informática

CP 7851, 50.732-970, Recife, PE, Brasil

Fone: (081)271-8430, Fax: (081) 271-4925

Internet: {hmg, ecdbcf}@di.ufpe.br

Abstract

One of the most problems with off-line signature recognition is the drastical reduction of useful information due to the fact that all dynamic features are reduced to a single static image. A great variety of promising techniques, that have been established as appropriated to general tasks of pattern recognition, may be used in off-line signature recognition to better the performance, like Neural Networks and Momentum. The main goal of this paper is to discuss and to compare Neural Networks (represented by MLP-Backpropagation paradigm) and Moment-based techniques (represented by Standard Moments) in off-line signature recognition. We start with a brief introduction to the off-line signature recognition problem. Afterwards, we give a description of the investigated techniques. At the end, a signature database, the design of experiments and results are presented, and conclusions related.

1 Introdução

O reconhecimento automático de manuscritos representa uma nova tecnologia para facilitar a interface homem-máquina e para integrar melhor a utilização do computador na sociedade [12]. Um dos pontos de interesse na análise de assinaturas (e

manuscritos em geral) é a possibilidade de reconhecimento através do processamento de características próprias do autor que são depositadas, mesmo que inconscientemente, sobre o papel quando o mesmo escreve. Inúmeros sistemas estratégicos de informação requerem algum tipo de processamento automático de escrita, como aqueles encontrados em bancos para verificação da autenticidade de documentos. Como exemplo prático desta necessidade, a verificação, ainda manual, das assinaturas de cheques são um ponto crítico do sistema bancário brasileiro, que movimentou, só no primeiro semestre de 1992 cerca de 1,7 bilhão de cheques e outros papéis de compensação [9].

Há duas abordagens utilizadas no reconhecimento de assinaturas: off-line (ou estática) e on-line (ou dinâmica) [8, 12]. O reconhecimento off-line utiliza apenas características extraídas a partir de imagens digitalizadas, como a distribuição de *pixels* ou as variações de tonalidade. Já o reconhecimento on-line considera características dinâmicas, tais como, a aceleração da caneta, a pressão desempenhada, a seqüência de movimentos primitivos, dentre outras. Um dos grandes problemas que dificultam o reconhecimento off-line é a redução drástica do número de características observáveis, ao se considerar as assinaturas na forma de imagens estáticas. Portanto, é fundamental a utilização de técnicas efetivas para superar este tipo de problema, tais como Redes Neurais e Momentum.

Vários modelos de redes neurais [5, 13] têm sido propostos, investigados e utilizados em aplicações práticas, das quais se destacam os problemas de reconhecimento de padrões. Destes modelos, MLP Backpropagation [11] foi um dos que adquiriu grande popularidade, devido a sua simplicidade e poder computacional. Evidentemente, outros modelos também compartilham de propriedades semelhantes, mas, foi escolhido MLP-Backpropagation por este ser mais amplamente conhecido e utilizado. A adequabilidade de Redes Neurais aos problemas de reconhecimento de padrões tem sido confirmada a cada dia. Em particular, o Departamento de Informática da UFPE possui várias teses de mestrado, recentemente defendidas, e publicações na área, que ratificam esta afirmativa [2, 7, 14, 3, 4, 1]. Redes

¹Desenvolvido com apoio financeiro do CNPq

Neurais podem ser utilizadas tanto para a extração de características quanto para a classificação de imagens de assinaturas. No processo de extração de características, uma rede neural é capaz de fornecer uma representação compacta das imagens. Redes Neurais também são bastante adequadas ao processo de classificação, uma vez que não requerem conhecimento *a priori* das propriedades estatísticas das classes, ao contrário de outras técnicas da teoria de decisão, como os classificadores bayesianos, por exemplo.

Momentum tem se apresentado como uma técnica robusta para a decomposição de uma imagem de forma arbitrária em um conjunto finito de características, uma vez que é baseada na aplicação direta de transformações lineares, não havendo a necessidade de heurísticas específicas da aplicação. Em Estatística, Momentum é utilizado para caracterizar a distribuição de variáveis randômicas, e em Mecânica, serve para caracterizar corpos em função de sua distribuição espacial de massa. Considerando que uma imagem pode ser tratada como uma função de distribuição de densidade bi-dimensional, é possível aplicar Momentum para a análise de imagens, através da extração de propriedades que têm analogia em Estatística e Mecânica.

Considerando as dificuldades com o reconhecimento off-line de assinaturas, procura-se, através deste artigo, discutir a aplicação destas duas técnicas promissoras para extração de características e reconhecimento de imagens, descrever alguns experimentos de reconhecimento em que as técnicas são combinadas em diferentes configurações, apresentar os resultados obtidos, e tecer comentários e conclusões com base nestes resultados. Nas seções seguintes, apresentamos brevemente o modelo MLP-Backpropagation e a técnica baseada em Momentum utilizada.

2 MLP-Backpropagation

Multi Layer Perceptron-Backpropagation (MLP-BP) é um modelo de rede neural artificial que utiliza neurônios do tipo McCulloch-Pitts organizados em múltiplas camadas, possuindo um algoritmo de aprendizagem supervisionada denominado Back-Propagation [11], que baseia-se na correção do erro entre as saídas desejadas e as produzidas pela rede. Nos neurônios deste modelo, usualmente, é utilizada a função sigmóide (equação 1):

$$f(z) = 1/(1 + e^{-z}) \tag{1}$$

onde $z = \sum_j x_{ij}w_{ij} + \Theta_i$, que corresponde a soma ponderada das entradas pelos pesos mais um *bias*, que é similar a um *threshold* (limiar) no neurônio McCulloch-Pitts.

A arquitetura de uma rede MLP-BP é constituída por duas ou mais camadas de neurônios. A primeira camada é denominada camada de entrada e serve para distribuir os dados através da rede. As camadas seguintes são chamadas de camadas escondidas, e a última camada, de saída, responde com os valores computados pela rede (fig 1).

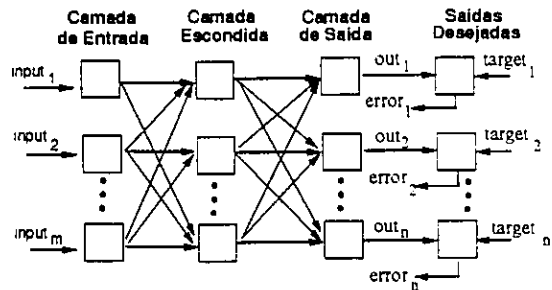


Figura 1: Arquitetura de uma rede MLP-BP

Toda informação codificada numa rede MLP se concentra nos pesos das conexões entre os neurônios, os quais podem ser vistos como parâmetros para as combinações e composições das funções de ativação (ou de transferência) dos neurônios. O processo de aprendizagem consiste no ajuste destes pesos de forma que todas as saídas produzidas pela rede representem um erro mínimo com respeito a um fator de vigilância. A capacidade de uma rede MLP aprender uma dada função está diretamente relacionada com a quantidade de neurônios nas camadas escondidas da rede.

Inúmeros sistemas para reconhecimento de padrões baseados no modelo MLP-Backpropagation tem sido propostos e investigados, como por exemplo o que foi definido em [12], para reconhecimento off-line de manuscritos. Várias outras aplicações podem ser encontradas em [5], [13] e [15].

3 Momentum

A fórmula geral para momentum bi-dimensional m_{pq} de ordem $p + q$, de uma função de distribuição de densidade, $f(x, y)$, é dada pela equação 2:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \tag{2}$$

Um conjunto de momentum de ordem n é formado por elementos do tipo m_{pq} , tais que $p + q \leq n$. Para uma imagem discretizada $g(x, y)$ de dimensões $M \times N$, a fórmula para momentum bi-dimensional é dada pela equação 3:

$$m_{pq} = \sum_{y=0}^{M-1} \sum_{x=0}^{N-1} x^p y^q g(x, y) \quad (3)$$

Valores de momentum de baixa ordem representam propriedades geométricas fundamentais de uma distribuição ou de um corpo. Por exemplo, no caso de imagens com função de distribuição binária, o momentum de ordem 0, $\{m_{00}\}$, representa a área total do objeto. Os valores de momentum de ordem 1, $\{m_{01}, m_{10}\}$, podem ser utilizados para calcular o centro de massa (\bar{x}, \bar{y}) , da imagem do objeto da seguinte forma:

$$\bar{x} = m_{10}/m_{00}, \quad \bar{y} = m_{01}/m_{00} \quad (4)$$

Se o objeto for posicionado de forma que seu centro de massa coincida com a origem do sistema de coordenadas ($\bar{x} = 0$ e $\bar{y} = 0$), então cada valor de momentum computado para aquele objeto é referido como momentum central, designado por μ_{pq} . Momentum de segunda ordem, $\{m_{02}, m_{11}, m_{20}\}$, conhecido como momentum de inércia, pode ser utilizado para determinar os eixos principais do objeto, sobre os quais há um momentum de segunda ordem mínimo e outro máximo (eixo principais maior e menor, respectivamente). A orientação do eixo principal mais próximo ao eixo x pode ser calculada pela equação 5:

$$\theta = \frac{1}{2} \tan^{-1} \left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right) \quad (5)$$

A representação de toda informação contida em um segmento de imagem requer uma quantidade de valores de momentum potencialmente infinita. Em termos práticos, a utilização de Momentum para o reconhecimento de imagens requer a seleção de um subconjunto de valores que contenha informação suficiente para caracterizar unicamente cada imagem. Um dos fatores que diferencia as técnicas baseadas em Momentum é o método para derivar valores invariantes a partir dos valores de momentum calculados sobre a imagem. As principais técnicas baseadas em Momentum que estão atualmente sendo pesquisadas incluem: Invariantes de Momentum, Momentum Rotacional, Momentum Ortogonal, Momentum Complexo e Momentum Padrão [10]. Há duas formas de se obter invariância: através de invariantes algébricos ou através de normalização de características. Em [4], demonstrouse, experimentalmente, que técnicas baseadas em normalização de características tendem a ser mais adequadas ao processamento de imagens de assinaturas do que invariantes algébricos.

Será utilizado Momentum Padrão, uma vez que representa uma técnica não baseada em in-

variantes algébricos, mas sim em um critério de normalização para escala, posição e orientação. Primeiramente, um conjunto de momentum $\{m_{pq}\}$, de uma ordem desejada, é computado de uma dada imagem, utilizando a equação 3. A partir de então, uma série de operações de normalização quanto a escala, posição e orientação é aplicada sobre o conjunto inicial de momentum, para derivar um conjunto padrão de momentum, $\{M_{pq}\}$. Todas as operações atuam apenas no domínio do momentum, não havendo necessidade de manipulação dos pixels da imagem, e combinam os valores de momentum originais de mesma ordem ou menor. Uma descrição detalhada das operações podem ser encontradas no trabalho anterior [4].

Em [10] são relacionados vários trabalhos que tratam da utilização de Momentum em aplicações como compactação de imagens, reconhecimento de caracteres ópticos, objetos sólidos, imagens aéreas, e compressão de imagens, dentre outras.

4 Base de Dados de Assinaturas

Para a realização dos experimentos, foi utilizada uma base de dados de assinaturas que está sendo desenvolvida no Departamento de Informática da UFPE. Atualmente, a base de assinaturas conta apenas com amostras verdadeiras de uma comunidade restrita. A intenção foi a de avaliar preliminarmente a adequabilidade de diferentes técnicas no processo mais geral (e mais simples) de reconhecimento de assinaturas, que trata da separação das classes apenas, não considerando, ainda, o problema mais complexo de reconhecimento (classificação) que é a verificação de assinaturas falsificadas. Em sua versão final, a base de dados deverá possuir um número significativo de assinaturas verdadeiras e correspondentes falsificações profissionais (adquiridas em agências bancárias [9]), além de englobar as culturas gráficas² de diferentes regiões do Brasil.

Na versão atual da base de dados, as assinaturas foram coletadas em formulários especiais de 16 x 24cm, com capacidade para 20 amostras. Cada formulário foi utilizado para representar as amostras de uma classe de indivíduo. Um total de 7 classes, com 20 assinaturas cada, foram coletadas entre membros do DI/UFPE. As imagens dos formulários foram digitalizadas em um scanner de mesa, utilizando a resolução de 150dpi e 256 tons de cinza. A combinação de técnicas para equalização de histograma e filtros digitais foi utilizada para melhorar a qualidade das imagens. Em seguida, foi aplicado um algoritmo heurístico (baseado no layout dos formulários) para segmentar as assinaturas em

²Entende-se por cultura gráfica o estilo utilizado na escrita das assinaturas, como o tipo das letras, a presença de sinais não alfabéticos etc

retina de dimensões iguais a 350 x 90 pontos. Depois de segmentadas, as imagens foram binarizadas segundo um threshold que permitiu filtrar ruídos sem haver deteriorização da imagem. Com o intuito de permitir avaliar a robustez das técnicas de reconhecimento, não houve normalização das imagens quanto a orientação e escala. Na figura 2 são apresentadas algumas assinaturas pertencentes à base de dados construída.

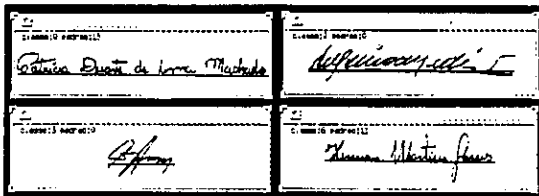


Figura 2: Assinaturas digitalizadas

5 Experimentos

Para cada uma das classes C_j ($j = 0, \dots, 6$) da base de dados, foram construídos aleatoriamente dois conjuntos de assinaturas, um para treinamento dos classificadores (L_j) e outro para os testes de reconhecimento (T_j). Definiu-se, a priori, que ambos os conjuntos teriam o mesmo número de assinaturas, sendo este igual a metade das amostras disponíveis por classe (=10). Foram definidas três configurações para os experimentos de reconhecimento, apresentadas na figura 3.

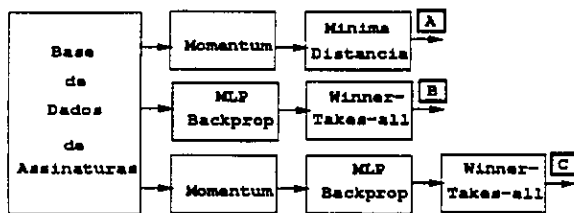


Figura 3: Configurações dos Experimentos

Na configuração **A**, Momentum foi utilizado para extração de características, e um classificador de distancia euclidiana mínima no espaço de características foi utilizado para reconhecimento. Foram considerados 6 valores para compor o vetor de características, assim definido: $x = (M_{02}, M_{20}, M_{12}, M_{21}, M_{30}, M_{03})$. Cada classe foi representada pelo vetor de características médio (protótipo) do conjunto de treinamento (equação 6).

$$m_j = \frac{1}{10} \sum_{x \in L_j} x \quad j = 0, 1, \dots, 6 \quad (6)$$

Utilizou-se um critério de rejeição com base nos diâmetros dos conjuntos de treinamento de cada classe. O diâmetro do conjunto de treinamento é definido como sendo a maior de todas as distâncias encontrada entre os padrões do conjunto. Quando a menor distancia $Dmin_j(x_i)$ computada entre um padrão de teste x_i e um dado protótipo m_j for superior a um limiar θ multiplicado pelo diâmetro δ_j da classe candidata C_j , então deve-se rejeitar o padrão de teste (equação 7). Os testes de reconhecimento com esta configuração foram realizados variando linearmente θ entre 0 e 2, com incremento de 0.01.

$$\text{Rejeitar } x_i \text{ se : } Dmin_j(x_i) > \theta \times \delta_j \quad (7)$$

Na configuração **B**, utilizou-se uma rede neural MLP-Backpropagation de duas camadas escondidas, recebendo diretamente como entrada uma retina de 350 x 90 pontos contendo a imagem de uma assinatura (figura 4). A conectividade entre os neurônios de duas camadas consecutivas é total. A camada de saída foi dimensionada com 7 neurônios (numerados de 0 a 6, cada qual representando uma classe) com o intuito de permitir uma classificação baseada no critério Winner-Takes-All (WTA). Desta forma, na fase de teste, o número da classe reconhecida pela rede é precisamente o número do neurônio que produziu o maior valor na camada de saída. Um critério para rejeição foi definido da seguinte forma: quando a diferença entre as duas maiores saídas $Omax'$, $Omax''$ produzidas pela rede for inferior a um limiar ϕ de rejeição, então deve-se rejeitar o padrão x_i de teste (equação 8). Para a avaliação das taxas de reconhecimento, foram utilizados valores de ϕ variando linearmente entre 0 e 1, com incremento de 0.01. Diferentes valores para o erro de treinamento³ da rede foram testados. Verificou-se que o valor 0.02 foi o que produziu os melhores resultados.

$$\text{Rejeitar } x_i \text{ se : } |Omax' - Omax''| < \phi \quad (8)$$

Na configuração **C**, acrescentou-se antes da rede neural uma etapa de extração de características, utilizando os mesmos valores de momentum da configuração **A**. Espera-se, com esta configuração, obter-se um ganho em relação às anteriores, uma vez que combina os resultados das duas

³erro de treinamento é maior diferença encontrada entre uma saída desejada e uma saída produzida pela rede, para todos os padrões de treinamento

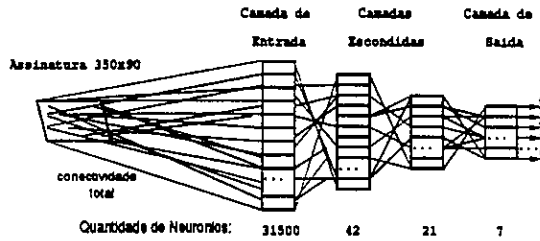


Figura 4: Arquitetura MLP-Backpropagation

| Config. | Acerto | Erro | Rejeição | Limiar |
|---------|--------|------|----------|-----------------|
| A | 87.1 | 0.0 | 12.9 | $\theta = 1.03$ |
| B | 82.9 | 0.0 | 17.1 | $\phi = 0.82$ |
| C | 92.9 | 0.0 | 7.1 | $\phi = 0.70$ |

Tabela 1: Índices de acerto (em %) com erro mínimo

técnicas. A arquitetura de rede utilizada foi quase idêntica àquela da figura 4, a menos da camada de entrada, que desta vez possui 6 neurônios (um para cada valor de momentum do vetor de características). O classificador e as estratégias de treinamento e teste foram praticamente os mesmos da configuração anterior. A única diferença é que o erro de treinamento mais adequado foi de 0.1.

6. Implementação

Todo o código fonte para definição e acesso à base de dados bem como para a implementação de Momentum e dos classificadores foi escrito na linguagem C-Unix. A rede MLP-Backpropagation foi implementada utilizando o simulador Aspirin/MIGRAINES [6]. Os experimentos foram executados numa estação de trabalho Sparc 2 da Sun microsystems.

7 Resultados

Foram computados sobre os conjuntos de padrões de teste, os índices médios de acerto, erro e rejeição obtidos pelas diferentes configurações e variação na faixa de valores assumida pelos limiares de rejeição. Na tabela 1, apresentam-se os limiares de rejeição de cada configuração que produziram os menores índices médios de erro. E, na tabela 2, apresentam-se os limiares de rejeição associados aos maiores índices médios de acerto. Na figura 5 são apresentados gráficos evolutivos dos índices de reconhecimento com respeito aos diferentes limiares de rejeição.

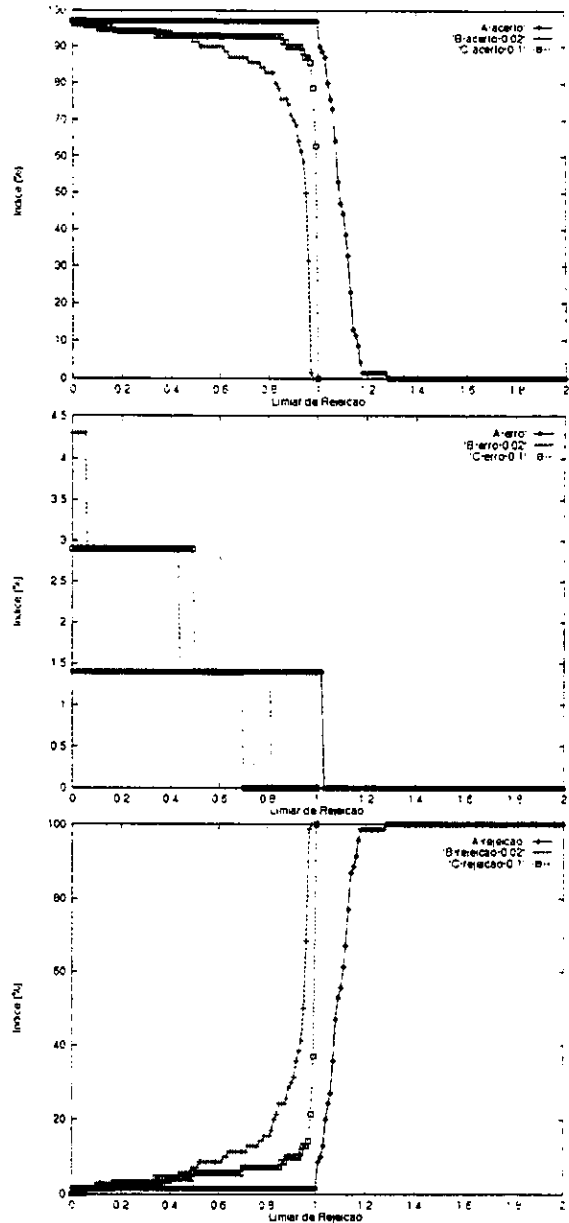


Figura 5: Gráficos evolutivos dos índices de acerto, erro e rejeição

| Config. | Acerto | Erro | Rejeição | Limiar |
|---------|--------|------|----------|-----------------|
| A | 97.1 | 1.4 | 1.4 | $\theta = 1.00$ |
| B | 92.9 | 1.4 | 5.7 | $\phi = 0.48$ |
| C | 97.1 | 2.9 | 0.0 | $\phi = 0.05$ |

Tabela 2: Maiores índices de acerto(em %)

8 Conclusões

Analisando as tabelas 1 e 2, verifica-se que a técnica MLP-BP (configuração **B**) sozinha não produziu bons resultados. Como era esperado, verificou-se que a configuração **C**, combinando MLP-BP e Momentum foi a que produziu o melhor índice médio de acerto (92.9%), quando fixando um erro mínimo (tabela 1), e um índice de acerto igual ao da configuração **A** (97.1%), quando o critério de erro mínimo não foi considerado (tabela 2).

Com base nestes resultados, pode-se dizer que a configuração **C** foi mais adequada ao processo de reconhecimento de assinaturas que as demais configurações. Assim, num sistema para reconhecimento off-line de assinaturas, em face da grande variabilidade e complexidade dos padrões, é interessante utilizar um processo de extração de características para reduzir a dimensionalidade dos dados e extrair a informação de interesse, mesmo quando se está utilizando uma técnica supostamente robusta no sentido de abstrair características relevantes dos dados de treinamento e generalizar para os demais padrões, como é o caso de Redes Neurais. Outra conclusão que pode ser extrapolada a partir dos experimentos e resultados é que, considerando os problemas com o processamento de imagens de manuscritos devido a redução drástica das características dinâmicas do sinal, é melhor integrar mais de uma técnica, como na configuração **C**, a fim de se obter melhores resultados.

Apesar do estágio inicial da base de dados de assinaturas, as configurações desenvolvidas e os resultados obtidos trazem uma importante contribuição tanto para a área de Redes Neurais quanto para a área de reconhecimento de padrões, na medida em que confirmaram experimentalmente as excelentes propriedades de Redes Neurais (MLP-BP), Momentum, e sua combinação no reconhecimento off-line de assinaturas. Espera-se que a construção de uma base de dados mais completa e o desenvolvimento de experimentos para a verificação de assinaturas falsificadas tornem operacionalmente mais práticas contribuições deste trabalho.

Referências

- [1] D. L. Bisset, E. C. D. B. C. Filho, and M. C. Fairhurst. A comparative study of neural network structures for practical application in a pattern recognition environment. In *Proc. First IEE International Conference on Artificial Neural Networks*, pages 378-382, London, UK, October 1989. IEE.
- [2] E. C. de Barros Carvalho Filho. Reconhecimento de assinaturas usando redes neurais artificiais. Master's thesis, Universidade Federal de Pernambuco, 1987.
- [3] A. C. P. de Leon Ferreira de Carvalho. Reconhecimento de Sequências utilizando neurônios booleanos. Master's thesis, Universidade Federal de Pernambuco, 1990.
- [4] H. M. Gomes and E. C. D. B. C. Filho. Reconhecimento off-line de assinaturas utilizando momentum. a ser apresentado na XX Conferência Latinoamericana de Informática - CLEIP'94, Cidade do México, México, 1994.
- [5] R. Hecht-Nielsen. *Neurocomputing*. Addison-Wesley Publishing Company, 1990.
- [6] R. R. Leighton. The Aspirin/MIGRAINES Neural Network Software - Release V6.0. Technical Report MP-91W00050, MITRE Corporation, october 1992. User's Manual.
- [7] T. B. Ludemir. Discriminação de Sequências com redes neuronais digitais. Master's thesis, Universidade Federal de Pernambuco, 1987.
- [8] R. Plamondon and G. Lorette. Automatic signature verification and writer identification - the state of the art. *Pattern Recognition*, 5(21):525-537, 1988.
- [9] Projeto assinatura eletrônica, maio 1993. Convênio de Cooperação Técnico-Científica firmado entre o Banco do Brasil (CEDIP/Recife) e o Departamento de Informática da UFPE.
- [10] R. Prokop and A. Reeves. A survey of moment-based techniques for unoccluded object representation and recognition. *CVGIP: Graphical Modules and Image Processing*, 54(5):438-460, september 1992.
- [11] D. E. Rumelhart and R. J. Williams. Learning internal representations by error propagation. *Parallel Distrib. Processing*, 1:318-62, 1986.
- [12] A. W. Senior. Off-line handwriting recognition: A review and experiments. Technical Report CUED/F-INFENG/TR 105, Cambridge University Engineering Department, december 1992.
- [13] P. K. Simpson. *Artificial Neural Systems*. Pergamon Press, 1990.
- [14] G. C. Vasconcelos. Modelo probabilístico da dinâmica de recuperação de erros em redes neurais, booleanas. Master's thesis, Universidade Federal de Pernambuco, 1991.
- [15] P. D. Wasserman. *Neural Computing: Theory and Practice*. Van Nostrand Reinhold, 1989.