

Introdução ao Reconhecimento de Padrões utilizando Redes Neurais

Caimi Franco Reis
Marcelo Portes de Albuquerque
Sérgio Borges de Castro

Centro Brasileiro de Pesquisas Físicas – CBPF
Rua Dr. Xavier Sigaud, 150
22290-180- Rio de Janeiro, RJ

Resumo

Esta nota técnica tem o intuito de explicar as idéias básicas que envolvem o processamento de imagens digitais, as redes neurais e o reconhecimento de padrões, além de mostrar como as redes neurais auxiliam no reconhecimento de padrões.

Sumário

Resumo

Introdução

Capítulo 1: Processamento de imagens digitais

1.1-Histograma de uma imagem

1.1.1-Equalização de um histograma

1.2-Técnicas utilizadas no processamento de imagens digitais

1.2.1-Suavização espacial

1.2.2-Realce de bordas

1.2.3-Binarização

1.2.4-Detecção de bordas

Capítulo 2: Redes neurais

2.1-Introdução às redes neurais artificiais

2.2-Histórico

2.3-Composição das redes

2.4-Classificação das redes

2.5-A Rede Neural Artificial

2.6-Backpropagation

Capítulo 3: Reconhecimento de padrões utilizando redes neurais artificiais

3.1-Reconhecimento de padrões

3.2-Técnicas para o reconhecimento de padrões

3.3-Relações entre redes neurais artificiais e reconhecimento de padrões

Conclusão

Introdução

O processamento digital de imagens consiste no conjunto de técnicas voltadas para a análise de dados de várias dimensões adquiridos por diversos tipos de sensores. O processamento é usado para melhorar o aspecto visual da imagem original, e fornecer alguns subsídios para a sua interpretação, gerando produtos que possam ser submetidos a outros processamentos posteriores. Dentre as áreas em que ele é utilizado estão: análise de recursos naturais e meteorologia por meio de imagens de satélites, transmissão digital de sinais de televisão, análise de imagens biomédicas, obtenção de imagens médicas por ultra-som, radiação nuclear ou técnicas de tomografia computadorizada, aplicações em automação industrial envolvendo o uso de sensores visuais em robôs.

O sistema de processamento de imagens é constituído de diversos processos, tais como: digitalização, pré-processamento, segmentação, pós-processamento, extração de atributos, classificação e reconhecimento. A digitalização é a conversão de um sinal da forma analógica para a forma digital. O pré-processamento refere-se ao processamento inicial de dados brutos para calibração radiométrica da imagem, correção de distorções geométricas e remoção de ruído. A segmentação consiste na etapa de processamento em que se analisa a imagem do ponto de vista da informação nela presente, ou seja, consiste em dividir a imagem em diferentes regiões, que serão posteriormente analisadas por algoritmos especializados em busca de informações de alto-nível. O pós-processamento é a etapa em que se corrigem os defeitos da segmentação, já que muitas vezes, o resultado da segmentação não é adequado. Para corrigir os defeitos residuais nesta etapa, utilizam-se as técnicas da morfologia matemática. A extração de atributos é a fase em que se realizam medidas sobre a imagem. Estas medidas podem ser muito sofisticadas, permitindo uma nova separação dos objetos em classes de similaridade, em função dos parâmetros medidos, fase esta chamada de classificação.

Já o reconhecimento de padrões é a área do processamento em que se utilizam técnicas computacionais para aproximar a performance das máquinas a do ser humano de modo a reconhecer padrões dentro de uma imagem. Há duas grandes motivações para o estudo nesta área: a necessidade das pessoas em se comunicarem com máquinas computacionais através de linguagens naturais, e o interesse na idéia de projetar e construir autômatos (máquinas inteligentes) que possam realizar certas tarefas com habilidades comparáveis à performance humana. O reconhecimento de padrões é uma ciência não-exata. Esta natureza não estruturada do padrão a ser reconhecido torna o problema de reconhecimento difícil de ser tratado por paradigmas tradicionais de computação. Desta forma, o uso de redes neurais artificiais aparece como modo alternativo de solução, pois, ao invés de criar procedimentos lógicos, a construção destas redes envolve o entendimento informal do comportamento desejado para atender ao problema. Neste trabalho, dar-se-á um enfoque em quatro técnicas de processamento de imagens: realce de bordas, binarização, detecção de bordas e suavização espacial.

Esta nota técnica foi elaborada com o objetivo de discutir no primeiro capítulo o processamento de imagens, suas utilidades e técnicas, bem como fundamentos de imagens digitais. No segundo capítulo será apresentada uma visão geral de redes neurais artificiais. No terceiro capítulo será focado o modo como as redes neurais artificiais auxiliam no reconhecimento de padrões.

Capítulo 1

Processamento digital de imagens

Processar uma imagem consiste em transformá-la sucessivamente com o objetivo de extrair mais facilmente a informação nela presente. O objetivo do processamento de imagens consiste em processar uma imagem de modo que o resultado final seja mais adequado que a imagem original para uma aplicação específica.

Antes de se iniciar a conceituação das técnicas de processamento propriamente ditas, procurar-se-á caracterizar um elemento importante na análise da imagem: seu histograma.

1.1-Histograma de uma imagem

O histograma de uma imagem representa, para cada nível de intensidade (cinza), o número de pixels com aquele nível. Ele indica se a imagem está distribuída adequadamente dentro dos possíveis níveis. Se a imagem não está utilizando todos os níveis disponíveis, pode-se alterá-los, para melhor uso. Este tipo de filtragem chama-se manipulação de histograma.

Usando uma linguagem matemática, o histograma de uma imagem digital, com níveis de cinza na escala $[0, L-1]$, é a função discreta

$$p(r_k) = n_k/n \quad (1.1)$$

onde r_k é o k -ésimo nível de cinza, n_k é o número de pixels na imagem com aquele nível de cinza, n é o número total de pixels na imagem, e $k = 0, 1, 2, \dots, L-1$.

Em outras palavras, $p(r_k)$ dá uma estimativa da probabilidade de ocorrência do nível cinza r_k . Um gráfico desta função para todos os valores de k mostra uma descrição global da aparência da imagem. Por exemplo, com a figura 1.2 foram construídos quatro tipos básicos de imagens, conforme mostra a figura 1.3, onde a figura 1.3(a) é uma imagem com características pretas, 1.3(b) imagem com características brancas, 1.3(c) uma imagem de baixo contraste e 1.3(d) de alto contraste. O histograma das figuras anteriores foi exibido na figura 1.4, onde (a) mostra que os níveis de cinza estão concentrados ao redor do preto na escala de níveis de cinza, ou seja, ele tem características predominantemente escuras ou pretas. O oposto é verdadeiro na figura 1.4(b). O histograma mostrado na Figura 1.4(c) tem uma forma concentrada no meio do gráfico, que indica uma escala pequena e corresponde a uma imagem de baixo contraste. Como todos níveis de cinza ocorrem em torno do meio da escala, a imagem aparecerá com baixo contraste. Finalmente a figura 1.4(d) mostra um histograma com um espalhamento significativo, correspondendo a uma imagem com alto contraste.



Figura 1.3(a) imagem com baixos níveis de cinza; (b) imagens com altos níveis de Cinza; (c)imagem com baixo constraste; (d) imagem com alto constraste;

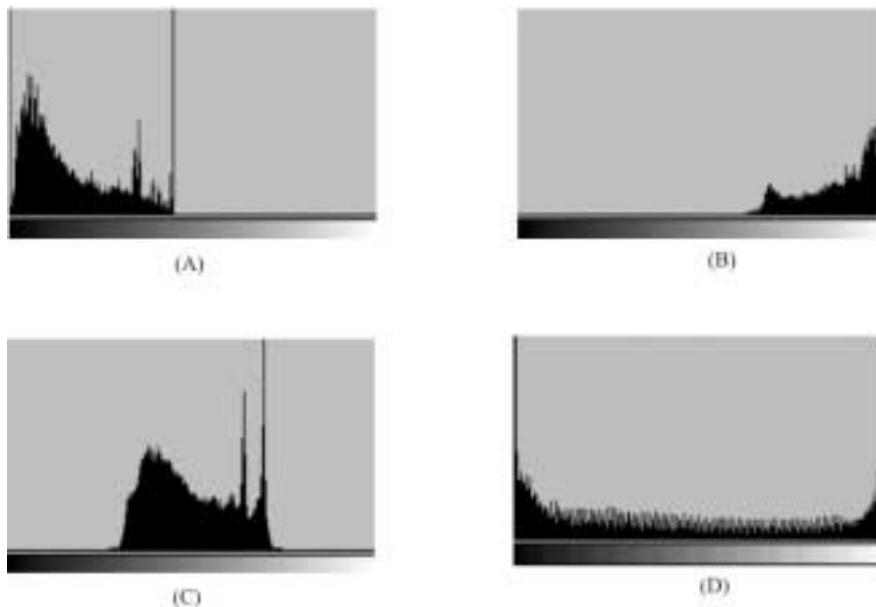


Figura 1.4-(a) histograma de uma imagem com baixo nível de cinza; (b) histograma com alto nível de cinza; (c) histograma de uma imagem de baixo constraste; (d) histograma de uma imagem de alto constraste.

1.1.1-Equalização do Histograma

Tomando-se a variável r para representar os níveis de cinza na imagem a ser realçada. Na parte inicial deste trabalho, a suposição é que os valores do pixels são quantidades contínuas que podem ser normalizadas de modo a ficarem no intervalo

[0,1], com $r = 0$ representando o preto e $r = 1$ representando o branco. Depois, considerou-se uma formulação discreta permitindo valores de pixels estarem no intervalo $[0, L-1]$.

Para qualquer r no intervalo $[0,1]$, aplica-se a transformação da forma

$$s = T(r) \quad (1.2)$$

que produz um nível s para cada valor de pixel r na imagem original. Foi assumido que a função de transformação dada na equação (1.2) satisfaz as seguintes condições:

- a) $T(r)$ é escolhida (auto-valor) e monotonamente crescente no intervalo $0 \leq r \leq 1$;
- b) Possui um único valor para cada nível de cinza;
- c) $0 \leq T(r) \leq 1$ para $0 \leq r \leq 1$.

A condição (a) preserva a ordem do preto para o branco na escala cinza, enquanto que a condição (c) garante um mapeamento consistente com a escala de valores de pixels.

1.2-Técnicas de processamento de imagens

As técnicas envolvem duas categorias principais: métodos que operam no domínio espacial e métodos que operam no domínio da frequência. O domínio espacial refere-se ao próprio plano da imagem, e as técnicas nesta categoria são baseadas na manipulação direta dos pixels de uma imagem. As técnicas de processamento no domínio da frequência se baseiam na modificação da transformada de Fourier de uma imagem.

Para exemplificar as técnicas apresentadas a imagem da figura 1.1 será usada várias vezes.



Figura 1.1- imagem padrão utilizada como exemplo

1.2.1-Suavização espacial

Suavização busca uma homogeneização dos pixels presentes nas diversas regiões das imagens, alterando pixels com níveis de cinza pouco semelhantes aos da vizinhança e que podem representar um ponto ruidoso. No domínio da frequência, suavização é obtida através de filtros passa-baixa, que reduz os detalhes da imagem, enquanto que os filtros passa-alta realçam os detalhes da imagem, conforme pode-se observar na figura 1.2. No domínio espacial, através de ações realizadas dentro dos limites de uma máscara, que se desloca sobre toda a imagem efetuando operações lineares e não-lineares baseadas em informações de uma vizinhança 3x3 do pixel atual (ou vizinhança-8). O nível de cinza do pixel central da janela de imagem definida pela

máscara é substituído por um valor que é função do método empregado e dos níveis de cinza da vizinhança definida pela janela. Estas técnicas geralmente incorporam características do ruído, o conhecimento a princípio sobre bordas e propriedades do sistema visual humano para obter o efeito desejado. Técnicas de suavização têm como objetivos principais a remoção de ruído e a uniformização dos níveis de cinza dos pixels na imagem.



Figura 1.2 (a) imagem após o uso de um filtro passa-baixa; (b) imagem após o uso de um filtro passa-alta

1.2.2-Realce de bordas

Bordas são características primitivas de uma imagem que são largamente utilizadas em sistemas de classificação e análise de imagens. Uma borda, é definida como sendo uma mudança ou descontinuidade local na luminosidade de uma imagem. O realce de bordas é obtido no domínio da frequência, através de filtros passa-alta, conforme a figura 1.3, e no domínio espacial, por máscaras utilizando operadores diferenciais e direcionais. Seus objetivos principais consistem em realçar finos detalhes em uma imagem ou realçar detalhes que tenham sido borrados por erro ou por um efeito natural de um método particular de aquisição de imagens.



Figura 1.3-imagem que foi realçada usando um filtro passa-alta

1.2.3-Binarização

Binarização de imagem é uma importante etapa do processamento em muitos sistemas de análise de imagem, onde se deseja segmentar determinados objetos. Geralmente, o método de segmentação é baseado na associação de informações relevantes da imagem, tais como: distribuição dos níveis de cinza, média local dos níveis de cinza, e orientação do gradiente dos pontos de imagem. Na figura 1.1 foi feita uma binarização para o nível 128 da escala dos níveis de cinza, resultando na figura 1.4.



Figura 1.4: Exemplo de uma imagem que foi binarizada

1.2.4-Deteccção de Bordas

Definindo borda como sendo a fronteira entre duas regiões com valores de níveis de cinza distintos e assumindo que estas duas regiões em questão são homogêneas de modo que a transição entre elas possa ser baseada unicamente nas discontinuidades dos níveis de cinza.

A idéia da maioria das técnicas de detecção de bordas está fundada no cálculo de uma variação dos valores que um pixel pode assumir em uma determinada imagem. A primeira derivada deste modelo apresentado é 0 em todas as regiões com níveis de cinza constantes e assume um valor constante na transição dos níveis de cinza. A segunda derivada, por outro lado, é zero em todas as regiões exceto no início e no fim das transições dos níveis de cinza. Baseado nisto é evidente que a primeira derivada detecta a presença de uma borda, enquanto que a segunda derivada pode ser usada para determinar se o pixel está sobre um fundo escuro ou sobre um objeto claro. A primeira derivada de qualquer ponto de uma imagem pode ser obtida avaliando a magnitude do gradiente naquele ponto, enquanto que a segunda derivada é dada por um laplaciano. Um exemplo da detecção de borda é mostrado na figura 1.5, em que se usa um pequeno raio de detecção e um alto de raio de detecção.

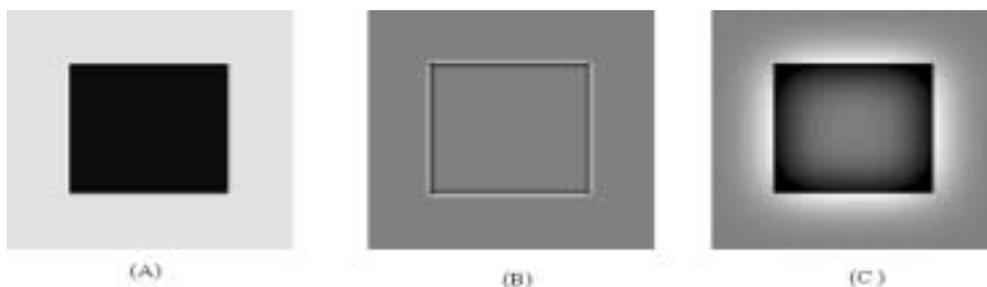


Figura 1.5- (a)imagem utilizada para a exemplificação; (b)imagem com pequeno raio de detecção; (c)imagem com alto raio de detecção

Capítulo 2

Redes neurais artificiais

2.1-Introdução às Redes neurais artificiais

As redes neurais artificiais são técnicas computacionais que apresentam um modelo inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento através da experiência. Uma grande rede neural artificial pode ter centenas ou milhares de unidades de processamento, enquanto que o cérebro de um mamífero pode ter muitos bilhões destas unidades. Apesar das redes neurais artificiais não permitirem uma única definição, as linhas seguintes seguem como uma tentativa de generalizar o modo como funciona uma rede.

De forma geral, a operação de uma célula da rede se resume em: cada sinal é apresentado à entrada, é multiplicado por pesos que indicam a influência deles na saída da unidade, e em seguida é feita a soma ponderada dos sinais que produz um nível de atividade.

2.2-Histórico

As primeiras informações mencionadas sobre redes neurais datam de 1943, em artigos de McCulloch e Pitts, em que sugeriam a construção de uma máquina inspirada no cérebro humano. Muitos outros artigos e livros surgiram após isso, porém, por um longo período de tempo, pouco resultado foi obtido. Até que em 1949 Donald Hebb escreveu um livro intitulado como *The Organization of Behavior* (A Organização do Comportamento) que perseguia a idéia de que o condicionamento psicológico clássico está presente em qualquer parte dos animais pelo fato de que esta é uma propriedade de neurônios individuais. Suas idéias não eram completamente novas, mas Hebb foi o primeiro a propor uma lei de aprendizagem específica para as sinapses dos neurônios. Este primeiro e corajoso passo serviu de inspiração para que muitos outros pesquisadores perseguissem a mesma idéia. Embora se tenha estudado e publicado muito nos anos que se seguiram (1940-1950), estes serviram mais como base para desenvolvimento posterior.

Também proveniente deste período de tempo foi a construção do primeiro neuro-computador, denominado Snark, por Mavin Minsky, em 1951, mas este nunca executou qualquer função de processamento de informação interessante.

O primeiro neuro-computador a obter sucesso (Mark I Perceptron) surgiu em 1957 e 1958, criado por Frank Rosenblatt, Charles Wightman e outros. Devido à profundidade de seus estudos, suas contribuições técnicas e a sua maneira moderna de pensar, muitos o vêem como o fundador da neuro computação na forma em que a temos hoje. Seu interesse inicial para a criação do Perceptron era o reconhecimento de padrões.

Após Rosenblatt, Bernard Widrow, com a ajuda de alguns estudantes, desenvolveram um novo tipo de elemento de processamento de redes neurais chamado de Adaline, equipado com uma poderosa lei de aprendizado, que diferente do Perceptron ainda permanece em uso.

Infelizmente, os anos seguintes foram marcados por um entusiasmo exagerado de muitos pesquisadores, que passaram a publicar mais e mais artigos e livros que faziam uma previsão pouco confiável para a época, sobre máquinas tão poderosas quanto o cérebro humano que surgiriam em um curto espaço de tempo. Isto tirou quase toda a credibilidade dos estudos desta área e causou grandes aborrecimentos aos técnicos de outras áreas.

Um período de pesquisa silenciosa seguiu-se durante 1967 a 1982, quando poucas pesquisas foram publicadas porque os cálculos matemáticos das redes neurais são complicados, e na época não havia computadores avançados.

Nos anos 80, muitos dos pesquisadores foram bastante corajosos e passaram a publicar diversas propostas para a exploração de desenvolvimento de redes neurais bem como suas aplicações. Porém talvez o fato mais importante deste período tenha ocorrido quando Ira Skurnick, um administrador de programas da DARPA (Defense Advanced Research Projects Agency) decidiu ouvir os argumentos da neuro-computação e seus projetistas, e divergindo dos caminhos tradicionais dos conhecimentos convencionais, fundou em 1983 pesquisas em neuro computação. Este ato não só abriu as portas para a neuro computação, como também deu à DARPA o status de uma das líderes mundiais em se tratando de "moda" tecnológica.

Foi em 1986 que o campo de pesquisa na neuro computação "explodiu" com a publicação do livro "Parallel Distributed Processing" (Processamento Distribuído Paralelo) editado por David Rumelhart e James McClelland. Desde 1987, muitas universidades anunciaram a formação de institutos de pesquisa e programas de educação em neuro computação.

2.3-O Neurônio Artificial e a Rede Neural Artificial

Assim como o sistema nervoso é composto por bilhões de células nervosas, a rede neural artificial também seria formada por unidades que nada mais são que pequenos módulos que simulam o funcionamento de um neurônio. Estes módulos, ligam-se uns aos outros, e esta ligação é chamada de sinapse, devendo funcionar de acordo com os elementos em que foram inspirados, recebendo e retransmitindo informações.

2.4-Funcionamento do Neurônio Artificial

O fisiologista Warren MacCulloch interpretou o funcionamento do neurônio biológico como sendo um circuito de entradas binárias combinadas por uma soma ponderada (com pesos) produzindo uma entrada efetiva:

No modelo geral de neurônio como mostra a figura 2.1 as entradas $W_i X_i$ são combinadas usando uma função F , para produzir um estado de ativação do neurônio (correspondente à frequência de descarga do neurônio biológico). As entradas chegam através dos dendritos e têm um peso atribuído pela sinapse. A função básica de um neurônio é somar as entradas e retornar uma saída, caso esse valor seja maior que o valor de soma (threshold).

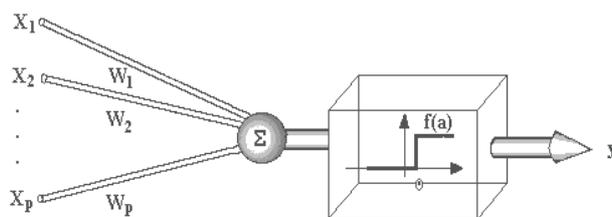


Figura 2.1-Modelo de McCulloch e Pitts

2.5-A Rede Neural Artificial (*Multilayer Perceptron*)

A rede neural artificial é um sistema de neurônios ligados por conexões sinápticas e dividido em neurônios de entrada, que recebem estímulos do meio externo, neurônios internos ou *hidden* (ocultos) e neurônios de saída. A forma de arranjar perceptrons em camadas é denominado Multilayer Perceptron (MLP). O *multilayer perceptron* foi concebido para resolver problemas mais complexos, os quais não poderiam ser resolvidos pelo modelo de neurônio básico. Um único perceptron ou uma combinação das saídas de alguns perceptrons poderia realizar determinadas operações, porém, seria incapaz de aprendê-las. Para isto são necessárias mais conexões, que só existem em uma rede de perceptrons dispostos em camadas. Os neurônios internos são de suma importância na rede neural pois provou-se que sem estes torna-se impossível a resolução de problemas linearmente não separáveis. Em outras palavras pode-se dizer que uma rede é composta por várias unidades de processamento, cujo funcionamento é bastante simples. Essas unidades, geralmente são conectadas por canais de comunicação que estão associados a determinado peso. As unidades fazem operações apenas sobre seus dados locais, que são entradas recebidas pelas suas conexões. O comportamento inteligente de uma Rede Neural Artificial vem das interações entre as unidades de processamento da rede.

A maioria dos modelos de redes neurais possui alguma regra de treinamento, onde os pesos de suas conexões são ajustados de acordo com os padrões apresentados. Em outras palavras, elas aprendem através de exemplos. Arquiteturas neurais são tipicamente organizadas em camadas, com unidades que podem estar conectadas às unidades da camada posterior.

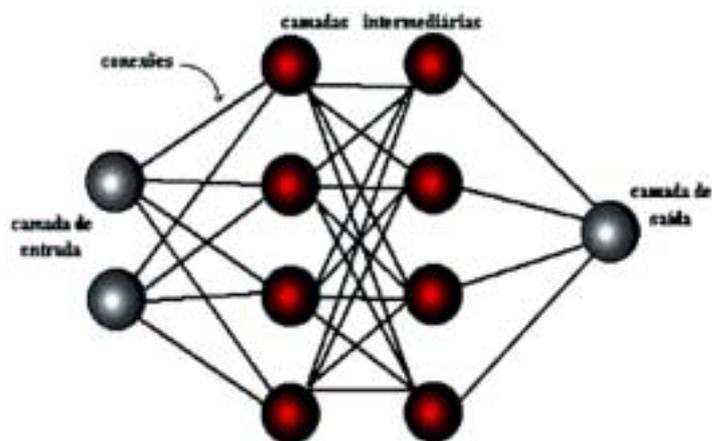


Figura 2.2-Organização em camadas .

A rede neural passa por um processo de treinamento a partir dos casos reais conhecidos, adquirindo, a partir daí, a sistemática necessária para executar adequadamente o processo desejado dos dados fornecidos. Sendo assim, a rede neural é capaz de extrair regras básicas a partir de dados reais, diferindo da computação programada, onde é necessário um conjunto de regras rígidas pré-fixadas e algoritmos.

Usualmente as camadas são classificadas em três grupos:

- Camada de Entrada: onde os padrões são apresentados à rede;
- Camadas Intermediárias ou Ocultas: onde é feita a maior parte do processamento, através das conexões ponderadas; podem ser consideradas como extratoras de características;
- Camada de Saída: onde o resultado final é concluído e apresentado.

Redes neurais são também classificadas de acordo com a arquitetura em que foram implementadas, topologia, características de seus nós, regras de treinamento, e tipos de modelos. Ao modo como se alteram os pesos durante o treinamento, denomina-se regra de treinamento e, dentre todas as regras existentes a mais utilizada é o backpropagation, que será tratado no próximo item.

2.6-Backpropagation

Durante o treinamento com o algoritmo backpropagation, a rede opera em uma sequência de dois passos, como mostra a figura 2.3. Primeiro, um padrão é apresentado à camada de entrada da rede. A atividade resultante flui através da rede, camada por camada, até que a resposta seja produzida pela camada de saída. No segundo passo, a saída obtida é comparada à saída desejada para esse padrão particular. Se estas não forem iguais, o erro é calculado. O erro é retropropagado da camada de saída até a camada de entrada. Os pesos das conexões das unidades da camada de saída e das camadas intermediárias vão sendo modificados conforme o erro é retropropagado.

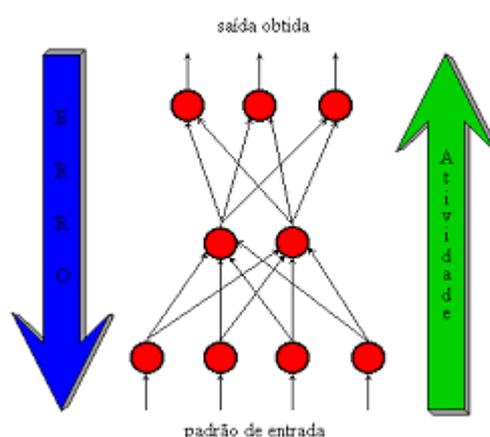


Figura 2.3-Esquema do funcionamento do Algoritmo Backpropagation

As redes que utilizam backpropagation trabalham com uma variação da regra delta, apropriada para redes com múltiplas camadas: a regra delta generalizada. A regra delta padrão essencialmente implementa um gradiente descendente no quadrado da soma do erro para funções de ativação lineares.

Entretanto, a superfície do erro pode não ser tão simples, como a ilustrada na figura 2.4, e suas derivadas mais difíceis de serem calculadas. Nestes casos, devem ser utilizadas redes com camadas intermediárias. Ainda assim, as redes ficam sujeitas aos problemas de procedimentos *hill-climbing*, ou seja, ao problema de mínimos locais. A figura também ilustra o processo de redução dos erros utilizando o backpropagation.

O erro cometido pela rede vai sendo progressivamente diminuído, podendo chegar ao mínimo global.

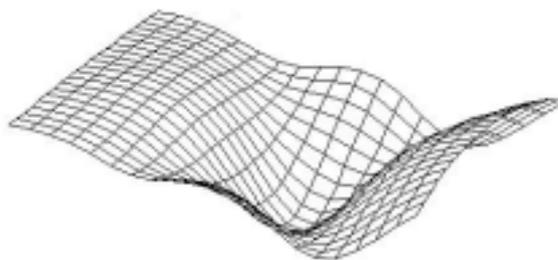


Figura 2.4- Exemplo de uma superfície de erro para um treinamento usando backpropagation

A regra delta generalizada funciona quando são utilizadas na rede unidades com uma função de ativação semi-linear, que é uma função diferenciável e não decrescente. Note que a função threshold não se enquadra nesse requisito. Uma função de ativação amplamente utilizada, nestes casos, é a função sigmoid.

O verdadeiro gradiente descendente requer que sejam tomados passos infinitesimais. Assim quanto maior for essa constante, maior será a mudança nos pesos, aumentando a velocidade do aprendizado, o que pode levar à uma oscilação do modelo na superfície de erro. O ideal seria utilizar a maior taxa de aprendizado possível que não levasse à uma oscilação, resultando em um aprendizado mais rápido.

O treinamento das redes MLP com backpropagation pode demandar muitos passos no conjunto de treinamento, resultando um tempo de treinamento consideravelmente longo. Se for encontrado um mínimo local, o erro para o conjunto de treinamento pára de diminuir e estaciona em um valor maior que o aceitável. Uma maneira de aumentar a taxa de aprendizado sem levar à oscilação é modificar a regra delta generalizada para incluir o termo momentum, uma constante que determina o efeito das mudanças passadas dos pesos na direção atual do movimento no espaço de pesos.

Desta forma, o termo momentum leva em consideração o efeito de mudanças anteriores de pesos na direção do movimento atual no espaço de pesos. O termo momentum torna-se útil em espaços de erro que contenham longas gargantas, com curvas acentuadas ou vales com descidas suaves, como o apresentado na figura 2.4.

As respostas geradas pelas unidades são calculadas através de uma função de ativação. Existem vários tipos de funções de ativação, as mais comuns são: Hard limiter, Threshold Logic e Sigmoid, como mostra a figura 2.5.

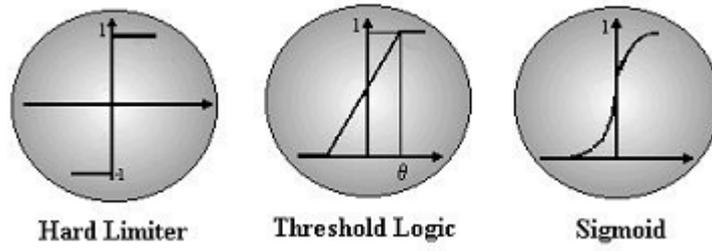


Figura 2.5-Tipos de funções de ativação mais utilizadas

Capítulo 3

Reconhecimento de Padrões Utilizando Redes Neurais

Notadamente as redes neurais têm sido uma técnica muito utilizada recentemente em tarefas de reconhecimento de padrões. Alguns autores renomados afirmam que entre todos os domínios de utilização de modelos de redes neurais, o reconhecimento de padrões é o que possui maior potencial. Nos próximos tópicos, procurou-se definir a tarefa de reconhecimento de padrões e seu relacionamento com redes neurais artificiais.

3.1 Reconhecimento de padrões

Reconhecimento de padrões é uma tarefa que causa pouca dificuldade para os seres humanos, e até mesmo para os animais, contudo são um grande desafio para a tecnologia moderna.

O termo “reconhecimento de padrões” foi introduzido no início da década de 60 e originalmente significava a detecção de formas simples. Para Bezdek e Pal, há várias definições para o termo reconhecimento de padrões, porém aquela que mais o caracteriza foi dada por Duda e Hart, em 1973, como sendo o “campo interessado no reconhecimento de regularidades significativas em ambientes ruidosos ou complexos, feito por máquinas”, ou a “procura por uma estrutura em dados.”

Há duas grandes motivações para estudos nesta área: a necessidade das pessoas em se comunicarem com máquinas computacionais através de linguagens naturais; e o interesse na idéia de projetar e construir autômatos (máquinas inteligentes) que possam realizar certas tarefas com habilidades comparáveis à performance humana. Estas tarefas envolvem percepção invariante em relação a equivalência de estímulos, posição, deslocamento, rotação, perspectiva, oclusão parcial, entre outros. Entre as áreas de aplicação do reconhecimento de padrões estão:

- Comunicação do homem com a máquina: reconhecimento automático da fala, reconhecimento da escrita, compreensão da fala, compreensão das imagens, processamento da linguagem natural;
- Defesa: reconhecimento, orientação e controle automáticos de alvos;
- Medicina: diagnose médica, análise de imagens, classificação de doenças;
- Veículos: controladores de automóveis, aviões, trens, barcos;
- Polícia e investigação: detecção criminal a partir da fala, escrita manual, impressões digitais, fotografias;
- Estudo e estimativa de recursos naturais: agricultura, extrativismo, geologia, ambiente;
- Sistemas domésticos: utensílios e
- Computadores: hardware e software difusos.

3.2-Técnicas para o reconhecimento de padrões

Algumas aplicações de reconhecimento podem requerer que seja encontrado uma ocorrência exata de um padrão, enquanto outras são satisfeitas por encontrar uma

ocorrência aproximada. Esta crítica distinção separa os métodos baseados em abordagens simbólicas dos métodos baseados em universo de características e teoria de decisão estatística. Alguns autores também denominam de abordagens sintáticas ou estruturais e de abordagens estatísticas ou numéricas.

Um deles explora a idéia da representação sintática (ou estrutural) de um padrão para o problema de reconhecimento, utilizando gramáticas para esta tarefa. Ao definir uma gramática para descrever os padrões, o projetista estará definindo componentes primitivos (símbolos) e as regras de como estes componentes serão agrupados para formar os padrões desejados.

Outra corrente de reconhecimento de padrões é a do reconhecimento numérico de padrões que define que todo padrão pode ser representado por um vetor numérico, conhecido como vetor característico ou vetor do padrão. Este vetor é comparado a um vetor representativo de uma classe de padrões, também conhecido como vetor de similaridade ou de proximidade. O padrão será reconhecido ou classificado de acordo com a maior similaridade entre ele e o vetor representativo de uma classe. Muitas das técnicas aplicadas ao problema de reconhecimento através de estatística, lógica difusa e redes neurais estão nesta corrente.

3.3-Relações entre redes neurais e reconhecimento de padrões

Técnicas e conceitos básicos a respeito de redes neurais artificiais têm sido amplamente cobertos pela literatura. Porém é importante destacar algumas vantagens do uso destas redes para que se possa melhor entender o seu relacionamento e sua aplicação em processos de reconhecimento de padrões.

O reconhecimento de padrões é, por sua própria natureza, uma ciência não exata. Enquanto alguns padrões podem ser identificados como bem estruturados ou estruturados adequadamente para serem definidos por uma gramática, como por exemplo a fabricação de um produto em uma linha de montagem, outros são de difícil modelagem ou difícil construção de uma gramática.

Fukushima, um autor da área de reconhecimento de padrões, ao desenvolver a idéia da representação sintática para o problema de reconhecimento, descreve o processo de definição de uma linguagem para a descrição dos padrões, afirmando que não há soluções gerais para o reconhecimento, e que a escolha é influenciada pela natureza dos dados disponíveis, da aplicação e da tecnologia. O projetista então vai realizar escolhas para a definição da gramática baseadas na sua experiência e vontade.

Esta natureza não estruturada do padrão a ser reconhecido torna o problema de reconhecimento difícil de ser tratado por paradigmas tradicionais de computação. Desta forma, o uso de redes neurais aparece como modo alternativo de resolução, pois segundo Iyengar e Kashyap, outros autores, ao invés de criar procedimentos lógicos, a construção destas redes envolve o entendimento informal do comportamento desejado para atender ao problema. Blum, outro importante autor da área, afirma que utilizando redes neurais há menor necessidade de se determinar, a princípio, quais são os fatores determinantes sobre o modelo que se está desenvolvendo.

As principais vantagens do uso de redes neurais sobre muitas técnicas tradicionais de reconhecimento de padrões são:

- Adaptatividade: habilidade de se ajustar a novas informações;
- Velocidade: via o paralelismo massivo;
- Tolerância a falhas: capacidade de oferecer boas respostas mesmo com falta, confusão ou dados ruidosos;

- Otimalidade: visto como taxa de erros em sistemas de classificação.

A capacidade de generalização de uma rede neural também se apresenta como vantagem no problema de reconhecimento de padrões. Tomando como exemplo o reconhecimento de escrita manual, pode-se imaginar a grande variedade de formas que uma mesma letra pode ser escrita, mesmo sendo por uma mesma pessoa. A inclinação, tamanho, pressão e traçado são algumas das variáveis que podem afetar o reconhecimento da escrita. A rede neural após aprender a distinguir alguns “As” de tamanhos diferentes de alguns “Bs” de tamanhos diferentes, será capaz de distinguir um A de qualquer tamanho de um B de qualquer tamanho. Desta forma, a capacidade de reconhecer padrões nunca antes vistos, porém semelhantes aos apresentados durante o treinamento, torna-se um importante diferencial perante muitas técnicas tradicionais.

De um modo geral, as redes neurais são um método de modelagem altamente recomendável para se lidar com sistemas abertos ou mais complexos, pouco entendidos e que não podem ser adequadamente descritos por um conjunto de regras ou equações.

Conclusão

Nesta nota pudemos observar que as redes neurais são uma técnica que auxiliam no reconhecimento de padrões devido às suas características e vantagens frente à natureza não estruturada dos padrões, que são um problema bastante estudado e abordado de diversas formas atualmente. E às técnicas de processamento de imagens, podem-se unir as redes neurais para se fazer o reconhecimento de padrões.

Bibliografia

- “Digital Image Processing”- R. C. Gonzales / R. Woods-Addison Wesley-1992.
- “Fundamentals of Digital Image Processing”. A. K. Jain – Prentice Hall – 1989
- <http://sputnik.dpi.inpe.br/spring/teoria/realce/realce.htm>
- <http://www.lcmi.ufsc.br/~ferdes/rap/DocumentoOnLine/>
- http://www.datacom.com.br/tecnologia/ocr_icr.htm
- http://www.eps.ufsc.br/disserta/hugo/cap_3/cp3_hug.htm
- <http://www.visgraf.impa.br/sibgrapi96/trabs/pdf/c12.pdf>