Aplicação de Redes Neurais Artificiais em Conjunto com o Método Vetorial da Propagação de Feixes na Análise de um Acoplador Direcional Baseado em Fibra Ótica

Pablo Luiz Braga Soares¹ José Patrocínio da Silva¹

Resumo: Redes Neurais Artificiais (RNAs), usualmente denominadas redes neurais, são modelos matemáticos que se assemelham às estruturas neurais biológicas. O Método dos Elementos Finitos (MEF) e o Método da propagação de feixes (MPF) são utilizados na discretização de domínios originais contínuos de problemas físicos. O objetivo deste artigo é apresentar a modelagem e projeto de um acoplador direcional formado por fibras óticas separadas por uma distância d, constantes ao longo da direção de propagação, aplicando RNAs. Os dados necessários para treinamento da RNA foram obtidos a partir da análise modal e da aplicação do MVPF como ferramenta de simulação numérica. A rede neural obteve como saída a potência nas duas fibras ao longo da distância de acoplamento.

Palavras-chave: Redes neurais artificiais. Acopladores direcionais de fibra ótica. Método vetorial da propagação de feixes.

Abstract: Artificial Neural Networks (ANN), commonly called neural networks are mathematical models that simulate biological neural structures. The Finite Element Method (FEM) and beam propagation method (BPM) are used in the original discretization of continuous domains of physical problems. The aim of this paper is to present the modeling and design of a directional coupler formed by optical fibers separated by a distance d, constant along the propagation direction, using RNAs. The data needed for training the ANN were obtained from the modal analysis and the application of VBPM as a tool for numerical simulation. The neural network achieved an output power in both the long distance fiber coupling.

Keywords: Artificial neural networks. Directional couplers fibers optics. Vector beam propagation method.

1 Introdução

O primeiro modelo artificial de um neurônio biológico data de 1943, resultado do trabalho do neuroanatomista e psiquiatra Warren McCulloch e do matemático Walter Pitts. Entretanto, mereceram ênfase naquilo que possibilite o conhecimento básico de seus fundamentos, bem como a descrição dos paradigmas de aprendizagem, evidenciando-se o modelo Perceptron Múltiplas Camadas (Multi Layer Perceptron - MLP) e o algoritmo backpropagation, responsável pelo ressurgimento das RNA's no final da década 1980. Muitos autores conceituam e descrevem redes neurais artificiais, sempre enfatizando as características que mais evidenciam a utilização que se dá a elas. Várias dessas definições podem ser encontradas em [1] e [2].

Acopladores direcionais formados por fibra ótica (Figura 1) são estruturas básicas importantes no projeto de dispositivos fotônicos e circuitos, como filtros, dispositivos multiplexadores e demultiplexadores. Em alguns casos e com o uso de algumas aproximações, existem soluções analíticas aproximadas [3] e [4]. Em geral, métodos numéricos, como o Método dos Elementos Finitos (Finity Element Method - FEM), são utilizados na modelagem dessas estruturas. Entretanto, isso requer conhecimento da teoria eletromagnética e grandes recursos

http://dx.doi.org/10.5335/rbca.2011.1803

¹ Mestrado em Ciência da Computação, UERN/UFERSA, Costa e Silva - BR 110 - Km 47 - Mossoró (RN) - Brasil {pabloufersa@gmail.com}, {patroc@ufersa.edu.br}

computacionais. Por outro lado, redes neurais artificiais (RNAs) têm a capacidade de aproximar funções não lineares, podendo, através de treinamento, mapear relações de entrada e saída. Essa habilidade permite, sem a necessidade de conhecer o objeto investigado, modelar um sistema conhecendo apenas os valores de entrada e saída. Tal funcionalidade é usada na modelagem de acopladores de fibras óticas [5] e também para análise e síntese de antenas de microfita na área de microondas [6] e [7]. A maior vantagem do uso de modelos baseados em redes neurais é a sua simplicidade, o pouco tempo e esforço computacional requeridos após a fase de treinamento, além da sua aplicação em problemas de síntese.

Na modelagem de dispositivos fotônicos, no que se refere à obtenção da formulação do VBPM, a eficiência e robustez na construção do modelo proposto representam o maior desafio a ser vencido. Nesse contexto, o MEF tem sido amplamente utilizado em frequências óticas. Esse tipo de aplicação despertou um grande interesse no uso do método devido a sua flexibilidade na discretização de estruturas com geometrias complexas. Outras características que tornam o uso do método bastante atraente são a possibilidade de análise de materiais anisotrópicos e a possibilidade do uso de condições de contorno absorventes eficientes [8], o que permite a redução do domínio e esforco computacional sem comprometer a exatidão dos resultados.

Na área de fotônica, existe um grande interesse na análise de novos e diversos dispositivos óticos integrados com características especiais, como variações tanto no espaço quanto no tempo. Essas aplicações têm sido alvo da atenção de diferentes autores, por exemplo: divisores de potência [9], [10], [11], [12], [13] e [14]; guias de variação longitudinal [13], [15], [16], [17] e [18]; filtros óticos; multiplexadores; demultiplexadores; roteadores; ressoadores óticos [14] e [11]; grades de Bragg [17]; chaveadores [18], [9] e [13]; estruturas fotônicas periódicas magneto-ótico [19], [20] e [21] etc.

Mais recentemente, a atenção de muitos pesquisadores tem se concentrado no uso da inteligência computacional para aplicações em controladores, nas fibras de cristais fotônicos (Photonics Crystals fibers, PCF's), nos dispositivos construídos a partir desses tipos de fibras, tais como os acopladores de fibras de cristais fotônicos. Portanto, devido à importância das aplicações, o trabalho aqui proposto, em termos globais, representa a etapa de desenvolvimento do simulador que faz a análise da propagação.

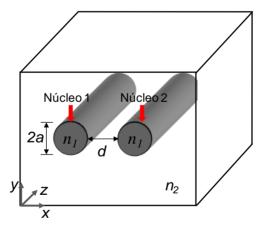


Figura 1. Acoplador direcional formado por 2 fibras óticas idênticas

O restante deste artigo está organizado da seguinte forma: na seção 2 é feita uma revisão sobre o funcionamento de uma rede neural artificial enfatizando a topologia, o paradigma de aprendizagem e algoritmo utilizado no artigo e, ao final da seção, é apresentada a MLP; na seção 3 é apresentado o funcionamento do método VBPM e onde se aplica; na seção 4 são feitas algumas considerações sobre acopladores direcionais com base em fibra ótica; na seção 5 é descrito o ambiente de simulação utilizado; na seção 6, resultados obtidos e na Seção 7 serão apresentadas as considerações finais.

2 Redes Neurais Artificiais

Segundo [2], em "sistemas inteligentes", as RNAs são modelos matemáticos que se assemelham às estruturas neurais biológicas (neurônios) e que têm capacidade computacional adquirida por meio de

aprendizado e generalização. O aprendizado de uma rede neural é um processo pelo qual os parâmetros livres da mesma são adaptados pelo processo de estimulação no ambiente onde a rede está inserida [22]. Esta etapa pode ser considerada como uma adaptação da RNA às características intrínsecas de um problema, onde se procura cobrir um grande espectro de valores associados às variáveis pertinentes. Isso é feito para que a RNA adquira, através de uma melhora gradativa, uma boa capacidade de resposta para o maior número de situações possíveis. Por sua vez, a generalização de uma RNA está associada à sua capacidade de dar respostas coerentes para dados não apresentados a ela durante o treinamento. Espera-se que uma RNA treinada tenha uma boa capacidade de generalização, independentemente de ter sido utilizado o aprendizado supervisionado ou não supervisionado durante o treinamento.

As RNAs diferenciam-se pela sua arquitetura e pela forma como os pesos associados às conexões são ajustados durante o processo de aprendizado. A arquitetura de uma rede neural restringe o tipo de problema no qual a rede poderá ser utilizada e é definida pelo número de camadas (camada única ou múltiplas camadas), pelo número de nós em cada camada, pelo tipo de conexão entre os nós (feedforward ou feedback) e por sua topologia [22]. O modelo inicial para um neurônio artificial foi proposto por [23] e posteriormente aprimorado por [24], que o nomeou de perceptron (Figura 2).

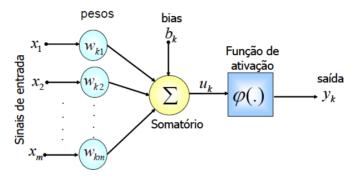


Figura 2. Modelo não linear de um neurônio artificial (Adaptado de [22])

Nele x_m são as entradas da rede; w_{km} são os pesos, ou pesos sinápticos, associados a cada entrada; b_k é o termo bias; u_k é a combinação linear dos sinais de entrada; $\varphi(.)$ é a função de ativação e y_k é a saída do neurônio. Pode-se dizer que é nos pesos que reside todo o conhecimento adquirido pela rede. Os pesos é que são os parâmetros ajustáveis que mudam e se adaptam à medida que o conjunto de treinamento é apresentado à rede. Assim, o processo de aprendizado supervisionado em uma RNA com pesos, resulta em sucessivos ajustes dos pesos sinápticos, de tal forma que a saída da rede seja a mais próxima possível da resposta desejada. Tipicamente, a ordem de amplitude normalizada da saída do neurônio está no intervalo [0, 1] ou alternativamente [-1, 1]. O modelo neural também inclui um termo chamado de "bias", aplicado externamente, simbolizado por b_k . O b_k tem o efeito do acréscimo ou decréscimo da função de ativação na entrada da rede, dependendo de se é positiva ou negativa, respectivamente. O bias serve para aumentar os graus de liberdade, permitindo uma melhor adaptação, por parte da rede neural, ao conhecimento fornecido a ela.

2.1 Topologias das Redes Neurais

A priori, quanto mais camadas de neurônios, melhor seria o desempenho da rede neural, pois aumenta a capacidade de aprendizado, melhorando a precisão com que delimita regiões de decisão. Mas, na prática, aumentar o número de camadas intermediárias às vezes se torna inviável devido ao trabalho computacional. Pesquisadores como [25] afirmam que com apenas uma camada intermediária na rede neural já é possível calcular uma função arbitrária qualquer a partir de dados fornecidos. De acordo com esses autores, a camada oculta deve ter por volta de (2i+1) neurônios, onde i é o número de variáveis de entrada. Outros, como [26], estudaram o número de camadas intermediárias necessárias à implementação de funções em RNAs. Seus resultados indicam que uma camada intermediária é suficiente para aproximar qualquer função contínua e que duas camadas intermediárias aproximam qualquer função matemática. [27] afirma que, havendo uma segunda camada intermediária na rede neural, esta deve ter o dobro de neurônios da camada de saída; no caso de apenas uma camada oculta, deverá ter s(i+1) neurônios, onde s é o número de neurônios de saída e i, o número de neurônios na entrada.

2.2 Paradigmas de Aprendizagem

A propriedade que é de importância primordial para uma rede neural é a sua habilidade de aprender a partir de seu ambiente e de melhorar o seu desempenho através da aprendizagem. Uma rede neural aprende acerca do seu ambiente através de um processo interativo de ajustes aplicados a seus pesos sinápticos e níveis de bias.

[22] define aprendizagem como um processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um processo de estimulação pelo ambiente no qual a rede está inserida. Apesar de existirem diversos métodos (regras como p.ex., correção do erro, aprendizagem hebbiana, aprendizagem competitiva e aprendizagem boltzmann) para o treinamento de redes neurais, podem ser agrupados em dois paradigmas principais: os aprendizados supervisionado e não supervisionado. Existem ainda os aprendizados por reforço e por competição, mas o método de aprendizado supervisionado é o mais usado em RNAs, quer de neurônios com pesos, quer de neurônio sem pesos, e possui este nome em função de que tanto entradas quanto saídas desejadas são fornecidas à rede durante o processo de treinamento [2].

2.2.1 Aprendizado Supervisionado

Aprendizagem supervisionada, que é também denominada "aprendizagem com um professor" [22], é um método desenvolvido pela disponibilização de um professor que verifique os desvios da rede a um determinado conjunto de dados de entrada.

Em termos conceituais, o professor é considerado como tendo conhecimento sobre o ambiente, sendo representado por um conjunto de exemplos de entrada-saída, [22], como ilustrado na Figura 3. Entretanto, o ambiente é desconhecido pela rede neural de interesse.

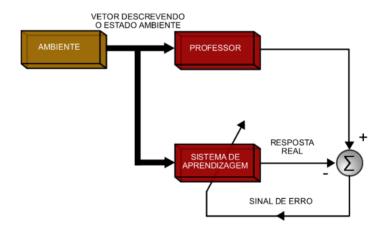


Figura 3. Diagrama em blocos de aprendizagem com o professor (Adaptado de [22])

Quando um vetor de treinamento é apresentado a uma rede neural que está usando o paradigma da aprendizagem supervisionada, em virtude do seu conhecimento prévio, o professor é capaz de fornecer à rede neural uma resposta desejada àquele vetor de treinamento. Na verdade, a resposta desejada representa a ação (saída) ótima a ser realizada pela rede neural. Os parâmetros (pesos sinápticos) da rede neural são, então, ajustados através da influência combinada do vetor de treinamento e do sinal de erro indicado pelo professor. O sinal de erro é definido como a diferença entre a resposta desejada e a resposta real da rede. Este ajuste é realizado passo a passo, iterativamente, com o objetivo de fazer a rede neural emular o professor. Dessa forma, o conhecimento disponível ao professor do ambiente é passado para a rede através do treinamento; ao alcançar esse objetivo, podemos, então, dispensar o professor e deixar a rede neural lidar sozinha com o ambiente.

Como medida de desempenho para o sistema, pode-se pensar em termos do erro médio quadrado ou da soma de erros quadrados sobre a amostra de treinamento. O sinal de erro (a diferença entre a resposta desejada, d, e a saída da rede, z) é dado por:

$$e(n) = d(n) - z(n) \tag{1}$$

em que o índice n varia no intervalo n = 1, 2, ..., N, sendo N o número total de exemplos de treinamento. Para evitar que valores de erro com sinais positivos anulem valores com sinais negativos, toma-se o erro quadrático de cada amostra, levando em consideração todas as saídas da rede, Ns:

$$E(n) = \frac{1}{2N_s} \sum_{k=1}^{N_s} [e(n)]^2$$
 (2)

para uma análise geral do treinamento utiliza-se a média dos erros quadráticos das amostras de todo o conjunto de treinamento. Esta medida, denominada MSE, é dada por:

$$E(t) = \frac{1}{N} \sum_{n=1}^{N} E(n)$$
(3)

em que, *t* denota o número de épocas de treinamento. Uma época é contada a cada apresentação à rede neural de todos os exemplos do conjunto de treinamento.

2.2.2 Aprendizado Não-Supervisionado

Segundo [22], na aprendizagem não supervisionada, ou auto-organizada, não há um professor externo para supervisionar o processo de aprendizado, como ilustrado na Figura 4. Em vez disso, são dadas condições para realizar uma medida independentemente da tarefa da qualidade da representação que a rede deve aprender, e os parâmetros livres da rede são otimizados em relação a esta medida.

O ESTADO DO AMBIENTE SISTEMA DE APRENDIZAGEM

VETOR DESCREVENDO

Figura 4. Diagrama em blocos da aprendizagem não-supervisionada (Adaptado de [22])

2.3 Algoritmo da Retropropagação (Backporpagation)

O tipo mais utilizado de rede neural é a MLP treinada com o algoritmo backpropagation (BP) ou retropropagação do erro, [28]. A difusão deste algoritmo reporta a década de 1980, representando um marco na utilização das redes neurais e tendo o trabalho de [29] como uma das referências principais.

Para treinar a rede são utilizados vetores de entrada associados aos seus respectivos vetores de saída desejada, até que a rede aproxime uma determinada função e possa, a partir daí, oferecer saídas adequadas a vetores de entrada diferentes daqueles com os quais foi treinada. O backpropagation padrão é um algoritmo gradiente descendente, por meio do qual os pesos das conexões entre os neurônios são atualizados ao longo de um gradiente descendente de uma determinada função. O termo backpropagation refere-se à forma como o gradiente é calculado para redes de múltiplas camadas não lineares.

Existe um número razoável de variações a partir do algoritmo básico, que são baseadas em técnicas de otimização, tais como o gradiente conjugado e métodos de Newton. Basicamente, a aprendizagem por retropropagação de erro consiste em dois passos através das diferentes camadas da rede: um passo para frente, Feed-forward (a propagação), e um passo para trás, *Feed-backward* (retropropagação) [22], como ilustrado na Figura 5.

Feed-forward → um padrão de atividade (vetor de entrada) se propaga pela rede, da camada de entrada até a camada de saída, e, finalmente, um conjunto de saídas é produzido como resposta real da rede. Durante esse passo os pesos sinápticos da rede são todos fixos;

Feed-backward → a resposta real é subtraída de uma resposta desejada (alvo) para produzir um sinal de erro. O erro se propaga na direção contrária ao fluxo de dados, indo da camada de saída até a primeira camada escondida, ajustando os pesos sinápticos das camadas.

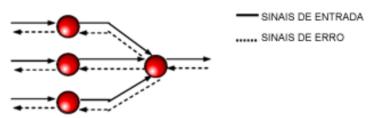


Figura 5. Representação Feed-Forward(Entrada) e Feed-backward(Erro) (Adaptado de [22])

O Quadro 1 apresenta sucintamente este algoritmo. Maiores detalhes acerca da regra de atualização de pesos e uma descrição matemática podem ser encontrados em [22]. Foi demonstrado por [26] que o backpropagation é um aproximador universal capaz de representar qualquer função.

Quadro 1. Descrição resumida do algoritmo de aprendizagem backpropagation

Inicialização
→ pesos iniciados com valores aleatórios e pequenos;
Treinamento
\rightarrow Loop até que o erro de cada neurônio de saída seja \leq tolerância, para todos os padrões do conjunto de
treinamento;
1. Aplica-se um padrão de entrada X_i com seu respectivo vetor de saída Y_i desejado;
2. Calculam-se as saídas dos neurônios, começando da primeira camada escondida até a camada de saída;
3. Calcula-se o erro para cada neurônio da camada de saída. Se erro ≤ tolerância, para todos os neurônios,

- 4. Atualizam-se os pesos de cada neurônio, começando pela camada de saída, até a primeira camada escondida;
- 5. Volta-se ao passo 1 caso a diferença entre a saída da rede e a resposta desejada seja maior que um determinado limite especificado pelo usuário

Um dos fatores que influenciam no aprendizado é a forma como se dá. O algoritmo básico pode ser desenvolvido tanto para o aprendizado incremental quanto para o aprendizado acumulativo. No modo de aprendizado incremental, os pesos são atualizados a cada apresentação de um novo padrão, mas tende a aprender melhor o último padrão apresentado.

No modo de aprendizado acumulativo, também chamado de modo batch, os pesos são ajustados apenas depois da apresentação de todos os padrões, ou seja, depois de um ciclo completo (após cada época) na apresentação dos padrões. Muitas das variações do algoritmo básico utilizam essa forma de aprendizado como tentativa de acelerar o processo de treinamento da rede.

2.4 O Perceptron Múltiplas Camadas (Multi Layer Perceptron)

volta ao passo 1;

Existem dezenas de diferentes modelos de RNA descritos na literatura, tais como MLP, Redes de função de Base Radial (RBF), Redes de Função Sample (SFNN), Redes de Fourier e Redes Wavelet. À primeira vista, o

processo de seleção pode parecer uma tarefa dificil diante do número de possibilidades, porém é provável que poucos modelos forneçam uma solução excelente [30].

Um perceptron de [23] calcula a combinação linear de entradas chamadas de "entradas da rede". Então, uma função de ativação possivelmente não linear é aplicada às entradas da rede para produzir uma saída. Um perceptron pode ter uma ou mais saídas. Segundo [31] ao introduzir na rede uma camada extra de neurônios, chamada de "intermediária" ou "camada oculta", e se o modelo inclui pesos, estimados entre a camada de entrada e a camada oculta, e a camada oculta usa função de ativação não linear, tal como a função logística, o modelo torna-se genuinamente não linear. O modelo resultante é chamado de Multi Layer Perceptron ou MLP, como ilustrado na Figura 6. De acordo com [22], a MLP é uma extensão do perceptron simples, capaz de trabalhar com problemas não linearmente separáveis. Esse avanço foi possível pela utilização de, pelo menos, uma camada entre a camada de entrada e a camada de saída. Essas camadas intermediárias, conhecidas como "camadas ocultas", trabalham como um reconhecedor de características, que ficam armazenadas nos pesos sinápticos.

Os perceptrons de múltiplas camadas têm sido aplicados com sucesso para resolver diversos problemas dificeis, através do seu treinamento de forma supervisionada com um algoritmo, conhecido como "algoritmo de retropropagação de erro" (error backpropagation) [22]. Um perceptron de múltiplas camadas tem três características distintivas:

- 1. o modelo de cada neurônio da rede inclui uma função de ativação não linear;
- 2. a rede contém uma ou mais camadas de neurônios ocultos, que não fazem parte da entrada ou da saída da rede;
- 3. a rede exibe um alto grau de conectividade, determinado pelas sinapses da rede.

É pela combinação dessas características, juntamente com a habilidade de aprender da experiência do treinamento, que a MLP deriva seu poder computacional. Essas mesmas características, entretanto, são também responsáveis pelas deficiências no estado atual de nosso conhecimento sobre o comportamento da rede. Primeiro, a presença de uma forma distribuída de não linearidade e a alta conectividade da rede tornam difícil a análise teórica de um perceptron de múltiplas camadas. Segundo, a utilização de neurônios ocultos torna o processo de aprendizagem mais difícil de ser visualizado.

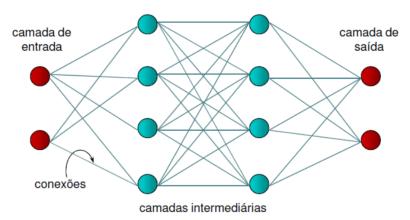


Figura 6. Representação de uma MLP

3 O Método Vetorial da Propagação de Feixes

Atualmente, o Método dos Elementos Finitos (MEF) tem sido utilizado em inúmeras áreas da engenharia e ciências da terra, que usam modelos de natureza caracterizada por equações diferenciais parciais. De fato, por se tratar de uma importante ferramenta computacional numérica e científica, o MEF têm tido um grande impacto na teoria e prática de métodos numéricos desde meados do século XX, inclusive na ótica e nas telecomunicações.

Infelizmente, dizer exatamente quando se deu início a este método é um tanto complexo, pois recai no problema básico de definir exatamente o que constitui o "Método dos Elementos Finitos". Para a maioria dos matemáticos, sua origem está associada ao apêndice de um artigo de [32], onde são discutidas aproximações lineares contínuas por parte do problema de Dirichlet em um domínio usando triângulos.

O Método dos Elementos Finitos é um método numérico utilizado na resolução de equações diferenciais parciais em domínios finitos. Através de uma discretização de domínios originais contínuos de problemas físicos, obtêm-se sistemas de equações lineares, que, em geral, são esparsos e podem ser resolvidos utilizando técnicas eficientes. A esparsidade das matrizes geradas é uma das principais vantagens do MEF, porque o sistema a ser resolvido é esparso e, portanto, de fácil resolução. Outra vantagem é sua versatilidade para a análise de estruturas com geometrias complexas. Em outras palavras, o MEF pode ser utilizado em domínios com características não homogêneas de formatos quaisquer. Assim, por exemplo, em regiões onde existe muita variação e/ou alta intensidade do campo eletromagnético usam-se malhas mais refinadas do que nas outras regiões onde a variação do campo e/ou a intensidade é menor e na situação é, em geral, difícil ou às vezes impossível de ser tratada eficientemente com outros métodos, como, por exemplo, o Método das Diferenças Finitas.

Em linhas gerais, o procedimento padrão para a solução de um problema qualquer utilizando o MEF pode ser resumido a um esquema de três passos básicos, conforme o fluxograma ilustrado na Figura 7.

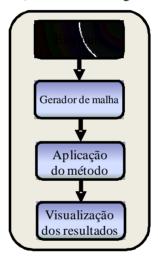


Figura 7. Fluxograma de um problema de elementos finitos

Primeiro, deve-se definir a malha dos elementos finitos para a estrutura que se deseja estudar, i.e., gerar as coordenadas dos nós da estrutura; indicar a conectividade entre os nós de cada elemento; entrar com as propriedades dos materiais presentes em cada elemento; por último, aplicar as condições de contorno apropriadas nos nós apropriados. Existem diversos programas, alguns comerciais e outros gratuitos, eficazes e eficientes. Como exemplos Gid, GMSH, Amtec etc.

Em seguida, aplica-se o código do MEF apropriado, podendo ser um problema de análise modal ou algum tipo de propagador, que consiste em: ler todas as informações da malha; montar as chamadas "matrizes elementares", que fazem a conversão dos operadores diferenciais do modelo em uma série de contribuições individuais, ou seja, atribui-se uma contribuição para cada elemento; somar todas as contribuições das matrizes elementares para montar as chamadas "matrizes globais" do sistema matricial; e finalmente, resolver o problema em questão.

Uma vez resolvido o sistema, o último passo consiste em retomar os resultados numéricos e resgatá-los fisicamente através da visualização das incógnitas, ou manipulação destas para se conseguir alguns outros parâmetros inerentes e particulares a cada tipo específico de problema.

Em linhas gerais, para se proceder com método elementos finitos em aplicações eletromagnéticas, partese das equações de Maxwell e obtém-se a equação de onda de Helmholtz no domínio da frequência,

$$\nabla \times \left(\bar{k}\nabla \times \vec{H}\right) - k_0^2 \vec{H} = 0 \tag{4}$$

onde k_0 é o número de onda no espaço livre, dado por $\omega \sqrt{\mu_0 \epsilon_0}$, sendo ω a frequência angular, $\bar{k} = 1/\bar{\epsilon}$, com $\bar{\epsilon}$ representando o tensor permissividade relativa do meio.

4 Acopladores Direcionais

Acopladores direcionais são componentes que, como o próprio nome sugere, acoplam um sinal, no caso eletromagnético, que se propaga por um canal (cabo coaxial, fibra ótica, guias de ondas retangulares ou circulares etc.) a outro canal do mesmo tipo. Esses componentes encontram grande aplicação na faixa de microondas (telecomunicações, radares, etc.) [33] e [34].

Acoplamento direcional é um fenômeno de troca de energia entre ondas eletromagnéticas que se propagam em diferentes estruturas de guiagem [35]. O dispositivo que resulta da disposição dessas estruturas de guiagem com o objetivo de controlar a troca de energia é denominado "acoplador direcional". Seja para a aquisição de amostras, seja para a divisão de sinais, o acoplador é um dispositivo fundamental em qualquer circuito ótico.

Acopladores direcionais podem ser construídos a partir de guias de onda e microlinhas para operar na faixa de microondas e a partir de fibras óticas e em ótica integrada para a faixa ótica. Qualquer que seja a faixa de operação, os modelos dessas estruturas são obtidos a partir das equações de Maxwell e da geometria e composição do material das estruturas de guiagem envolvidas, sendo o dispositivo descrito, em qualquer caso, por meio de matriz espalhamento e dos parâmetros diretividade, acoplamento e isolação [36], [37] e [38].

A Figura 8 ilustra como se comporta o acoplamento ótico para um acoplador de fibras óticas. A luz é lançada na fibra 1 e, após percorrer a distância L, parte de sua energia é transferida para a fibra2. O acoplamento poderá ocorrer parcialmente, quando apenas uma parte da energia é transferida, ou totalmente, quando toda energia presente na fibra 1 é transferida para a fibra 2. Este tipo dispositivo é bastante utilizado em divisores de potência, polarizadores, multiplexadores e outros dispositivos óticos para os quais se deseja compartilhar a energia.

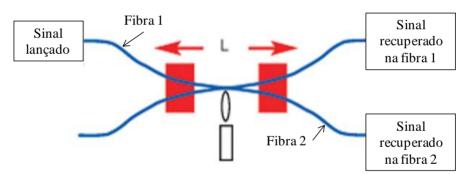


Figura 8. Acoplamento de um acoplador de fibras óticas

5 Simulação

A simulação foi desenvolvida utilizando o software interativo Matlab, que, segundo [39], integra a capacidade de fazer cálculos, visualização gráfica e programação em um ambiente fácil de usar, além de possuir uma variedade de funções desenvolvidas para uma gama de problemas.

A estrutura (Figura 1) foi excitada com um feixe correspondente ao modo E_{11}^x , lançado na fibra da esquerda, como ilustrado na Figura 9. Este feixe possui índice de refração efetivo neff ($\beta z/k_0$) = 1,4529, calculado isoladamente para a fibra da esquerda usando-se a análise modal [8]. Como os núcleos dos acopladores estavam próximos, lado a lado, a uma distância d=4,0 µm, suficiente para que a energia do núcleo excitado se transfira para o núcleo adjacente, como ilustrado nas Figuras 10 e 11, o comprimento de onda usado nesta simulação foi centrado em 1,55 µm; o passo de propagação considerado no VBPM foi de 0,1.

Para esta análise utilizaram-se as equações de Sellmeier para obter os índices de refração da casca (n_2) e do núcleo (n_1) , [8]. O índice de refração da casca é constituído de sílica pura e o índice de refração do núcleo é composto por sílica dopada com germânio.

Na simulação utilizou-se o paradigma da aprendizagem supervisionada [22], onde o conjunto de entrada e saída (potência na fibra) foi obtido após a utilização do VBPM. O acoplador ficou modelado com 6000 elementos finitos. Para efeito de treinamento da rede neural utilizaram-se como entrada cinco características dos elementos finitos: a primeira é o número do elemento finito que varia entre 1-6000; a segunda, terceira e quarta são os três vértices que formam o triângulo do elemento finito, respectivamente; por fim, o índice de refração, que pode ser n_1 ou n_2 , dependendo de onde o elemento finito se encontra, no núcleo ou na casca, respectivamente. Como saída desejada utilizou-se a potência das duas fibras.

A topologia utilizada na rede foi do tipo de alimentação direta, mais precisamente do tipo MLP, onde três camadas foram suficientes para mapear os valores de entrada na saída esperada. A camada de entrada foi modelada com cinco neurônios, que representam as características escolhidas dos elementos finitos; a função de ativação utilizada na camada oculta foi do tipo sigmoide com 11 neurônios, pois, de acordo com [25], a camada oculta deve ter por volta de (2i+1) neurônios, onde i é o número de variáveis de entrada; na camada de saída usou-se a função linear com dois neurônios, que representam a potência gerada pela rede neural na fibra da esquerda e direita, respectivamente.

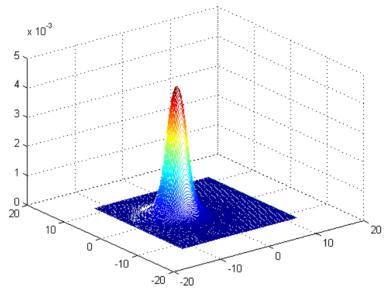


Figura 9. Feixe lançado na fibra da esquerda em $z = 0 \mu m$

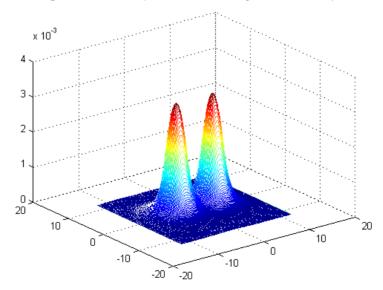


Figura 10. Transferência de energia entre as fibras da esquerda para direita para $z = 1500 \mu m$ (transferência parcial).

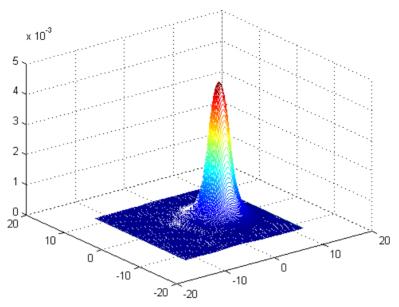


Figura 11. Transferência de energia entre as fibras da esquerda para direita para $z = 3000 \mu m$ (transferência total).

6 Resultados

A Figura 27 mostra a evolução da componente hx, do campo magnético transversal, na direção de propagação.

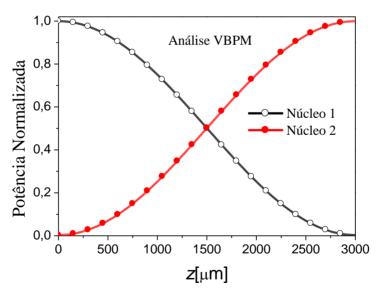


Figura 12. Evolução da potência normalizada das componentes transversais do campo ao longo da direção de propagação obtida através do VBPM.

Com a rede neural já pronta, foi realizada uma simulação antes de ser treinada. A rede apresentou os resultados ilustrados na Figura 13.

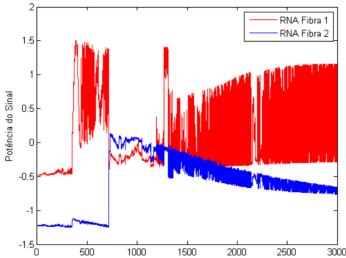


Figura 13. Evolução da potência normalizada das componentes transversais do campo ao longo da direção de propagação obtida antes do treinamento da RNA.

Para efeito de treinamento utilizou-se o algoritmo de backpropagation descrito na Seção 2.3, e os critérios utilizados para terminar o treinamento da rede foram que o erro médio quadrático, equação (3) atingisse valor $< 10^{-10}$ ou que a quantidade de 1000 épocas se esgotasse.

A Figura 14 mostra o erro médio quadrático sendo calculado no decorrer do treinamento da rede neural após cada época.

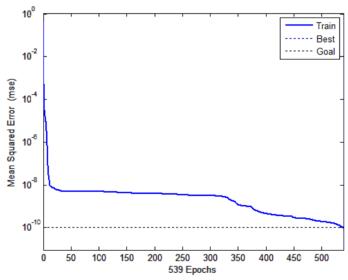


Figura 14. MSE em função do número de épocas

A rede terminou seu treinamento após as 539 épocas com o MSE de 9.7×10^{-11} . A Figura 15 mostra a evolução da componente hx, do campo magnético transversal, na direção de propagação obtida pela rede neural.

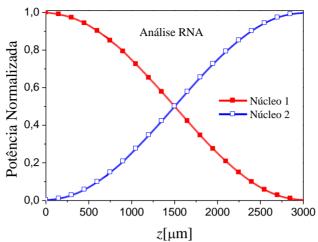


Figura 15. Evolução da potência normalizada das componentes transversais do campo ao longo da direção de propagação obtida após o treinamento da RNA.

Após o treinamento da rede neural, a mesma apresentou resultados bem próximos aos encontrados com o VBPM, pois, analisando o gráfico da Figura 15 com o gráfico da Figura 12, os resultados apresentam curvas semelhantes.

7 Considerações Finais

Neste trabalho, redes neurais modulares de alimentação direta por meio do paradigma da aprendizagem supervisionada foram aplicadas em conjunto com a teoria de acopladores e o método da propagação vetorial de feixes. O algoritmo responsável pelo treinamento foi o algoritmo backpropagation, que atende às necessidades em termos de capacidade de generalização e precisão nas estimativas geradas pela RNA.

Os resultados obtidos mostram a versatilidade das redes neurais para análise de estruturas fotônicas. De acordo com diversas estruturas neurais e algoritmos de aprendizagem propostos por vários pesquisadores, redes neurais artificiais possuem certas características exclusivas de sistemas biológicos, as quais entram em conflito com os tradicionais métodos computacionais.

Para a aplicação aqui proposta, o sistema de computação baseado em redes neurais tem a capacidade de receber ao mesmo tempo várias entradas e distribuí-las de maneira organizada, gerando uma ou mais saídas. Geralmente, as informações armazenadas por essa rede neural são compartilhadas por todas as suas unidades de processamento, característica que contrasta com os recursos computacionais usados na solução específica, aplicada para solucionar um sistema de equações diferenciais, como o usado pelo VBPM, onde a informação fica confinada em um determinado endereço de memória. Esse processo necessita que a cada passo de propagação o processador acesse informações armazenadas na memória, levando o processo a um excessivo custo computacional.

Os atributos de uma rede neural, tais como aprender por meio de exemplos e generalizações redundantes, possibilita a aproximação de qualquer função contínua não linear com o grau de correção desejado para esta aplicação. Além disso, essa habilidade afasta a necessidade do conhecimento do complexo formalismo eletromagnético indispensável na aplicação dos métodos de onda completa, como o aplicado para desenvolver o VBPM.

Como trabalho futuro pretende-se utilizar as redes neurais artificiais em conjunto com o VBPM na análise de fibras de cristais fotônicos com vários núcleos.

Referências

[1] HERTZ, J.; KROGH, A.; PLAMER, R. G. Introduction to the theory of neural computation, [S.l.]: Santa Fe Institute Studies in the Sciences of Complexity; Addison-Wesley, 1991. (Lecture Notes, 1).

- [2] BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDERMIR, T. B. Redes neurais artificiais: teoria e aplicações. Rio de Janeiro: LTC, 2000. 262 p.
- [3] OKAMOTO, K. Fundamentals of Optical Waveguides. 2. ed. San Diego: Academic Press, 2005. 584 p.
- [4] LEE, D. L. Electromagnetics principles of integrated optics. New York: Wiley and Sons, 1986. 331 p.
- [5] LI, J.; BAO, Z. Neural network model of optical fiber direction coupler design, in Proceedings of the SPIE, v. 5178, p. 238-245, 2004.
- [6] TÜRKER, N.; GUNE, F.; YILDIRIM, T. Artificial neural networks applied to the design of microstrip antennas. Microwave Review, Istanbul, v. 12, n. 1, p. 10-14, jun. 2006.
- [7] OUCHAR, A.; AKSAS, R.; BAUDRAND, H. Artificial neural network for computing the resonant frequency of circular patch antennas. Microwave and Optical Technology Letters, v. 47, n. 6, p. 564-566, 2005.
- [8] DA SILVA, J. P. Simulação por elementos finitos da propagação de feixes óticos em estruturas fotônicas. 2003. 187 f. Tese (Doutorado em Engenharia Elétrica) Departamento de Microondas e Ótica, Universidade Estadual de Campinas, campinas, SP, 2003.
- [9] DAVANÇO, M. I.; RUBIO-MERCEDES, C. E.; HERNÁNDES-FIGUEROA, H. E. "Novel boundary condition for the finite-element solution of arbitrarily planar junction", IEEE Photonics Technology Letters, v. 13, n. 1, p. 46-47, jan. 2001.
- [10] RUBIO-MERCEDES, C. E. Análise de estruturas fotônicas por elementos finitos no domínio da freqüência. 2002. 186f. Tese (Doutorado em Engenharia Elétrica) Departamento de Microondas e Ótica, Universidade Estadual de Campinas, campinas, SP, 2002.
- [11] RUBIO-MERCEDES, C. E.; HERNÁNDES-FIGUEROA, H. E. "Padé boundary conditions for the finiteelement solution of arbitrary planar junctions", IEEE Journal of Lightwave Technology 2002.
- [12] RODRIGUEZ-ESQUERRE, V. F.; RUBIO-MERCEDES, C. E; HERNÁNDES-FIGUEROA, H. E. "Assessment of time and frequency domain finite element algorithms for optical junctions", Anais do 2001 SBM0/IEEE MTT-S International Microwave and Optoelectronics Conference. Belém, Brasil, p. 465-468, ago 06-10, 2001.
- [13] TSUJI, Y. "Finit element method using port truncation by perfectly matched layer boundary conditions for optical waveguide discontinuity problems", IEEE Journal of Lightwave Technology, v. 20, n. 3, p. 463 468, mar. 2002.
- [14] TAFLOVE, A.; HAGNESS, S. C., Computational Electrodynamics, the finite-difference time-domain method, Artech House, In, Second Edition, 2000.
- [15] KOSHIBA, M. Optical Waveguide Theory By the Finite Element Method, KTK Scientific Publishers, 1992.
- [16] PINHEIRO, H. F.; HERNÁNDES-FIGUEROA, H. E., "Novel finite-element formulation for vectorial beam propagation analysis in anisotropic medium", IEEE Photonics Technology Letters, v. 12, n. 2, p. 155-157, fev. 2000.
- [17] SAFAAI-JAZI, A.; MORTAZAI, A. "Graded-index linear tapers", Proceedings of the 26th Southeastern Symposium, p.490-494, mar. 1994.
- [18] ZIOLKOWSKI, R.W.; JUDKINS, J.B. "Nonlinear finit-difference time-domine modeling of linear and nonlinear corrugated waveguides", Journal of the Optical Society of America B (JOSA B), v. 11, p 1565-1575, Set. 1994.

- [19] JOANNOPOULUS, J. D; MEADE, R.D; WINN, J.N. Photonics Crystals: molding the flow of light, Princenton University Press, 1995.
- [20] VERNER, Douglas.H., MITTRA, Raj., editors, Frontiers in Electromagnetics, IEEE Press, N.Y. 2000.
- [21] LEVY, M. et al. "Flat-top response in one-dimensional magnetic photonics bandgap structures with Faraday rotation enhancement", Journal of Lightwaver Technology, v.19, n.12, p.1964-1969, dez. 2001.
- [22] HAYKIN, S. Redes neurais: princípios e prática. 2. ed. Porto Alegre: Bookman, 2001. 900 p.
- [23] MACCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity, Bulletin of Mathematical Biophysics, [S.l.], v. 5, p. 115-133, 1943.
- [24] ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. Psychological Review, Washington, v. 65, n. 6, p. 386-408, 1958.
- [25] HECHT-NIELSEN, R. Neurocomputing. Boston: Addison-Wesley Longman, 1989.
- [26] CYBENKO, G. Approximation by superpositions of a sigmoidal function. The Canadian journal for Chemical Engineering, Illinois, v. 69, n. 4, p. 997-1009, 1989.
- [27] LIPPMANN, R.P. An introduction to computing with neural nets. IEEE ASSP Magazine, [S.l.], v. 4, n. 2, p. 4-22. 1987.
- [28] PATNAIK, A.; MISRA, R. K. ANN Techniques in Microwave Engineering. Microwave Magazine, IEEE, [S.l.], v. 1, n. 1, p. 55-60, mar. 2000.
- [29] RUMELHART, D.; HINTON, D.; WILLIAMS, G. Learning internal representations by error propagation. San Diego: Institute for Cognitive Science, 1985. 34 p.
- [30] MASTERS, T. Signal and image processing with neural networks a C++ Sourcebook. New York: John Wiley & Sons, 1994. 417 p.
- [31] SARLE, W. S. Neural networks and statistical models. ANNAUL SAS USER GROUP INTERNATIONAL CONFERENCE, 19., 1994, Cary, NC: Proceedings... Cary, NC: [s.n.], 1994.
- [32] COURANT, R. Variational methods for the solution of problems of equilibrium and vibration, Bull. Am. Math. Soc, [S.l.], v. 49, p. 1-23, 1943.
- [33] FEYNMAN, R. P. The Feynman lectures on physics. São Paulo: Pearson Education, 2008. 544 p.
- [34] LEVY, R. Directional couplers, in advances in microwaves. New York: Leo Young, 1966. v.1.
- [35] COLLIN, R. E. Foundations for microwave engineering. 2. ed. [S.l.]: Wiley India Pvt., 2009. 944 p.
- [36] MARCUSE, D. Theory of dielectric optical waveguides. 2. ed. New York: Academic Press, 1991. 380 p.
- [37] YARIV, A. Optical electronics. 4. ed. Oxford: Oxford University Press, 1995. 736 p.
- [38] CHEO, P. K. Fiber optics e optoelectronics. 2. ed. New Jersey: Prentice Hall, 1990. 448 p.
- [39] HANSELMAN, D. C.; LITTLEFIELD, B. C.; MARTINS, C. S. Matlab 6: curso completo. São Paulo: Pearson Prentice Hall, 2003. 700 p.