

# Reconhecendo Instrumentos Musicais Através de Redes Neurais Artificiais

Carlos Roberto Ferreira de Menezes Júnior, Eustáquio São José de Faria, Keiji Yamanaka

Faculdade de Engenharia Elétrica (Programa de Pós Graduação) – Universidade Federal de Uberlândia – Uberlândia, MG – Brasil

carlosmenezesjunior@gmail.com, eustaquio@pucminas.br, keiji@ufu.br

**Resumo.** *A Inteligência Artificial desenvolve-se no sentido de modelar o cérebro através da criação de redes neurais artificiais (RNA), que se caracterizam por possuir as mesmas propriedades cognitivas e associativas do cérebro humano. É percebido na literatura que inúmeras redes neurais foram desenvolvidas para resolver problemas das mais diversas naturezas. Este trabalho, em questão, procura aplicar os conceitos de Redes Neurais Artificiais no reconhecimento de instrumentos musicais. Foi desenvolvida, e testada, uma rede neural, através do algoritmo Backpropagation, para reconhecimento, usando dezesseis harmônicos, de cinco instrumentos musicais: violão, piano, flauta, sax tenor e trombone. Os resultados podem ser vistos nas seções seguintes.*

**Palavras-Chave.** *Redes Neurais Artificiais, Harmônicos, Inteligência Artificial, Backpropagation.*

## 1. Introdução

O rápido avanço nas pesquisas e o desenvolvimento dos computadores levaram diversos pesquisadores, na área de inteligência artificial, a acreditar que a construção de computadores pensantes fosse uma tarefa relativamente simples. Obviamente, este era um pensamento prematuro, equivocado justamente por estar pautado apenas na euforia causada pela rápida evolução digital. Para os pioneiros nesta área, o cérebro era comumente separado da mente. Segundo esta visão, o cérebro se constituía apenas como o meio físico de atuação da mente, a qual deveria ser a verdadeira responsável pelo pensamento. Assim, a analogia entre cérebro/mente e *hardware/software* se fazia imediata. A mente poderia ser vista como um *software* instalado em um *hardware* chamado cérebro (WHITE, 1992).

Considerava-se necessário, portanto, desenvolver *softwares* que representassem o comportamento da mente, enquanto que a modelagem do cérebro era visualizada como secundária neste processo. Acreditava-se que, assim como foi possível a construção de máquinas que voam sem, no entanto, bater as asas, seria também possível criar uma máquina de pensar sem a necessidade de criação de um cérebro artificial. Contudo, os fracassos sucessivos desta abordagem acabaram por mostrar a importância da modelagem do cérebro na criação de uma inteligência artificial. A partir de então, a área da inteligência artificial começou a desenvolver-se no sentido de modelar o cérebro através da criação de redes neurais artificiais (RNA), que se caracterizam por possuir as mesmas propriedades cognitivas e associativas do cérebro humano (PORTUGAL e FERNANDES, 1996).

Em Fausett (1994), podem ser encontradas algumas das diversas Redes Neurais Artificiais mais conhecidas. Essas redes podem e são usadas, na literatura, para resolver

problemas de aproximação, de reconhecimento de padrões, entre outros, nas mais distintas áreas do conhecimento (Medicina, Matemática, Lingüística, Música, etc.).

Este trabalho, em questão, procura aplicar os conceitos de Redes Neurais Artificiais no reconhecimento de instrumentos musicais.

As seções posteriores esclarecem o funcionamento de uma Rede Neural Artificial, os conceitos matemáticos relacionados às notas musicais e ao som, a Rede Neural Artificial desenvolvida para reconhecimento de instrumentos musicais e as considerações finais do trabalho.

## **2. O Funcionamento de uma Rede Neural**

No intuito de se construir uma rede neural, o funcionamento dos neurônios biológicos pode ser modelado como sendo um circuito binário que possui várias entradas que são combinadas através de uma soma ponderada (corpo celular), gerando a entrada efetiva do neurônio que por sua vez determinará o repasse desse estímulo a outros neurônios da rede, se o estímulo for excitatório, isto é, se ultrapassar um limiar. Isto é feito através das conexões sinápticas (RIBEIRO e CENTENO, 2001).

Para incitar o comportamento das sinapses, os terminais de entrada do neurônio artificial recebem pesos cujos valores podem ser maiores ou menores que zero, dependendo das sinapses serem inibitórias ou excitatórias. O neurônio dispara quando a soma dos impulsos por ele recebidos ultrapassa o limiar de excitação. A ativação do neurônio ocorre através da função de ativação que aciona ou não a saída, em função do valor da soma ponderada das entradas (BRAGA *et al.*, 2000).

Segundo Freeman e Skapura (1992), e Rumelhart, Hinton, e Williams (1986), um dos algoritmos de aprendizado neural mais utilizados desde a invenção das redes neurais é o algoritmo denominado *Backpropagation*. Este algoritmo é aplicado em redes Perceptron de multicamadas.

Para que o aprendizado aconteça, é necessária uma base de exemplos, representada por um conjunto de dados que descrevem os casos que serão apresentados para a Rede Neural Artificial. O aprendizado das redes que usam *Backpropagation* é conhecido como um aprendizado do tipo supervisionado, onde as redes aprendem a responder de modo similar às respostas apresentadas junto com os casos, formados por um conjunto de dados de entrada mais as saídas desejadas.

Usando um processo iterativo, apresenta-se à rede neural os exemplos contidos na base para que ocorra a adaptação dos pesos (simulando o reforço e a inibição das conexões sinápticas existentes entre os neurônios reais). O aprendizado surge desta adaptação de pesos e, por consequência, surge também a adaptação do comportamento da rede, que deverá aprender a responder aos estímulos de entrada de acordo com os exemplos que lhe foram apresentados.

## **3. Conceitos Matemáticos Básicos de Notas Musicais e de Acústica**

Ao analisar a história da música ocidental, percebe-se que os avanços científicos e tecnológicos de cada período influenciaram o fazer musical e, conseqüentemente, o arcabouço teórico que serve de base para o mesmo (ABDOUNUR, 1999).

Segundo Abdounur (1999), o som pode ser entendido sob alguns conceitos matemáticos razoavelmente simples.

### 3.1. O som

O som é formado pela perturbação do ar como resultado da vibração de um certo corpo elástico. As ondas sonoras (Figura 1) que são utilizadas na música como notas musicais têm vibrações regulares, ou seja, frequências regulares. Cada nota musical tem uma frequência específica e é na relação entre essas notas – entre essas frequências – que foi estruturada toda a teoria musical no ocidente.

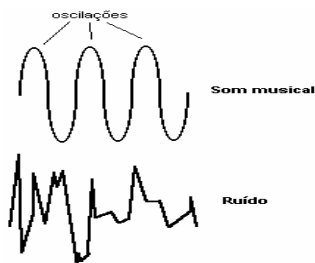


Figura 1. Representação Gráfica das Ondas Sonoras.

### 3.2. Harmônicos, Timbre e a Série Harmônica

Um som musical produzido por qualquer instrumento não é constituído apenas por uma nota, ou seja, por apenas uma frequência, e sim por uma série de outras notas resultantes de vibrações secundárias. A estas outras frequências, deu-se o nome de harmônicos (Figura 2).

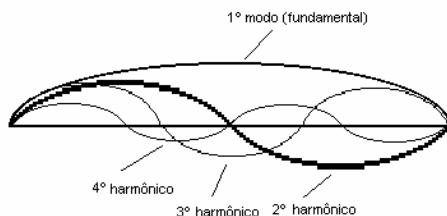


Figura 2. Harmônicos Vibrando em um Corpo Elástico.

O conjunto de harmônicos que acompanha uma nota fundamental forma a Série Harmônica. Toda nota gera uma série harmônica identicamente proporcional. Ao tocar a nota dó de um violão, por exemplo, a corda vibrará inteira e vibrará, também, de forma a dividir a corda pela metade, logo depois, vibrará em três partes iguais, quatro partes iguais e assim por diante, tudo simultaneamente (Figura 3).

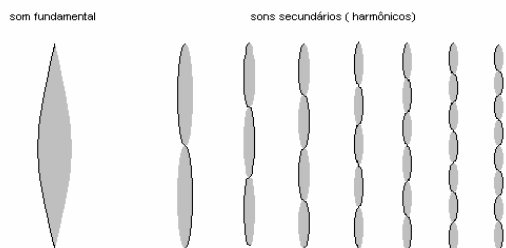
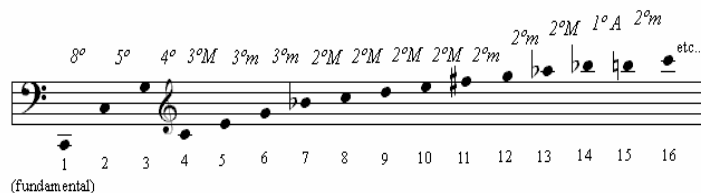


Figura 3 – Vibração dos Harmônicos em uma Corda.

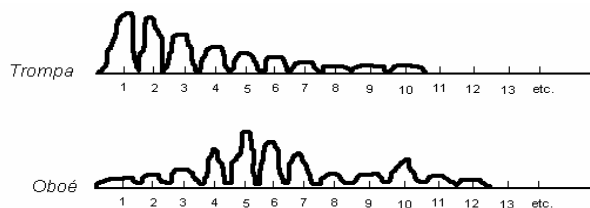
A figura 4 mostra um exemplo dos intervalos musicais formados pelos harmônicos, pegando como base a nota dó.



**Figura 4 – Intervalos Musicais usando-se a nota dó**

Observa-se que o primeiro intervalo a ser formado na série harmônica é a 8° justa, depois é uma 5° justa, 4° justa, 3°maior, 3° menor, etc. A amplitude dos harmônicos mais próximos da nota fundamental é maior do que a dos mais afastados, logo, quanto mais longe da fundamental, menos audível fica o harmônico. Esta relação de intervalos entre os harmônicos é a mesma para qualquer nota fundamental.

É a série harmônica que define a nota que está sendo executada por um determinado instrumento. São as diferentes intensidades dos harmônicos que permitem o reconhecimento dos diversos timbres. Cada instrumento musical tem um gráfico de intensidade de harmônicos característico (Figura 5).



**Figura 5 – Vibração dos Harmônicos em uma Corda Usando a Trompa e o Oboé**

Nem todo instrumento gera a série harmônica completa. A Clarineta, por exemplo, só produz os números ímpares. O que define esta diferença na série harmônica é o material de que é feito o instrumento, o modo como os sons são produzidos e como ressoam.

#### 4. Reconhecedor de Instrumentos Musicais - RIM

Este trabalho implementa uma rede neural artificial, utilizando o algoritmo *Backpropagation* (algoritmo da retro propagação do erro), para detecção de instrumentos musicais a partir de arquivos de áudio digitalizados – em formato WAV.

A rede reconhece cinco instrumentos diferentes, sendo que, para o seu treinamento, foram utilizados três amostras diferentes de cada instrumento. Os instrumentos reconhecidos são: violão, piano, flauta, sax tenor e trombone. É importante ressaltar também que a rede reconhece os instrumentos usando apenas a fundamental “la 440 Hz”.

##### 4.1. Pré-processamento

Para desenvolvimento do trabalho, realizou-se um pré-processamento no intuito de transformar os dados de entrada da rede em padrões de fato reconhecíveis e processáveis por ela. O pré-processamento foi constituído das seguintes atividades:

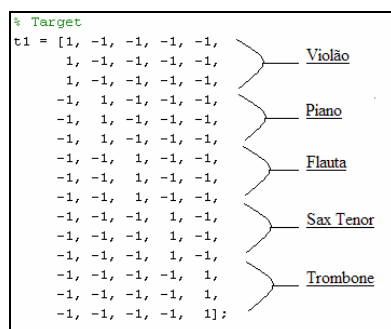
1. Abertura do arquivo padrão de entrada (WAV) e cálculo da FFT (Transformada de Fourier);

2. Cálculo do valor dos harmônicos para o padrão de entrada usando a fundamental “la 440”;
3. Normalização dos dados de entrada ( $-1 \leq \text{dados} \leq 1$ ) segundo figura 6;

$$X_{\text{normalizado}} = \frac{2 \cdot (X_{\text{entrada}} - \text{Mínimo})}{(\text{Maximo} - \text{Mínimo})} - 1$$

**Figura 6. Fórmula de Normalização dos Valores de Entrada.**

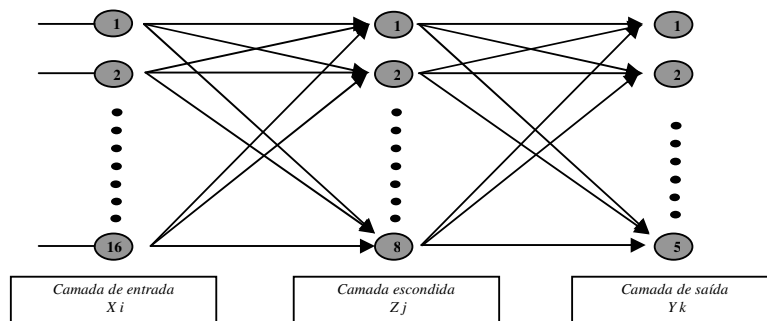
4. Repetição das atividades anteriores até que seja processado o último arquivo padrão de entrada;
5. Definição dos Targets da rede (Figura 7).



**Figura 7. Targets da Rede Neural.**

## 4.2. Arquitetura da Rede

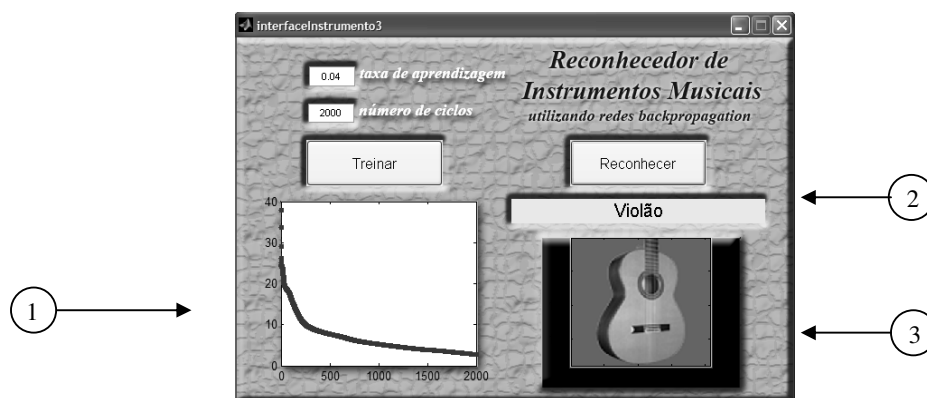
Conforme pode ser visto na figura 8, a rede possui 16 neurônios na camada de entrada, pois foram usados 16 harmônicos no padrão de entrada, 8 neurônios na camada escondida e 5 neurônios na camada de saída (representando capacidade para reconhecer 5 instrumentos diferentes).



**Figura 8. Arquitetura da Rede Neural.**

## 4.3. O Programa RIM

Procurou-se desenvolver uma interface simples e de fácil utilização. Na janela principal do RIM (Figura 9), podem ser preenchidas duas caixas de edição: taxa de aprendizagem ( $\alpha$  - alfa) e número de ciclos (épocas).



**Figura 9. RIM Reconhecendo um Violão.**

Após escolhidos a taxa de aprendizagem e o número de ciclos, o usuário deve clicar no botão “Treinar”. O gráfico do erro quadrático total por ciclos é mostrado, na janela do RIM, logo abaixo deste botão (área 1 da figura 9). A descrição e a figura do instrumento reconhecido podem ser encontrados nas áreas 2 e 3, da figura 9, respectivamente.

#### 4.4. Reconhecendo Instrumentos

A tabela 1 mostra o reconhecimento de instrumentos com ruídos, usando o RIM, trabalhando com 100, 500, 1000, 1500 e 2000 ciclos.

**Tabela 1. Reconhecimento de Instrumentos Com Ruídos.**

Número de Ciclos	Reconhecimento do Violão	Reconhecimento do Piano	Reconhecimento da Flauta	Reconhecimento do Sax Tenor	Reconhecimento do Trombone
100	N	N	N	N	S
500	S	N	S	S	S
1000	S	N	S	S	S
1500	S	S	S	S	S
2000	S	S	S	S	S

Observando a tabela 1, nota-se que somente os trombones foram reconhecidos quando a rede foi treinada com 100 ciclos. Em 1000 ciclos, apenas os pianos não foram reconhecidos pelo RIM. Em 1500 ciclos, todos os instrumentos foram reconhecidos com êxito.

## 5. Considerações Finais

### 5.1 Conclusões

Conforme esperado, com a simplicidade do teste efetuado, a rede neural artificial conseguiu reconhecer 100% dos padrões treinados a partir de uma quantidade específica de ciclos de treinamento – neste caso, a partir de 1500 ciclos.

Acredita-se que a normalização, executada na atividade de pré-processamento, desempenhou um papel importante no reconhecimento dos padrões, contribuindo para a eficácia do processo de treinamento.

Durante o levantamento do referencial teórico, percebeu-se que o conjunto de treinamento é extremamente importante para o desenvolvimento da rede neural. Procurando-se encontrar o conjunto de dados relevantes para reconhecimento de instrumentos musicais, descobriu-se que 16 harmônicos seriam suficientes. Estes harmônicos foram calculados usando-se a fundamental “1a 440Hz”. É importante ressaltar que, para o cálculo dos harmônicos, fez-se

necessário o cálculo da transformada de Fourier em cada um dos arquivos “WAV” padrões de entrada da rede.

## 5.2 Trabalhos Futuros

Propõem-se, como trabalhos futuros:

- Aplicação de um teste estatístico para confirmar e validar os resultados experimentais no RIM – se possível, usando arquivos WAV de entrada com uma diversidade maior de tipos de ruídos;
- Aplicação de testes com técnicas como leave-one-out, por exemplo, visando definir diversas taxas de erro;
- O desenvolvimento de uma rede neural artificial capaz de reconhecer instrumentos musicais a partir de qualquer nota musical;
- O desenvolvimento de uma rede neural capaz de reconhecer instrumentos distintos em um único arquivo de som.

## Referências

- ABDOUNUR, O. J.. **Matemática e Música. O Pensamento Analógico na Construção de Significados**. São Paulo: Escrituras. 1999.
- BRAGA, A .P. ; LUDEMIR, T. B. ; CARVALHO, A . C. P. F.. **Redes neurais artificiais – teoria e aplicações**. Rio de Janeiro: LTC Livros Técnicos e Científicos Editora S. A, 2000.
- FAUSETT, L. V. **Fundamentals of Neural networks: architectures, algotithms, and applications**. Florida Institute of Tecnology, 1994.
- FREEMAN, James A. & SKAPURA, David M. **Neural Networks: Algorithms, Applications, and Programming Techniques**. Addison-Wesley Publishing, Reading, 1992. 401p.
- PORTUGAL, Marcelo S.; e FERNANDES, Luiz Gustavo L.. **Redes Neurais Artificiais e Previsão de Séries Econômicas: Uma Introdução**. Departamento de Economia, Universidade Federal de Minas Gerais, Brasil, Revista Nova Economia, vol. 6(1), pp 51-74, Julho. 1996.
- RIBEIRO, Selma Regina Aranha; e CENTENO, Jorge Silva. **Classificação do Uso do Solo Utilizando Redes Neurais e o Algoritmo MAXVER**. Anais do X SBSR. Foz do Iguaçu – PR, Brasil, 2001.
- RUMELHART, D. ; HINTON, G; e WILLIAMS, R. **Learning Internal Representations by Error Propagation**. In: Parallel Distributed Processing: Explorations in the Microstructure of Cognition - Vol. 1. Cambridge: MIT Press, 1986.
- WHITE, H.. **Artificial Neural Networks: Approximation and Learning Theory**. Oxford: Blackwell Publishers, 1992.