

## **Aprendizagem Profunda Aplicada ao Problema de Contar o Número de Presentes em um Ambiente**

### **Deep Learning Applied to the Problem of Counting the Number of Gifts in an Environment**

DOI:10.34117/bjdv7n4-123

Recebimento dos originais: 07/03/2021

Aceitação para publicação: 06/04/2021

#### **Millena Michely de Medeiros Campos**

Graduação em Engenharia Elétrica pela Universidade Federal do Rio Grande do Norte  
Sidia Instituto de Ciência e Tecnologia, Manaus, AM  
Av. Darcy Vargas, 654 - Parque Dez de Novembro, Manaus - AM, 69055-035  
E-mail: millena.campos@sidia.com

#### **José Martins de Castro Neto**

Mestrado em Engenharia Elétrica e da Computação pela Universidade Federal do Rio Grande do Norte  
Sidia Instituto de Ciência e Tecnologia, Manaus, AM  
Av. Darcy Vargas, 654 - Parque Dez de Novembro, Manaus - AM, 69055-035  
E-mail: jose.neto@sidia.com

#### **Vicente Angelo de Sousa Jr.**

Doutor em Engenharia de Teleinformática (2009) pela Universidade Federal do Ceará  
Universidade Federal do Rio Grande do Norte  
59078 970 Avenida Senador Salgado Filho, 3000 - Lagoa Nova, Natal – RN  
E-mail: vicente.sousa@ufrn.edu.br

#### **Julian Wendel Ribeiro**

Estudante de Engenharia Elétrica com habilitação em Telecomunicações pela  
Universidade Federal de Juiz de Fora  
Universidade Federal de Juiz de Fora  
Campus Universitário, Rua José Lourenço Kelmer, s/n - São Pedro, Juiz de Fora - MG,  
36036-900  
E-mail: julian.wendel@engenharia.ufjf.br

#### **Rafael da Silva Macêdo**

Graduação em Engenharia Elétrica pela Universidade Federal de Juiz de Fora  
Universidade Federal de Juiz de Fora  
Campus Universitário, Rua José Lourenço Kelmer, s/n - São Pedro, Juiz de Fora - MG,  
36036-900  
E-mail: rafael.macedo2016@engenharia.ufjf.br

#### **Alvaro Augusto Machado de Medeiros**

doutorado em Engenharia Elétrica pela Universidade Estadual de Campinas  
Universidade Federal de Juiz de Fora  
Campus Universitário, Rua José Lourenço Kelmer, s/n - São Pedro, Juiz de Fora - MG,  
36036-900  
E-mail: alvaro@engenharia.ufjf.br

## RESUMO

O sinal recebido contém informações importantes sobre o meio que se propaga, as quais podem ser extraídas e utilizadas para diversos fins. Tais informações são comumente obtidas aplicando-se ferramentas estatísticas. Este trabalho propõe o uso de aprendizado de máquina para contar o número de pessoas em um ambiente indoor usando sinais de radiofrequência. Se valendo da variabilidade inserida no canal sem fio pela presença e movimento de pessoas, os dados obtidos têm a sua dimensionalidade reduzida, via ferramenta de Deep Learning, e posteriormente analisados por três classificadores. Os resultados demonstram que é possível obter o número de presentes em um ambiente com pelo menos 90\% de acurácia.

**Palavras-chave:** Contagem de pessoas, redução de dimensionalidade, sinal de RF.

## ABSTRACT

The received signal contains important information about the propagating medium, which can be extracted and used for different purposes. Such information is commonly obtained by applying statistical tools. This work proposes the use of machine learning to count the number of people in an indoor environment using radio frequency signals. Taking advantage of the variability inserted in the wireless channel by the presence and movement of people, the data obtained has its dimensionality reduced, through the Deep Learning tool, and subsequently analyzed by three classifiers. The results demonstrate that it is possible to obtain the number of gifts in an environment with at least 90 \% accuracy.

**Keywords:** Counting people, dimensionality reduction, RF signal.

## 1 INTRODUÇÃO

Com a permanente transformação da tecnologia, cada vez mais dispositivos móveis se incorporaram em diversos aspectos do cotidiano da humanidade, fazendo com que o sinal de radiofrequência (RF) seja utilizado para outros fins além de carregar a informação transmitida.

Uma das aplicações propostas na literatura é a detecção de pessoas em ambientes internos por meio da análise de um sinal de RF [1]-[3]. Porém, várias dessas abordagens não determinam o número de pessoas já presentes no ambiente, apenas a presença ou não de indivíduos. Além disso, são baseadas em tecnologias de redes sem fio como Wi-Fi e ZigBee. Este artigo tem o objetivo de propor uma solução, baseada em sinal de RF simples (um portadora senoidal), para contagem de pessoas em um ambiente já povoado, utilizando redes Autoencoder, para redução de dimensionalidade, e três classificadores: o SVM, a Rede Perceptron de múltiplas camadas (Multi-layer Perceptron - MLP) e a Rede RBF. Com o Autoencoder permitiu-se reduzir a complexidade dos dados avaliados e consequentemente diminuir o tempo de treinamento dos diferentes classificadores.

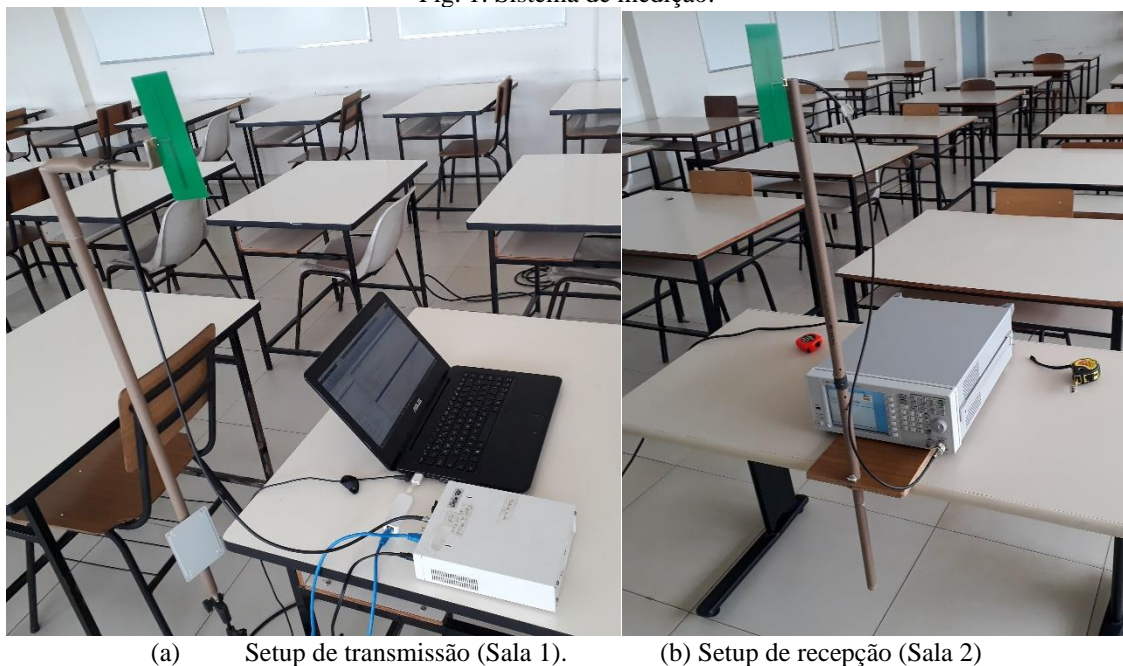
## 2 MATERIAIS E MÉTODOS

Nesta seção são abordados o setup e a campanha de medição, bem como os dados medidos foram organizados.

O sinal emitido consistiu em uma portadora em 915 MHz gerada pelo RF Agilent N9310A [4], acoplado a uma antena dipolo de microfita na polarização vertical, a uma altura de 1,40~m do chão, como mostrado na Fig.1(a).

No sistema de recepção (Fig.1(b)) foi utilizada a plataforma de rádio definido por software `\emph{Universal Software Radio Peripheral}` (USRP) N210, juntamente com a placa WBX que opera nas frequências de 50-2200 MHz. A antena utilizada foi a mesma do sistema transmissor, a uma altura de 1,40~m do chão. Para o controle da USRP N210 foi utilizado um computador dotado de um script com o intuito de automatizar e salvar as medidas para posterior tratamento estatístico.

Fig. 1. Sistema de medição.



As campanhas de medição foram realizadas na Universidade Federal de Juiz de Fora (UFJF) no prédio de engenharia, utilizando duas salas vizinhas. Para garantir que as medições não fossem influenciadas por outras pessoas, as campanhas de medição foram realizadas durante o período de férias.

Foram realizadas três campanhas de medição em um cenário sem linha de NLOS (non-line-of-sight), variando o número de pessoas entre 1-3, com uma distância entre o receptor e o transmissor de 7,7 m. Cada campanha consistiu em 1000 amostras a uma taxa

de amostragem de 0,1 s. As pessoas se movimentaram com uma velocidade constante, tentando manter a uma distância fixa entre si.

As estatísticas descritivas foram calculadas em médias móveis janeladas do sinal, sendo a janela de tamanho 100, de forma a mapear o comportamento do sinal para cada caso ao longo do tempo. Comumente são atribuídas diferentes métricas para caracterizar o sinal, foram adotadas neste trabalho média, desvio padrão, kurtosis, entropia e os momentos 4, 6, 8 e 10 ( $m_4$ ,  $m_6$ ,  $m_8$  e  $m_{10}$ ).

### 3 SOLUÇÃO PROPOSTA

O Autoencoder é uma rede neural treinada com o objetivo de copiar a sua entrada na saída. Internamente, essa rede possui dois estágios com uma, ou um conjunto de camadas escondidas. Esses estágios são o Encoder e o Decoder, e representam a codificação e decodificação da entrada, respectivamente. O Encoder é responsável por representar a entrada da rede em uma dimensão menor do que a dimensão de entrada. Já o Decoder faz o caminho reverso, i.e, tem como entrada os dados da saída do Encoder e tenta reconstruir os dados originais na sua saída. Esse modelo prioriza aspectos mais importantes da entrada para poder reconstruí-la na saída, assim ele é capaz de aprender propriedades úteis dos dados e representá-los na saída do Encoder [5].

Uma vez que a dimensionalidade da entrada é reduzida pelo Encoder, sua saída pode ser utilizada como entrada para classificadores. A vantagem de reduzir a dimensionalidade é a simplificação da montagem das redes subsequentes. Neste trabalho, a saída do Encoder é usada como entrada para três classificadores.

Na etapa de pré-simulação, os dados são normalizados com objetivo de evitar que a rede trabalhe com alta variância. Além disto, são separados 75 % dos dados para treinamento e 25 % para validação. A rede Autoencoder tem os seguintes parâmetros: arquitetura 8-16-12-4-12-16-8, 50 épocas, 16 batches, função de ativação Relu, otimizador adagrad e a métrica sendo o erro médio quadrático. A camada de saída do Encoder é composta por 4 neurônios, indicando uma diminuição de 8 para 4 na dimensão dos dados de entrada.

Na saída do Encoder são aplicados três classificadores, sendo eles:

- Rede MLP: arquitetura 64-32-16-3, 30 épocas, 64 batches, função de ativação Relu e sigmoide (saída), otimizador rmsprop e métrica acurácia;
- Rede SVM: parametrizada por  $C=6$ ,  $\gamma=5$  e kernel RBF;

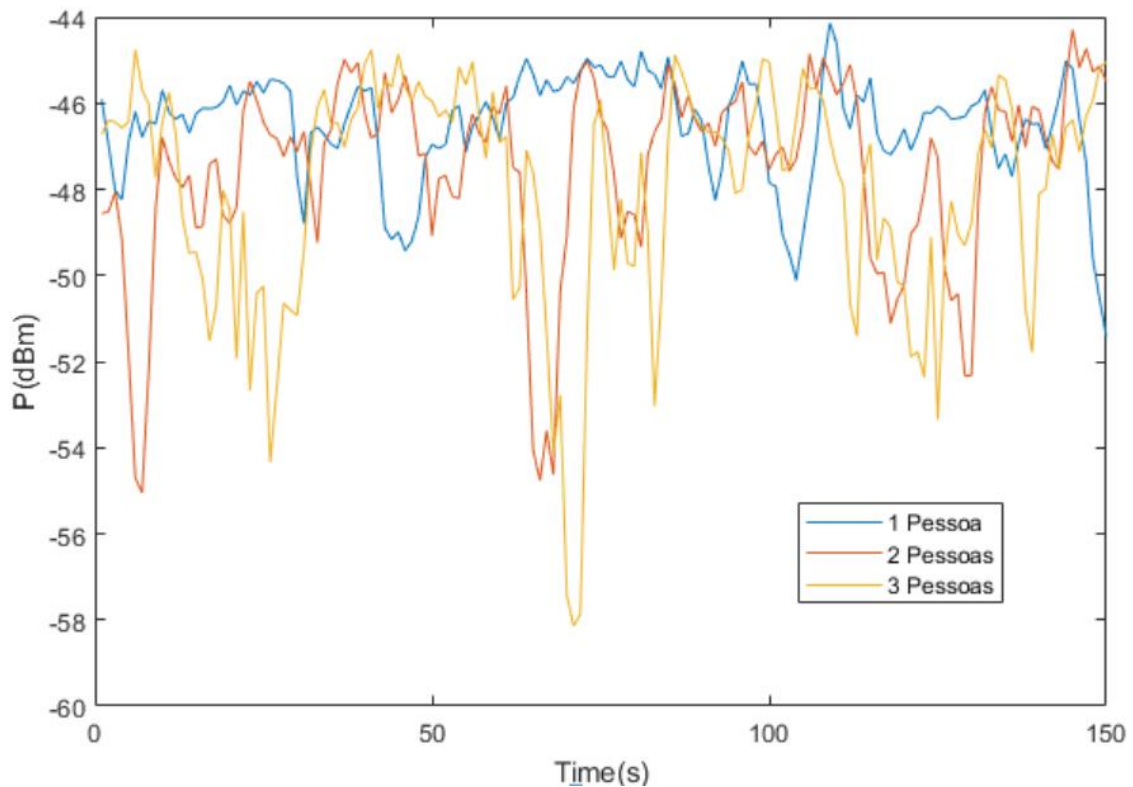
• Rede RBF: com 100 núcleos iniciados pelo algoritmo k-means e a variância calculada por  $d_{\max}/N_{\text{nucleos}}$ , sendo  $N_{\text{nucleos}}$  igual ao número de classes, neste caso 3, e  $d_{\max}$  a máxima distância entre dois núcleos.

Nota-se que a classificação é realizada levando em conta que há três classes, 1 pessoa, 2 pessoas ou 3 pessoas.

#### 4 RESULTADOS E CONSIDERAÇÕES FINAIS

Uma primeira análise é feita no sinal recebido em função do tempo na Fig2. Percebe-se que um aumento de pessoas na sala influencia diretamente a intensidade dos deep fades e a variação do sinal, porém, por uma simples inspeção visual não é possível separar os casos e fazer uma avaliação quantitativa.

Fig. 2. Potência do sinal recebido vs tempo.



Para o Autoencoder, o treinamento foi realizado seguindo os parâmetros indicados na Seção III. Obteve-se um erro médio quadrático de 0,003 tanto para o treinamento, como para a validação. Assim, ver-se que o Autoencoder consegue aproximar muito bem a saída com a entrada, mostrando também que a saída do Encoder (4 dimensões) contém informações suficientes para a reconstrução do sinal de entrada.

Para os classificadores MLP, SVM e RBF, a Tabela 1 apresenta a diagonal principal das matrizes de confusão para os dados de validação, i.e., os valores de acurácia por classe para cada classificador. Para o caso de apenas uma pessoa, os classificadores acertam praticamente 100 % das vezes. Já para 2 pessoas o desempenho não é excelente, mas é próximo de 97 % para o SVM e a MLP. No geral, nota-se que o desempenho para a MLP e o RBF é um pouco inferior ao desempenho da SVM para todas as classes, porém os resultados ainda são acima de 90 % para cada classe (exceto para a classe com 3 pessoas para a rede RBF, com 89 % de acurácia). Além disso, pela análise da tabela é possível notar que os classificadores tendem a errar mais à medida que o número de pessoas aumenta.

Tabela 1: Acurácia por classe para cada classificador.

Número de Pessoas	Classificadores		
	MPL	SVM	RBF
1	99,92%	99,92%	99,34%
2	96,70%	97,11%	95,30%
3	90,39%	90,64%	89,06%

Por meio das análises dos resultados, notou-se que apenas o uso do SVM para o problema de classificação, dado as entradas com dimensão reduzida pelo Autoencoder, já apresenta ótimos resultados de acurácia. Novas medições, incluindo mais pessoas e com diferentes comportamentos (em movimento, paradas, sentadas, em pé), são objetos de trabalhos futuros.

## REFERÊNCIAS

- [1] S. Ikeda, H. Tsuji, and T. Ohtsuki, "Indoor event detection with eigenvector spanning signal subspace for home or office security," *IEICE Transactions*, vol. 92-B, pp. 2406–2412, 07 2009.
- [2] M. Sekine and K. Maeno, "Activity recognition using radio doppler effect for human monitoring service," *Journal of Information Processing*, vol. 20, p. 396, 04 2012.
- [3] C. Hekimian-Williams, B. Grant, Xiuwen Liu, Zhenghao Zhang, and P. Kumar, "Accurate localization of rfid tags using phase difference," in *2010 IEEE International Conference on RFID (IEEE RFID 2010)*, April 2010, pp. 89–96.
- [4] A. Technologies, *Agilent N9310A RF Signal Generator - User's Guide*, jun 2006.
- [5] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, <http://www.deeplearningbook.org>.