



**UNIVERSIDADE FEDERAL DO TOCANTINS
CÂMPUS UNIVERSITÁRIO DE PALMAS
CURSO DE CIÊNCIA DA COMPUTAÇÃO**

**UMA ABORDAGEM DE APRENDIZAGEM PROFUNDA PARA
IDENTIFICAÇÃO DE PESSOAS ATRAVÉS DO SOM DOS PASSOS**

LEONARDO REZENDE COSTA

PALMAS (TO)

2018

LEONARDO REZENDE COSTA

UMA ABORDAGEM DE APRENDIZAGEM PROFUNDA PARA IDENTIFICAÇÃO
DE PESSOAS ATRAVÉS DO SOM DOS PASSOS

Trabalho de Conclusão de Curso I apresentado
à Universidade Federal do Tocantins para
obtenção do título de Bacharel em Ciência da
Computação, sob a orientação do(a) Prof.(a)
Dr. Rafael Lima de Carvalho.

Orientador: Dr. Rafael Lima de Carvalho

Coorientador: Me. Tiago da Silva Almeida

PALMAS (TO)

2018

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da Universidade Federal do Tocantins

- C837a Costa, Leonardo Rezende.
Uma abordagem de aprendizagem profunda para identificação de pessoas através do som dos passos. / Leonardo Rezende Costa. – Palmas, TO, 2018.
40 f.
- Monografia Graduação - Universidade Federal do Tocantins – Câmpus Universitário de Palmas - Curso de Ciências da Computação, 2018.
Orientador: Rafael Lima de Carvalho
Coorientador: Tiago da Silva Almeida
1. Computação Ubíqua. 2. Som dos passos. 3. Aprendizado Profundo. 4. Redes Neurais Artificiais. I. Título

CDD 004

TODOS OS DIREITOS RESERVADOS – A reprodução total ou parcial, de qualquer forma ou por qualquer meio deste documento é autorizado desde que citada a fonte. A violação dos direitos do autor (Lei nº 9.610/98) é crime estabelecido pelo artigo 184 do Código Penal.

Elaborado pelo sistema de geração automática de ficha catalográfica da UFT com os dados fornecidos pelo(a) autor(a).

LEONARDO REZENDE COSTA

UMA ABORDAGEM DE APRENDIZAGEM PROFUNDA PARA IDENTIFICAÇÃO
DE PESSOAS ATRAVÉS DO SOM DOS PASSOS

Trabalho de Conclusão de Curso I apresentado à UFT – Universidade Federal do Tocantins – Câmpus Universitário de Palmas, Curso de Ciência da Computação foi avaliado para a obtenção do título de Bacharel e aprovada em sua forma final pelo Orientador e pela Banca Examinadora.

Data de aprovação: 7 / 12 / 2018

Banca Examinadora:

Prof. Dr. Rafael Lima de Carvalho

Prof. Dr. Alexandre Tadeu Rossini da Silva

Prof. Dr. Warley Gramacho da Silva

AGRADECIMENTOS

Agradeço de coração a todo mundo que contribuiu direta e indiretamente ao longo do curso para o desenvolvimento deste trabalho. Em especial a quatro mulheres muito especiais: Valquiria, minha mãe, que é a mulher mais incrível que eu conheço; à Ana Paula, minha irmã, que sempre deu apoio da melhor maneira possível; à Jade, que dividiu comigo os bons e maus momentos desses últimos quatro anos e à Lara, a amiga que deu dicas acadêmicas preciosíssimas.

Conheci muitas pessoas ao decorrer do curso. É difícil citar todo mundo, mas os ensinamentos de paciência, amor e sabedoria estarão sempre guardados em minha mente. Thaylon, valeu por sempre estar disposto a tirar alguma dúvida e por estar sempre estimulando as pessoas a buscarem mais. Herinson, obrigado por mostrar o lado científico que vai muito além de desenvolvimento. Caio, valeu por me mostrar essa visão diferente de mundo, sempre muito sábio e muito empático. Osmir, por sempre manter a humildade e cativar as pessoas com essa simpatia que ainda vai longe. Felipe Reis, que tá SEMPRE empolgado e vê o melhor nos outros. Jonatha, que tem um senso de humor impecável e tá sempre disposto a ajudar. E aos que estiveram mais próximos nesse último ano: Feitosa, Larissa, Felipe Barbosa, Matheus e Jessivan. Vocês são demais.

Quero também enaltecer alguns professores que eu admiro muito: Glenda, por trabalhar sempre com ética, Rossini pela ótima preocupação em dar uma boa aula, Magalhães, pela atenção e dicas práticas sem igual, Hellena, que sempre bate na tecla certa mas a gente demora a entender e um agradecimento especial aos meus orientadores Rafael Lima e Tiago Almeida pelo apoio e por serem exemplos inquestionáveis de caráter, no melhor estilo professor/pesquisador em que alguém poderia se inspirar.

RESUMO

Com o advento da Computação Ubíqua, a tecnologia passou a fazer parte do cotidiano do ser humano. Deste modo, na construção de ambientes “inteligentes”, são empregados sistemas ligados à rotina dos indivíduos que ali habitam. Assim, uma das principais necessidades é o reconhecimento do indivíduo que convive no espaço em questão. O objetivo deste trabalho é identificar os indivíduos em um ambiente inteligente através de informações obtidas pelo som de seus passos. Para isso, foram desenvolvidas e testadas diferentes configurações de redes neurais de aprendizado profundo com o propósito de classificar as pessoas que participaram de um experimento realizado em Carvalho e Rosa (2010). A arquitetura de rede neural proposta é composta por um encadeamento de redes neurais e seus resultados alcançaram até 98,57% de acurácia.

Palavra-chave: Computação Ubíqua. Som dos passos. Aprendizado Profundo. Redes Neurais Artificiais.

ABSTRACT

With the advent of pervasive computing, the technology has become part of human daily life. Thus, in the construction of "intelligent" environments, systems are used that are linked to the routine of the individuals who live there. So one of the main needs is the recognition of the individual who lives in the space. The objective of this work is to identify individuals in an intelligent environment through information obtained by the sound of their steps. To this end, different configurations of neural networks of deep learning were developed and tested in order to classify the people who participated in an experiment realized in Carvalho and Rosa (2010). The proposed neural network architecture is composed of a chain of neural networks and the results achieved to 98.57 % accuracy.

Keywords: Pervasive Computing. Footstep Sound. Deep Learning. Artificial Neural Network.

LISTA DE FIGURAS

Figura 1 – Gráfico que relaciona as amostras de cada espécie de acordo com o peso	15
Figura 2 – Gráfico que relaciona as amostras de cada espécie de acordo com o brilho da escama.	16
Figura 3 – Gráfico que relaciona duas <i>features</i> peso e brilho das escamas, para as espécies A e B.	17
Figura 4 – Conjunto de dados que podem ser agrupados em três categorias diferentes. Um exemplo de <i>clustering</i>	18
Figura 5 – Exemplos de problemas que envolvem reconhecimento de padrões e classificação. Traduzido de Samarasinghe (2016)	18
Figura 6 – Exemplo de pontos gerados em um plano qualquer, onde se encontra uma função parecida com a realidade do conjunto.	19
Figura 7 – Exemplo de pontos previstos em uma função que relaciona a quantidade de ocorrências de chuva por um período de tempo. Traduzido de Samarasinghe (2016).	19
Figura 8 – <i>Perceptron</i> formado com apenas um neurônio.	20
Figura 9 – Diagrama de um perceptron	21
Figura 10 – Diagrama de uma Rede Neural do tipo <i>Multilayer perceptron</i> com três camadas: entrada, intermediária e saída.	22
Figura 11 – Diagrama de uma Rede Neural de Aprendizado Profundo	23
Figura 12 – Configuração do experimento realizado por Carvalho e Rosa.	26
Figura 13 – Gráfico que compara a <i>feature</i> Envelope Espectral de 6 indivíduos diferentes usando o mesmo calçado.	28
Figura 14 – Padrão de encapsulamento das características extraídas no trabalho anterior	29
Figura 15 – Fluxograma que representa o processo de desenvolvimento da Rede Neural Artificial.	30

Figura 16 – Acurácia e perda durante o treinamento da <i>feature loudnesszwicker</i>	32
Figura 17 – Acurácia e perda durante o treinamento da <i>feature lpcCeps</i>	32
Figura 18 – RNA que combina as <i>features loudnesszwicker</i> e <i>lpcCeps</i> em um único vetor de características	33
Figura 19 – Arquitetura multirede que combina as saídas das RNA's anteriores em uma nova rede.	34
Figura 20 – Acurácia e perda durante o treinamento da arquitetura multirede que combina três redes neurais	34

SUMÁRIO

1	INTRODUÇÃO	11
2	REFERENCIAL TEÓRICO	13
2.1	Classificação	13
2.1.1	Exemplo da fábrica de peixes enlatados	13
2.2	Rede neural artificial	17
2.3	Aprendizado profundo (<i>Deep learning</i>)	21
3	TRABALHOS RELACIONADOS	24
4	METODOLOGIA	26
4.1	Experimento	26
4.2	Desenvolvimento da rede neural	28
5	RESULTADOS	31
5.1	Descrição do experimento	31
5.2	Uma arquitetura multirede	33
6	CONCLUSÃO	36
	REFERÊNCIAS	38

1 INTRODUÇÃO

Na medida que o desenvolvimento tecnológico aumentou a capacidade de processamento em dispositivos cada vez menores, mais e mais itens do cotidiano passaram a receber algum aprimoramento por circuitos microcontrolados. A tendência é que esse aprimoramento aconteça de maneira discreta e é neste contexto que surge a Computação Pervasiva.

Também chamada de Computação Ubíqua, a Computação Pervasiva acontece quando a tecnologia se encontra nos objetos do dia-a-dia de uma maneira quase natural. Ela está na pulseira que monitora frequência cardíaca, no tênis que colhe informações sobre o caminhar do usuário, na geladeira que melhora a sua eficiência energética de maneira autônoma e vários outros objetos que estão incorporados na nossa rotina com tecnologia embarcada sem que isso chame a atenção. A principal vantagem desse ramo é a junção de praticidade com conforto. Isso se dá pelo aprimoramento de procedimentos rotineiros através de uma tecnologia que quase não é percebida pelo usuário.

Uma das aplicações da Computação Ubíqua é o reconhecimento de indivíduos através de informações colhidas sem interação direta. Nessa área, as informações podem ser usadas para procedimentos de segurança, como é sugerido em Larsen et al. (2013) e Shoji, Takasuka e Yasukawa (2004). Podem também ser usadas para identificar indivíduos de grupos específicos, como o sistema proposto para identificar se uma pessoa idosa sofreu uma queda e, conseqüentemente, precisa de ajuda (KHAN et al., 2015). Outra aplicação é o uso de sistemas de identificação em casas ou ambiente de trabalho para o aumento de conforto e otimização de recursos, como proposto em Orr e Abowd (2000), Rosa et al. (2004), Savio e Ludwig (2007), Carvalho e Rosa (2010), Ding et al. (2011) e Collotta e Pau (2015).

São diversas as características que podem ser usadas para distinguir as pessoas no dia-a-dia. Uma delas é o reconhecimento biométrico por impressão digital ou por leitura da íris. Outra é o reconhecimento por imagem, amplamente utilizado em sistemas de segurança. Uma terceira característica, que recentemente vem sendo explorada para esse fim, é o caminhar.

Recentemente a análise do caminhar vem sendo considerada como ferramenta para reconhecimento biométrico. Aristóteles foi a primeira pessoa a tomar notas sobre o caminhar de humanos e animais, como pode ser visto em “*On the Gait of Animals*” Aristóteles (350 A.C.) (CONNOR; ROSS, 2018). No trabalho de Connor e Ross é possível conferir uma análise completa da evolução histórica da análise sobre o caminhar do homem. Nele, pode se ver que são muitas as áreas de interesse que envolvem a análise e o reconhecimento pelo modo que as pessoas andam. A quantidade de características é suficiente para

identificar uma pessoa, seja por imagem, por dispositivos que capturam a pressão dos passos, ou pelo som que os passos emitem. Esta última área foi explorada, tendo o som dos passos como característica de classificação nos trabalhos de Carvalho e Rosa (2010) e Shoji, Takasuka e Yasukawa (2004).

Uma estratégia comumente utilizada para classificação é a aplicação de algoritmos de redes neurais no reconhecimento de indivíduos. O conceito de redes neurais não é recente: McCulloch e Pitts propuseram o que seria o primeiro modelo de rede neural artificial (RNA) ainda em 1943, no trabalho McCulloch e Pitts (1943). Entretanto, com o baixo poder computacional, esse ramo da inteligência artificial foi pouco explorado até meados dos anos 2000. Foi então que o aumento da capacidade de processamento das unidades gráficas (ou *Graphics Processing Unit* (GPU), em inglês), juntamente com a vasta base de dados que é disponível nas mais diversas áreas formou o ambiente ideal para a aplicação de redes neurais artificiais nos problemas do cotidiano. Para resolver problemas mais complexos, surgiu a rede neural de aprendizado profundo (ou *Deep Neural Network*, DNN, em inglês), que nada mais é do que uma RNA com muitas camadas de profundidade (por isso o nome). Uma das maiores aplicações de redes neurais de aprendizado profundo é a classificação de objetos através de imagens.

Dado esse contexto, o objetivo deste trabalho é unir a análise do caminhar com a classificação através de uma rede neural de aprendizado profundo. Para isso, os dados colhidos no trabalho de Carvalho e Rosa (2010) foram estudados para montar um modelo de classificação e então uma rede neural de aprendizado profundo foi desenvolvida, como descrito no Capítulo 4. O objetivo é explorar o som dos passos como característica e investigar se uma DNN pode classificar os indivíduos do experimento de maneira eficiente. Como descrito no Capítulo 5, os resultados se mostraram satisfatórios, alcançando mais de 98% de precisão através de apenas duas características.

2 REFERENCIAL TEÓRICO

Neste capítulo os temas abordados são referentes às bases do trabalho: classificação, redes neurais e, sua variação, redes neurais de aprendizado profundo. Alguns exemplos são dados para melhor entendimento da abordagem proposta que é descrita no capítulo 4.

2.1 Classificação

Com o avanço da tecnologia, o desejo de desenvolver máquinas capazes de reconhecer padrões surgiu de forma quase natural. As áreas de aplicação envolvem as diversas áreas de conhecimento, como: científica, médica, industrial, agropecuária, governamental, militar etc.. A necessidade de resolver diversos problemas fez com que o entendimento sobre o reconhecimento de padrões ganhasse mais ênfase. Muitas vezes as aplicações tem inspiração na maneira como ocorre na natureza.

Existem diversos algoritmos para classificação nos mais diferenciados contextos. Mas o que todas tem em comum é a maneira de preparar o problema para ser resolvido. Desse modo, antes de classificar o que quer que seja, deve-se estruturar o problema, pensar nas características que melhor representam os indivíduos e estudar como avaliar o resultado. Elaborando um pouco mais esse contexto, é possível conferir em Duda, Hart e Stork (2012) um exemplo de classificação voltada ao meio industrial.

2.1.1 Exemplo da fábrica de peixes enlatados

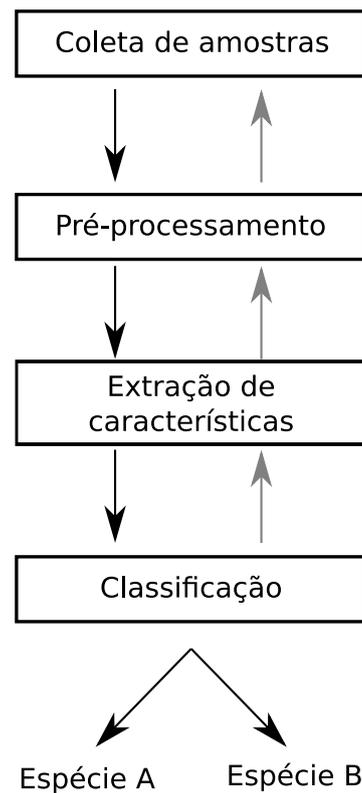
Pode-se supor uma fábrica que deseja aumentar o ritmo de produção de peixe enlatado. Nessa situação hipotética, foi percebido que um dos maiores gargalos no fluxo de produção é a separação entre as duas espécies de peixe que são descarregadas juntas na fábrica. O objetivo, então, é automatizar a classificação entre a Espécie A e a Espécie B.

As duas espécies tem claras diferenças físicas, como tamanho, peso, posição da boca, formato das barbatanas etc.. Por meio dessas características, ou *features*, pode-se desenvolver um modelo para cada grupo. Modelo este que é o conjunto de descrições diferentes (normalmente expressas na forma matemática) que representa as espécies de maneira genérica. É importante que o modelo seja genérico para englobar todos os indivíduos do grupo.

O protótipo do sistema para classificação na fábrica consiste no fluxograma representado pela Figura ???. O primeiro passo foi fotografar exemplares de cada espécie para gerar um conjunto de amostras. Os sinais capturados pela câmera fotográfica são então

processados para simplificar as próximas operações sem perder informações relevantes. Nesse caso, um processo de simplificação seria a segmentação para retirar o fundo da imagem e deixar somente o peixe. Depois, as informações simplificadas são encaminhadas para um extrator de *features*. Esse passo tem como objetivo separar as informações relevantes. Por fim, os dados são encaminhados para o classificador que sinaliza o resultado para aquela amostra.

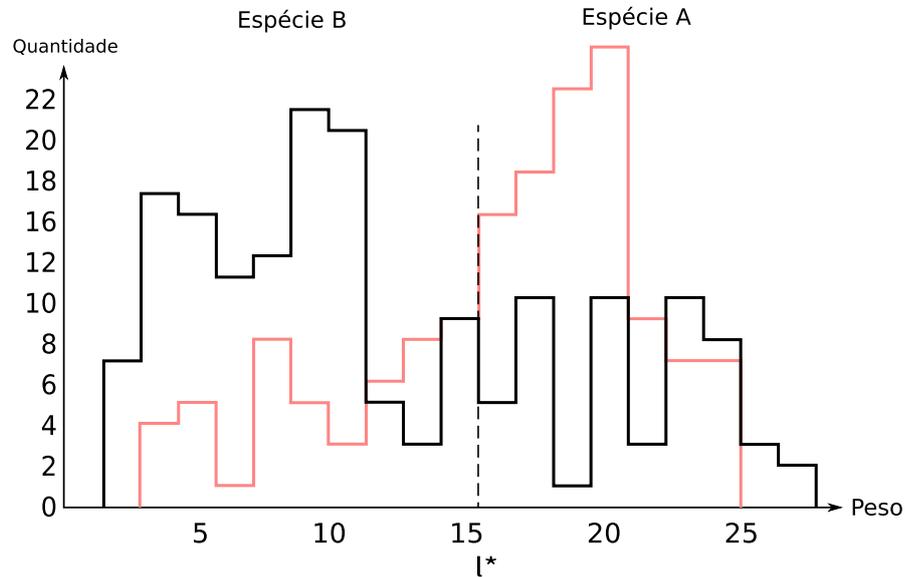
captionFluxograma que representa o funcionamento do protótipo do classificador usado



para diferenciar as espécies de peixe.

A primeira *feature* percebida foi o tamanho do peixe. A média de tamanho entre os peixes da Espécie A é maior que a Espécie B. Sendo assim, pode-se desenvolver um modelo onde a *feature* que diferencia as duas espécies é o peso. Para isso, peixes das duas espécies foram pesados e o conjunto de amostras pode ser visto na Figura 1. O objetivo então é encontrar um valor l que marca o ponto de separação entre os dois grupos de peixe.

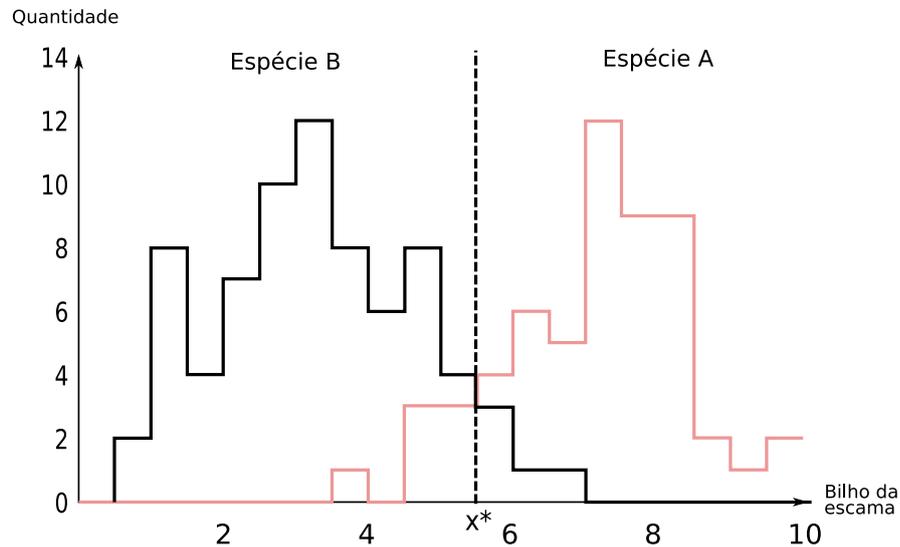
Figura 1 – Gráfico que relaciona as amostras de cada espécie de acordo com o peso (DUDA; HART; STORK, 2012).



Pela Figura 1 perceber-se que a *feature* peso não produz resultado satisfatório, pois o melhor valor possível para l ainda gera bastante ambiguidade.

Testa-se então o brilho das escamas (desta vez usando a câmera fotográfica proposta no protótipo) como critério de classificação. Supondo que uma quantidade razoável de amostras foi fotografada, segmentada e avaliada (matematicamente), tem-se o gráfico representado na Figura 2.

Figura 2 – Gráfico que relaciona as amostras de cada espécie de acordo com o brilho da escama.



Mais uma vez o resultado não é satisfatório, pois não foi encontrado um delimitador que minimize a ambiguidade. Nesse contexto, é necessário ressaltar a importância do custo. Imagine que a Espécie B tenha um valor de mercado muito elevado em relação à Espécie A. Um cliente que recebesse por engano uma lata com um peixe da Espécie B talvez não ficaria frustrado, mas o contrário é bastante improvável. Assim, a necessidade de classificar os peixes corretamente está diretamente ligada ao custo para a fábrica.

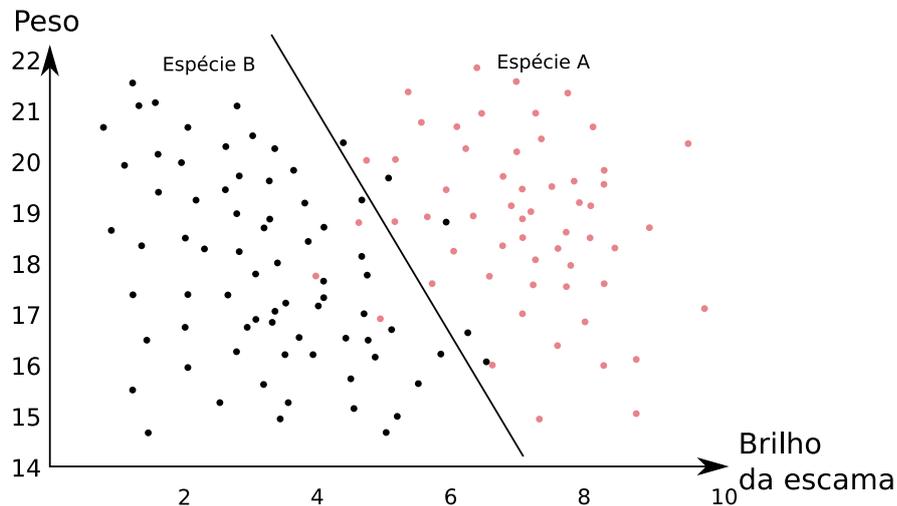
Agora pode se pensar no modelo como a combinação de mais de uma *feature*. Se as duas *features* anteriores forem combinadas: peso como x_1 e brilho das escamas como x_2 , o modelo pode ser dado como um vetor X num espaço de *features* com duas dimensões, como,

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (1)$$

Se as amostras forem colocadas em um gráfico que relaciona as características x_1 e x_2 , pode se encontrar um resultado como o representado pela Figura 3. Pode-se observar que os dois grupos estão separados de uma maneira muito mais satisfatória que as das Figuras 1 e 2. Isso mostra que o modelo que combina as duas *features* gera um classificador melhor. Nesse ponto, pode parecer que quanto mais características forem reunidas em um único modelo, maior a qualidade do classificador desenvolvido, mas esta relação é errônea. Quanto menos complexo for o modelo, melhor. Isso se dá pelo fato de que modelos complexos tendem a viciar o classificador no conjunto que foi usado para treinamento. Desta forma, o classificador não se comporta bem com dados novos, o que

não satisfaz o objetivo do projeto.

Figura 3 – Gráfico que relaciona duas *features* peso e brilho das escamas, para as espécies A e B.



A qualidade das *features* escolhidas está diretamente ligada à generalização do modelo. Normalmente a melhor abordagem é procurar por características fortes dos objetos a serem classificados. Depois, separar as mais relevantes e descartar principalmente as *features* redundantes para manter o modelo simples e o mais generalizado possível.

2.2 Rede neural artificial

Uma Rede Neural Artificial (RNA) é um modelo preditivo inspirado na maneira como o cérebro funciona (GRUS, 2015). O conceito de RNA surgiu na neurobiologia e a estratégia é modelar um neurônio para compor uma rede de neurônios interligados que trocam informações e processam os dados submetidos para alcançar um resultado específico.

Essa categoria possui grande variedade de aplicações. Dentre elas, estão o agrupamento de indivíduos com características semelhantes (também chamado de *clustering*), reconhecimento de padrões (que é o foco deste trabalho), aproximação de funções e previsão.

Figura 4 – Conjunto de dados que podem ser agrupados em três categorias diferentes. Um exemplo de *clustering*.

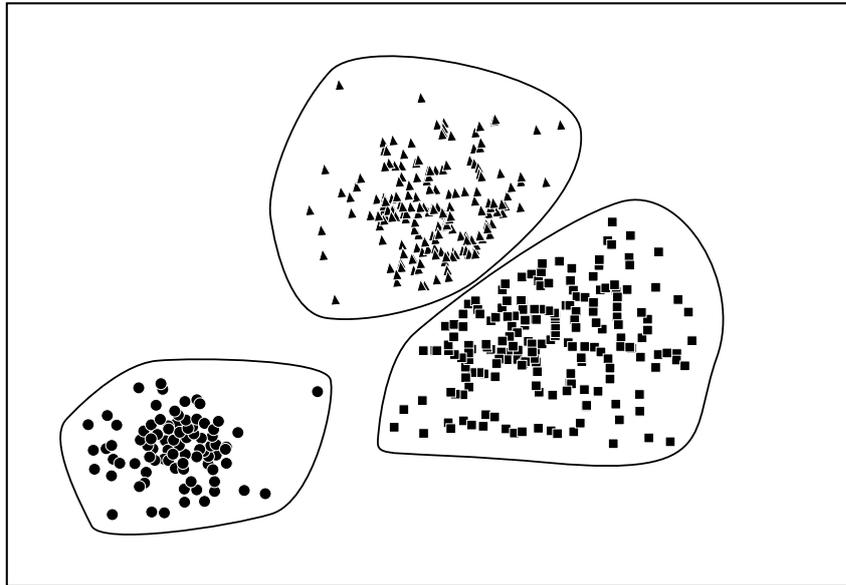
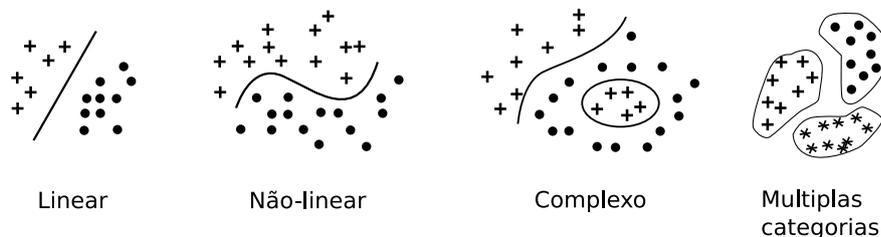


Figura 5 – Exemplos de problemas que envolvem reconhecimento de padrões e classificação. Traduzido de Samarasinghe (2016)



A Figura 4 mostra um exemplo de problema do tipo *clustering*. RNAs são uma boa escolha para encontrar critérios de agrupamento de informações naturalmente semelhantes, cujo critérios de aproximação não são conhecidos. Esse ramo da Inteligência Artificial também é eficiente em reconhecimento de padrões usando limites não-lineares de classificação, como ilustrado na Figura 5. RNAs também são muito eficazes para ajustar modelos em um conjunto de dados. Isso se dá pela eficiência em encontrar modelos complexos e não lineares para informações n-dimensionais. É por esse motivo que elas são chamadas de aproximadoras universais (SMITH, 1993). A Figura 6 ilustra a função ajustada para um conjunto de dados qualquer. Outra área de destaque é na previsão, onde os próximos passos são previstos de acordo com o histórico. Para isso, o modelo é a síntese dos padrões formados através de eventos passados. Um exemplo dessa área é a previsão do tempo,

onde o conhecimento sobre o passado define o comportamento futuro. A Figura 7 contém um gráfico que relaciona a frequência de chuva em relação ao tempo.

Figura 6 – Exemplo de pontos gerados em um plano qualquer, onde se encontra uma função parecida com a realidade do conjunto.

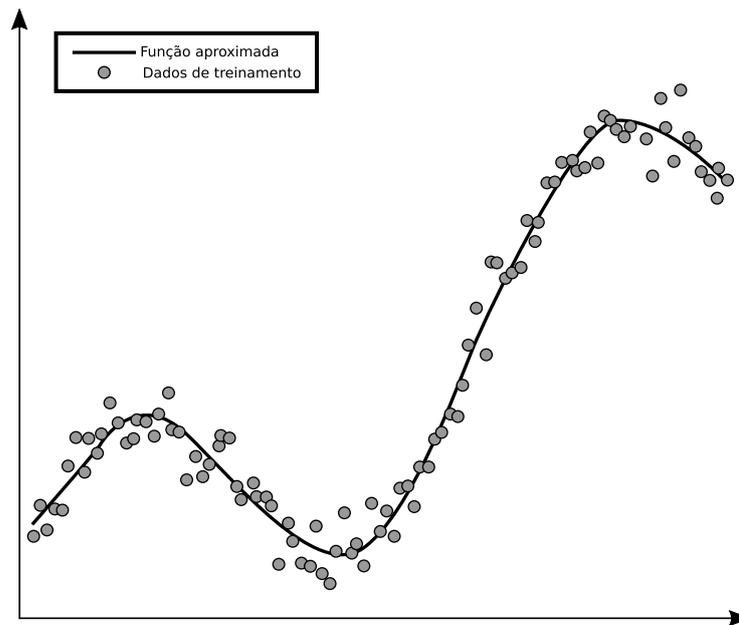
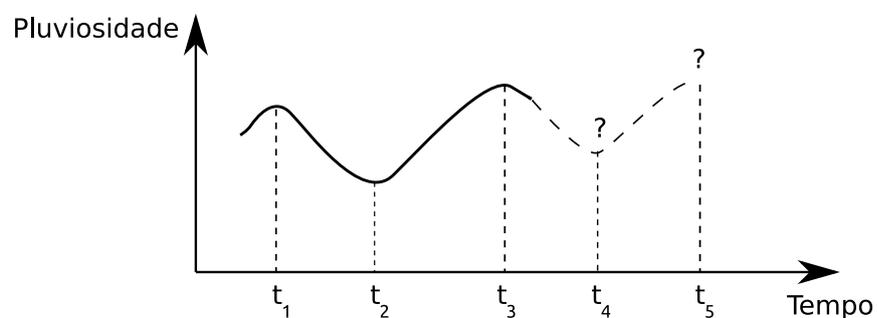


Figura 7 – Exemplo de pontos previstos em uma função que relaciona a quantidade de ocorrências de chuva por um período de tempo. Traduzido de Samarasinghe (2016).



Um tipo de RNA bastante usado como exemplo introdutório é o *Perceptron* (GÉRON, 2017). Isso se dá pela simplicidade de sua arquitetura. Ele foi criado por Frank

Rosenblatt em 1957, baseado em princípios biológicos sobre aprendizado. O *Perceptron* se baseia em um neurônio chamado de Unidade de Limiarização Linear (ou *linear threshold unit (LTU)* em inglês). Nessa unidade, números são usados como entrada/saída e cada entrada é associada a um peso. Então, a LTU calcula a soma das entradas, influenciadas por seus respectivos pesos, e processa a informação referente à soma, retornando o resultado como saída.

A maneira mais comum de processamento é pelas funções de limiarização *heaviside step function* e *sign*, que podem ser definidas como,

$$\text{heaviside}(z) = \begin{cases} 0 & \text{se } z < 0 \\ 1 & \text{se } z \geq 0 \end{cases} \quad (2)$$

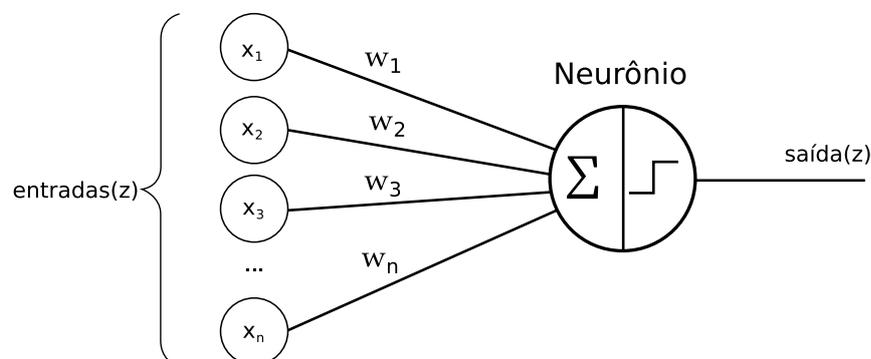
e

$$\text{sign}(z) = \begin{cases} -1 & \text{se } z < 0 \\ 0 & \text{se } z = 0 \\ +1 & \text{se } z > 0 \end{cases} \quad (3)$$

respectivamente.

O *Perceptron* pode ser usado como um classificador binário. Para isso é feita uma combinação linear das entradas, como ilustrada na Figura 8. Quando o resultado dessa combinação ultrapassa um limiar (*threshold*) estabelecido, a saída é disparada como “positiva”. Caso contrário, é disparada como “negativa”.

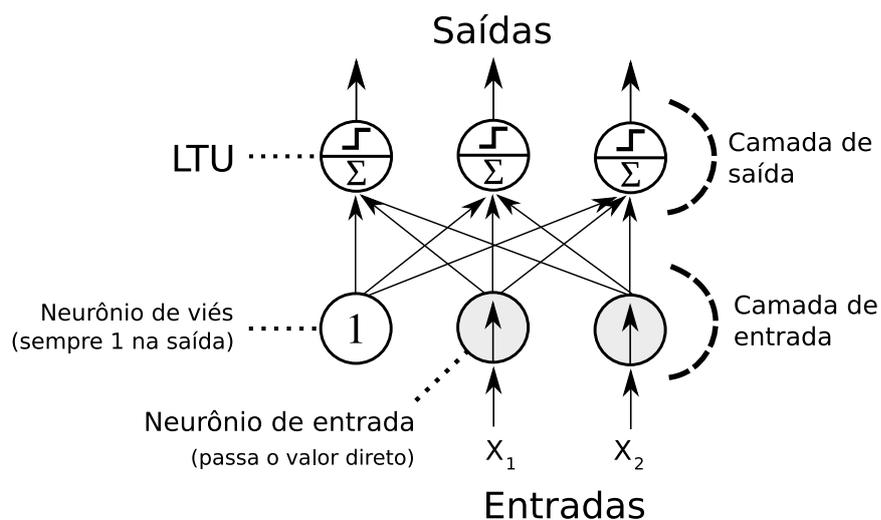
Figura 8 – *Perceptron* formado com apenas um neurônio.



A composição do *Perceptron* é tida como uma única camada de LTUs, com todos os neurônios ligados à todas as entradas, como ilustrado no diagrama da Figura 9. Essa

ligação é normalmente representada pelo uso de neurônios que não fazem nada além de emitir como saída o que recebem como entrada. Além disso, um neurônio extra é posto na entrada com saída constante em um (1). Esse neurônio é chamado de neurônio de viés (ou *bias neuron* em inglês).

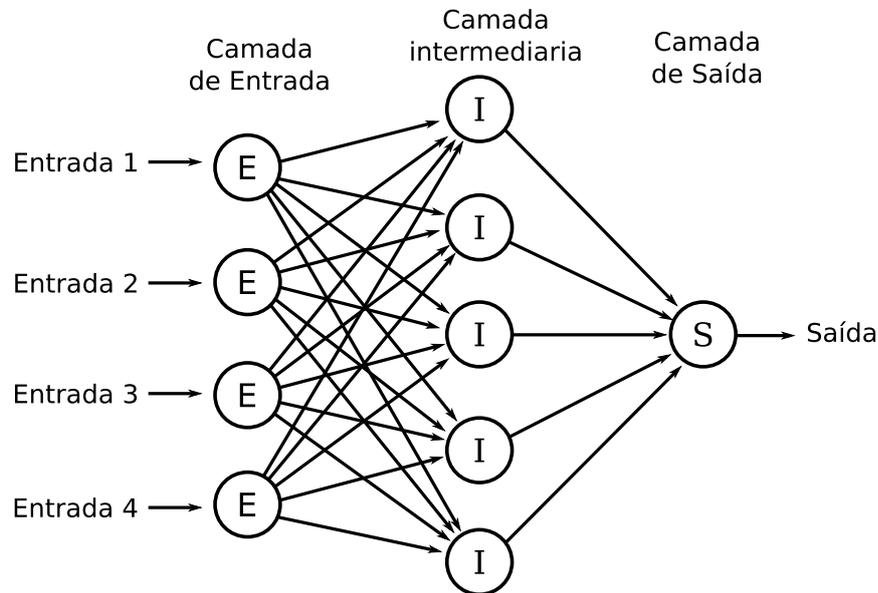
Figura 9 – Diagrama de um perceptron (GÉRON, 2017)



2.3 Aprendizado profundo (*Deep learning*)

Um tipo de RNA bastante utilizado é a *Multilayer Perceptron* (MLP), que consiste em conjuntos de neurônios formando camadas, onde a saída de uma camada serve como entrada para outra. A Figura 10 ilustra uma MLP simples, com uma camada de entrada (representada pela letra E), uma camada intermediária (representada pela letra I) e uma camada de saída (representada pela letra S). Esse tipo de RNA é capaz de mapear sistemas de entrada e saída complexos e é uma rede de neurônios bastante flexível, que se adapta livremente às informações apresentadas.

Figura 10 – Diagrama de uma Rede Neural do tipo *Multilayer perceptron* com três camadas: entrada, intermediária e saída.



A ligação entre os neurônios é chamada de “peso sináptico”. Ela é conhecida como “parâmetro livre”, pois, seus valores se adaptam durante o treinamento e é através dessa adaptação que a rede aprende sobre informações de entrada.

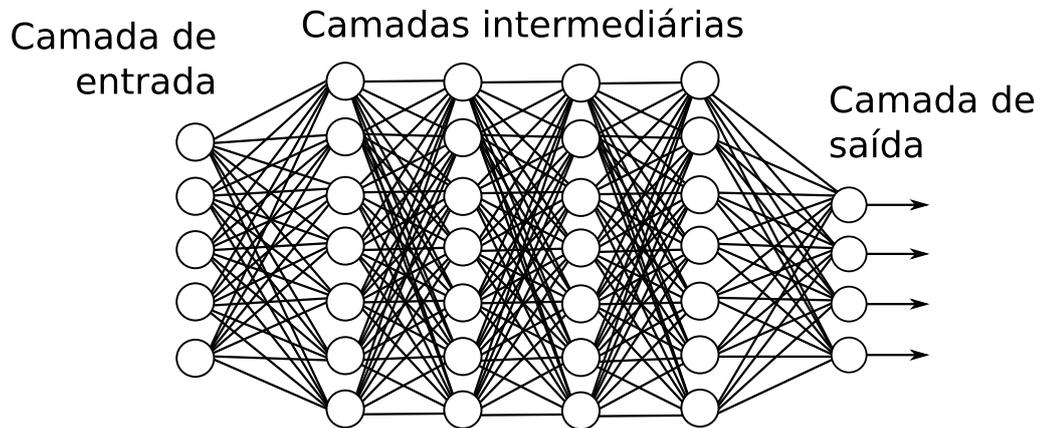
Samarasinghe (2016) descreve que o funcionamento da *Multilayer Perceptron* em três etapas, como:

- A primeira camada transmite a informação da entrada para a camada intermediária (também chamada de “camada escondida”) através do peso entrada-intermediária (ligação que está entre os neurônios de cada camada);
- A camada intermediária processa a informação e envia para a camada de saída pelo peso intermediária-saída;
- A camada de saída processa os dados e traduz o resultado final.

A rede é treinada ao ter como entrada várias amostras identificadas. Assim, o aprendizado se dá pela mudança de relação entre as camadas através do peso, até que a rede passe a identificar os novos dados. Por fim, os pesos finais são os parâmetros otimizados da rede.

Quando uma RNA tem duas ou mais camadas intermediárias, ela é chamada de Rede Neural de Aprendizado Profundo. A imagem da Figura 11 ilustra o diagrama de uma DNN.

Figura 11 – Diagrama de uma Rede Neural de Aprendizado Profundo



3 TRABALHOS RELACIONADOS

Há pouco material que explora o reconhecimento através do som dos passos. A maior parte dos trabalhos relacionados se baseia em dispositivos e sensores para capturar informações dos passos, como é o caso de Orr e Abowd (2000). Este trabalho propõe um sistema de identificação usando uma espécie de ladrilho com sensores de pressão para colher os dados sobre o indivíduo. Nele, Orr e Abowd ressaltam que seu *Smart Floor* (ou “Piso Inteligente” em tradução livre) tem vantagens em relação à sistemas que usam biometria para reconhecimento. A principal delas é que o usuário não precisa mudar seu comportamento para que os dados sejam lidos. Eles também dizem no artigo que conseguiram atingir 93% de sucesso na identificação durante os testes.

O artigo Shoji, Takasuka e Yasukawa (2004) tem objetivos bastante semelhantes aos propostos por este trabalho. Shoji, Takasuka e Yasukawa fizeram um sistema para identificar indivíduos através do som de seus passos. Foi gravado o som de cinco pessoas caminhando individualmente, dez vezes cada uma. Todos os participantes usaram o mesmo tipo de calçado, eram do mesmo sexo e tinham a mesma faixa de idade (22 a 24 anos). Sendo assim, a amostra é razoavelmente pequena e isto implica no peso que os testes representam. No artigo, Shoji, Takasuka e Yasukawa falam que as características retiradas foram: Análise mel-cepestral, intervalo de caminhada e grau de similaridade do envelope espectral. Essas características passaram por um processo de refinamento cujo objetivo era eliminar ruídos e realçar suas principais propriedades para depois serem processadas pelo algoritmo *K-Means*. Os resultados da identificação não foram muito bons. Os autores dizem que é preciso aumentar o número de amostras e trabalhar melhor a extração de características e a diminuição do ruído.

O trabalho Rosa et al. (2004) é mais voltado para o tema Casa Inteligente. Ele discorre sobre conceitos relacionados ao projeto da casa, como os atuadores, os sensores e as áreas de atuação. Rosa et al. defendem que o objetivo de tornar uma casa em uma casa inteligente é aumentar o conforto e reduzir gastos com, principalmente, água e energia. Isso se alcança identificando os indivíduos que habitam o local e aprendendo seu comportamento. O objetivo desta pesquisa é justamente propor um método de identificação para um ambiente inteligente como o proposto em Rosa et al. (2004).

Assim como Orr e Abowd (2000), o trabalho Savio e Ludwig (2007) propõe um piso equipado com sensores para coletar dados das pessoas que caminham no ambiente. Savio e Ludwig afirmam que o carpete proposto tem mais usabilidade e custo menor que o piso do trabalho citado anteriormente. Eles também defendem que a identificação através dos passos contribui mais para criar um sistema pervasivo, pois a sensação de estar sendo monitorado é menor. Sendo assim, o sistema proposto tem como objetivo

identificar o indivíduo e estimar sua localização no ambiente. Ambiente este, que pode ser tanto doméstico quanto corporativo e a sua aplicação pode ser tanto para conforto e otimização energética, quanto para vigilância. O problema é que mais uma vez a solução depende de um produto que, na prática, representa um custo de aplicação mais elevado.

O artigo Carvalho e Rosa (2010) é a base do trabalho proposto. Nele, Carvalho e Rosa realizaram um experimento (que será descrito no Capítulo 4) com 15 indivíduos caminhando sobre um palco de madeira. Eles gravaram o som do caminhar de cada indivíduo, filtraram os dados coletados, prepararam as características e as processaram com os métodos *K-Means* e *K-NN*. Os dados provenientes deste experimento são os que serão processados no trabalho proposto. No artigo, vê-se uma justificativa mais detalhada sobre cada um das cinco características extraídas das gravações. Características essas que são: frequência dos passos, envelope Espectral, coeficientes cepstrais, coeficientes mel-cepstrais e sonoridade. Os resultados da classificação variaram entre 33.5% à 97.5%.

Em Singh et al. (2017), Singh et al. propõem o uso dos passos como característica de identificação em uma rede de aprendizado profunda. Para isso, foram usados sensores de pressão em uma malha, apresentada no trabalho Zhou et al. (2017), para colher as informações referentes às pegadas. A maior proposta do artigo é a mudança de domínio das informações para serem processadas. Apesar da malha capturar informações referentes à pressão dos passos, Singh et al. transformaram as informações em imagens. Dessa forma, cada ponto na malha é representado como um pixel na imagem e quanto maior a pressão aplicada ali, mais claro é o ponto. Sendo assim, toda a base de dados foi primeiramente transformada em matrizes com valores normalizados entre 0 e 255, para depois ser usada no treinamento da rede de aprendizado profundo. Por fim, os resultados do trabalho variam entre 71,99% e 87,66% de precisão, sendo o melhor valor obtido através de uma rede neural convolucional profunda, usando toda a sequência de passos.

4 METODOLOGIA

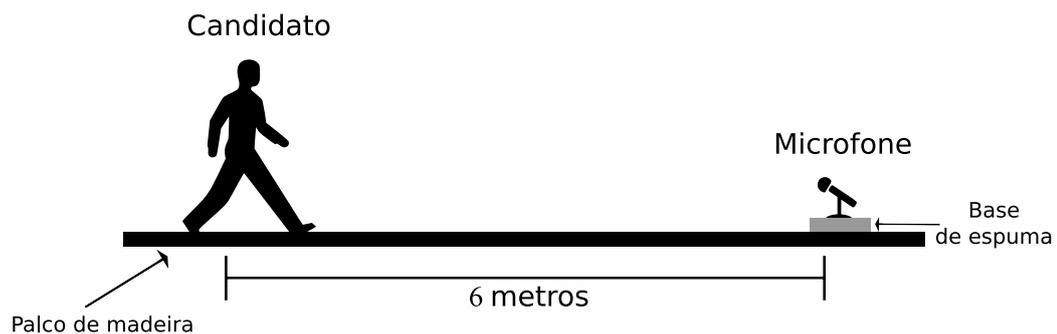
Neste capítulo são descritos os passos seguidos pelo experimento realizado no trabalho prévio de Carvalho e Rosa, assim como o processo de extração dos dados provenientes desse experimento. Após, a arquitetura da rede utilizada no trabalho é explicada, além das ferramentas e *frameworks* utilizados para o seu desenvolvimento.

4.1 Experimento

Este trabalho se baseia no artigo de Carvalho e Rosa (2009), onde Carvalho e Rosa colheram as amostras de quinze indivíduos de gêneros distintos, utilizando seis tipos de calçados diferentes. Estas mesmas amostras serão utilizadas para treinar e testar a rede de aprendizado profundo proposta neste trabalho.

O experimento foi realizado em uma sala fechada, onde os indivíduos caminharam por um palco de madeira de 6m x 1m (como pode ser conferido na Figura 12). Em uma extremidade do palco, foi fixado um microfone modelo padrão de mercado da marca *Leadership*, conectado a uma placa de som ADI ADI1986A *SoundMAX 6-channel* de alta definição. O som foi capturado em uma taxa de amostragem de 44 KHz com 16 *bits* de quantização. As amostras foram separadas em seis subconjuntos, de acordo com a variedade de calçados, como pode ser visto na Tabela 1.

Figura 12 – Configuração do experimento realizado por Carvalho e Rosa.



Carvalho e Rosa perceberam que os indivíduos caminhavam de modo mais natural na metade do trajeto. Isso se dá pelo fato da aceleração no início e da desaceleração no final do percurso. Sendo assim, cada sequência gravada foi editada para permanecerem somente quatro estímulos referentes ao meio do percurso.

Depois de editadas, as sequências passaram pela aplicação do critério de *Fisher* (YANG; YANG, 2003) para eleição das partes mais discriminantes. Durante a seleção de características, os valores 0,05, 0,1, 0,2 e 1 foram estabelecidos como percentual P,

Tabela 1 – Subconjuntos das amostras colhidas no experimento de Carvalho e Rosa.

Nome	Quantidade de indivíduos no subconjunto
Coturno	três homens e uma mulher
Tênis	oito homens
Sapato	seis homens
Calçados distintos	duas mulheres e quatro homens
Tênis e Sapato	seis homens

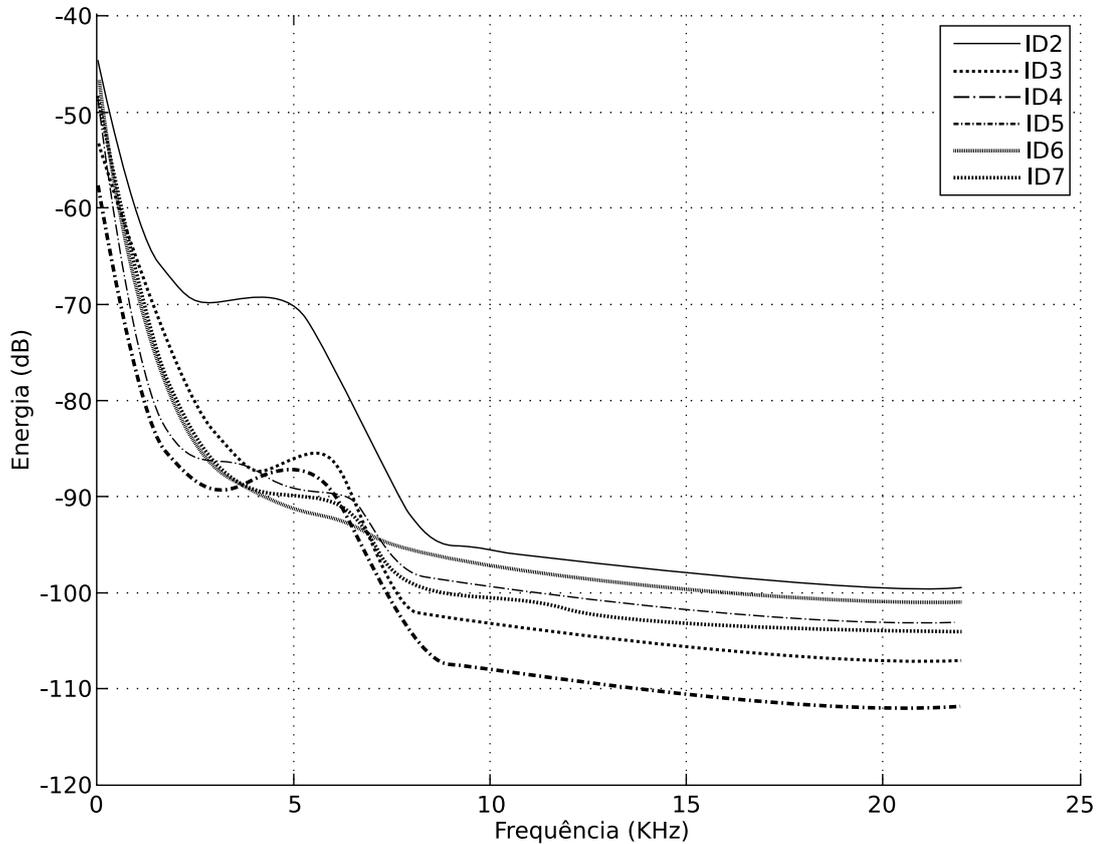
correspondentes às proporções dos elementos utilizados de cada característica. Assim, após a aplicação do critério de *Fisher*, os dados ficaram aptos para serem usados como entrada para a rede neural.

As *features* retiradas no processo descrito anteriormente são: frequência dos passos, envelope espectral, coeficientes cepestrais, coeficientes mel-cepestrais e sonoridade. Cada uma delas passou por um processo de filtragem para deixar a informação mais relevante. A Tabela 2 mostra a relação dos nomes citados anteriormente com as classes que armazenam os vetores de características. O gráfico da Figura 13 mostra uma comparação entre os dados referentes ao envelope espectral de seis indivíduos. Todos os seis indivíduos estavam usando o mesmo tipo de calçado. Isso então demonstra que realmente há uma variação bastante discriminante no caminhar dos indivíduos e se espera que esse conjunto de *features* será suficiente para classificação dos participantes do experimento.

Tabela 2 – Relação entre o nome das características e as classes extraídas no experimento, cujo termo é usado ao decorrer do trabalho para especificar a *feature*.

Característica	Nome da <i>feature</i>
Sonoridade	<i>loudnesszwicker</i>
Coeficiente cepestral	<i>lpcCeps</i>
Coeficientes mel-cepestral	<i>mfccmalcom</i>
Coeficientes mel-cepestral 2	<i>mfccmalcom2</i>
Envelope espectral (Periodograma)	<i>envwang</i>
Envelope espectral (Peridograma de Hamming)	<i>envwang2</i>
Envelope espectral (Método de Welch)	<i>envoliver</i>

Figura 13 – Gráfico que compara a *feature* Envelope Espectral de 6 indivíduos diferentes usando o mesmo calçado.



4.2 Desenvolvimento da rede neural

A biblioteca escolhida para se desenvolver a rede de aprendizado profundo foi a TensorFlow¹. Ela se trata de uma biblioteca originalmente desenvolvida pela equipe Google Brain² e hoje é um projeto de código aberto, cujo objetivo é alcançar alta performance em computação numérica.

Foi utilizada a Interface de Programação de Aplicações (API) Keras³ para implementar o código da rede neural. Através dela foi possível testar diferentes configurações de RNAP para tentar alcançar o melhor resultado.

O passo inicial foi estudar como as *features* extraídas no trabalho de Carvalho e Rosa estavam estruturadas. Assim, pôde-se notar que todos os arquivos seguiam o padrão

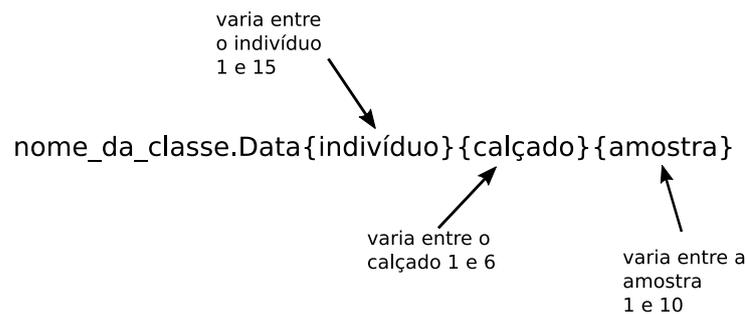
¹TensorFlow. About TensorFlow. Disponível em <<https://www.tensorflow.org/about/>> Acesso em 29 de outubro de 2018

²Google Brain. Bringing the benefits of AI to everyone. Disponível em <<https://ai.google/about/>> Acesso em 29 de outubro de 2018

³Keras. Keras: The Python Deep Learning library. Disponível em <<https://keras.io/>> Acesso em 29 de outubro de 2018

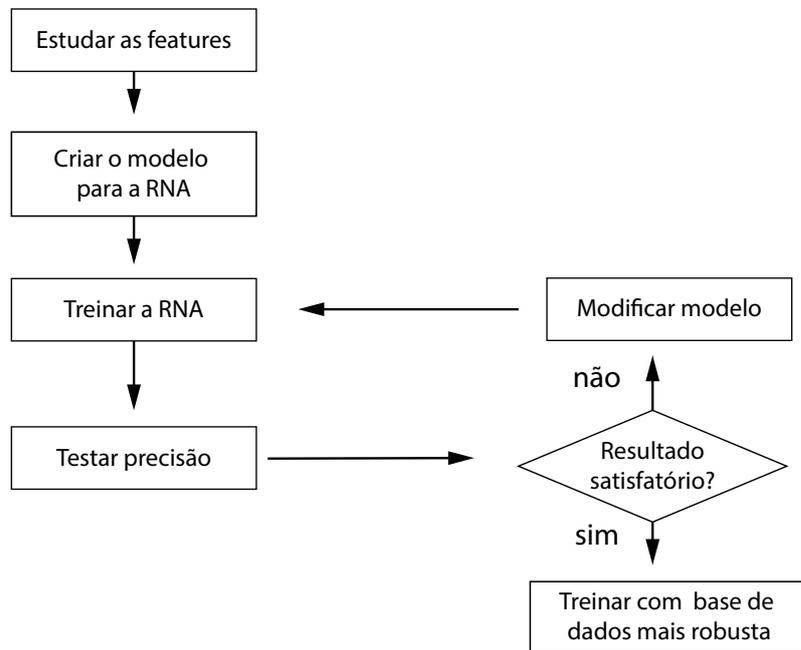
de encapsulamento representado na Figura 14. O tipo final do dado que representa a amostra pode variar entre *Array*, *Matriz* e *Cells*. Depois, cada uma das amostras passou pelo processo de conversão para o formato *csv*, assim como a geração do conjunto de saídas da rede. Por fim, os dados foram normalizados para serem usados como entrada na RNAP.

Figura 14 – Padrão de encapsulamento das características extraídas em Carvalho e Rosa (2010)



Com o conjunto de características, o próximo passo foi modelar a classe que representa o som dos passos de maneira abrangente. Depois, o modelo foi testado através do treinamento e avaliação da rede, utilizando $\frac{4}{5}$ dos dados colhidos no experimento. Isso faz com que o restante ($\frac{1}{5}$ dos dados) seja destinado à fase de testes. Feito isso, a rede então foi testada *feature a feature* para encontrar quais são as mais discriminantes e depois combiná-las. Por fim, a rede passou por um treinamento mais robusto para aumentar a sua precisão. Neste momento é importante ressaltar que o treinamento não pode ser longo a ponto de induzir a erros por vales no espaço de busca da RNA no conjunto de dados usados como entrada. Este roteiro segue o fluxograma apresentado na Figura 15.

Figura 15 – Fluxograma que representa o processo de desenvolvimento da Rede Neural Artificial.



O processo de aprendizado se resume em:

1. atribuir valores aleatórios para os pesos da RNA;
2. calcular o quanto o resultado atual está distante do resultado esperado;
3. calcular o valor de correção, que modifica os pesos para aproximar o resultado atual do resultado esperado;
4. ajustar os pesos da RNA de modo a aproximar o resultado atual do esperado;
5. voltar ao passo 2 e repetir uma quantidade n suficientemente grande para aumentar a precisão da RNA.

5 RESULTADOS

Este Capítulo contém o detalhamento das arquiteturas desenvolvidas e seus respectivos resultados. Um teste inicial foi realizado e, com base nele, outras arquiteturas de redes neurais foram modeladas e aplicadas.

5.1 Descrição do experimento

Com as *features* devidamente tratadas, cada uma foi usada para treinar individualmente uma DNN de arquitetura simples. A rede foi formada a partir de duas camadas ocultas de tamanhos 240 e 120 neurônios cada. A *feature* que demonstrou melhores resultados foi a *loudnesszwick*: 73,8% de acurácia logo no primeiro teste. Depois de algumas execuções com mais gerações, esta mesma característica alcançou 93,33% de precisão na base de teste. A Tabela 3 mostra os resultados para cada *feature* nesse modelo básico de DNN.

Tabela 3 – Resultados obtidos através da primeira rede testada com duas camadas intermediárias de formato 240x120x15 neurônios

Resultados das <i>features</i> na rede básica	
<i>Feature</i>	Acurácia
<i>loudnesszwick</i>	93,33%
<i>lpcCeps</i>	88,09%
<i>mfccmalcom2</i>	87,66%
<i>mfccmalcom</i>	78,76%
<i>envwang2</i>	9,52%
<i>envoliver</i>	4,76%
<i>envwang</i>	4,76%

As Figuras 16 e 17 mostram os gráficos de acurácia e perda durante o treinamento das *features loudnesszwick* e *lpcCeps*, respectivamente. Em cada geração, os pesos da rede são atualizados e se mede quantos resultados, em porcentagem, foram previstos corretamente. Essa medida converge de maneira promissora na *feature loudnesszwick*, como pode ser visto no gráfico da esquerda da Figura 16. Já a perda nos mostra a distância do resultado previsto para o resultado real. No gráfico da direita na Figura 16 pode ser visto que a perda chega a zero quando a base de treino alcança 100% de acurácia, mas mostra resultados ruins na base de teste à medida que as gerações vão passando. Já na Figura 17, pode ser visto que a evolução durante o treinamento para *lpcCeps* não alcança uma acurácia tão alta quanto a *feature* anterior, mas o gráfico da perda sugere que o modelo é melhor.

Figura 16 – Acurácia e perda durante o treinamento da *feature loudnesszwicker*.

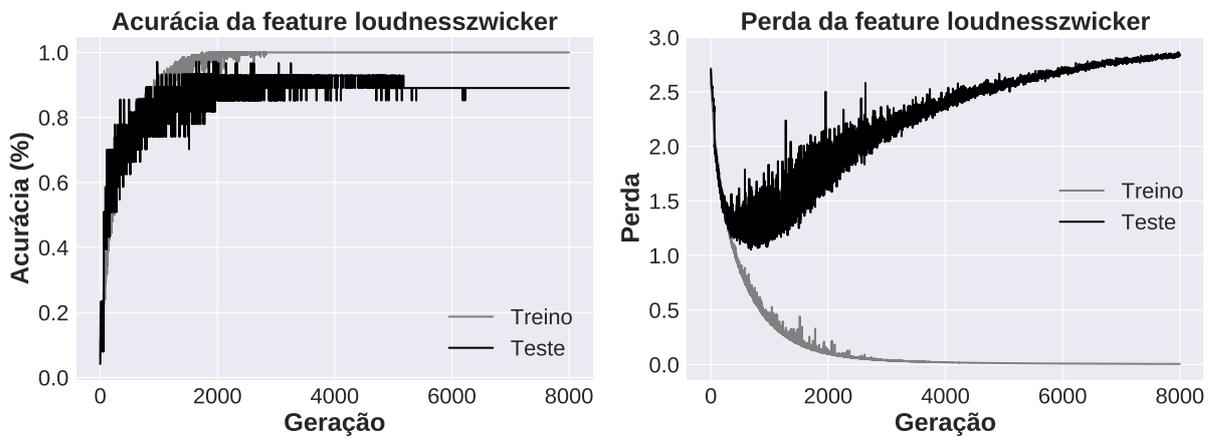
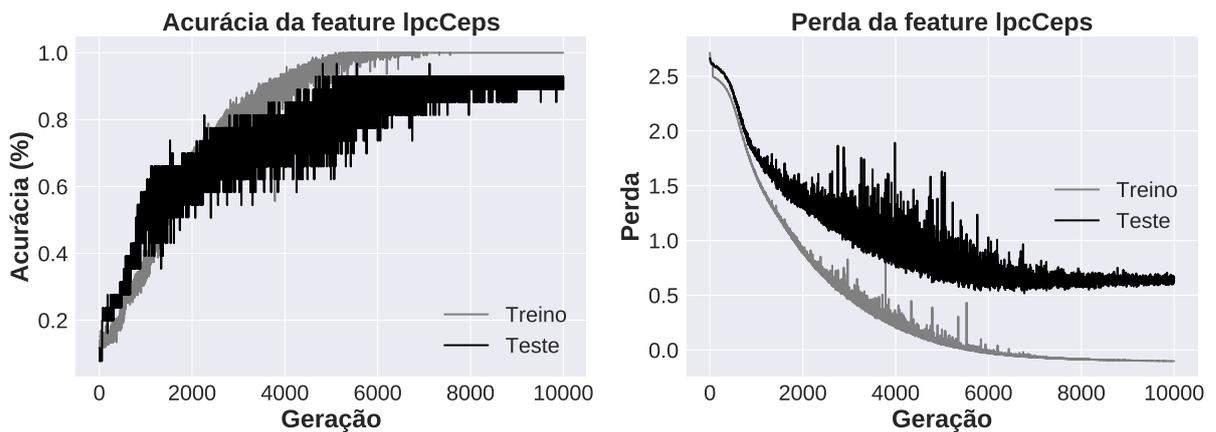
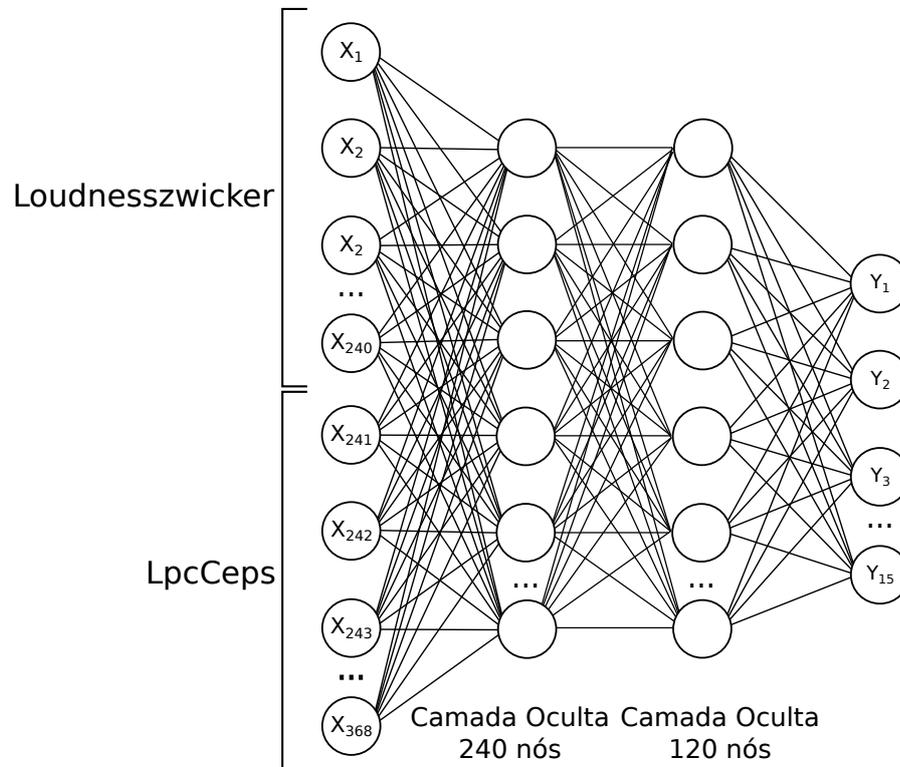


Figura 17 – Acurácia e perda durante o treinamento da *feature lpcCeps*.



Visto que as *features* que convergiram melhor foram a *loudnesszwicker* e *lpcCeps*, pode-se concluir que ambas as características se mostraram determinantes na caracterização dos indivíduos. Assim, um modelo foi desenvolvido combinando-as. É importante frisar que todas os conjuntos tem a mesma quantidade de amostras, o que possibilita a sua combinação. A Figura 18 representa graficamente o modelo que continua com a mesma configuração de camadas intermediárias. Como pode ser visto na parte esquerda da Figura 18, os conjuntos de características de ambas as *features* são combinados amostra por amostra, resultando em um vetor de 364 posições. A rede foi instanciada 50 vezes com pesos aleatórios, mas não atingiu resultados tão bons quanto os apresentados anteriormente. A acurácia máxima alcançada foi de 90,9%, contra os 93,33%, conforme é apresentado na Tabela 3.

Figura 18 – RNA que combina as *features loudnesszwicker* e *lpcCeps* em um único vetor de características

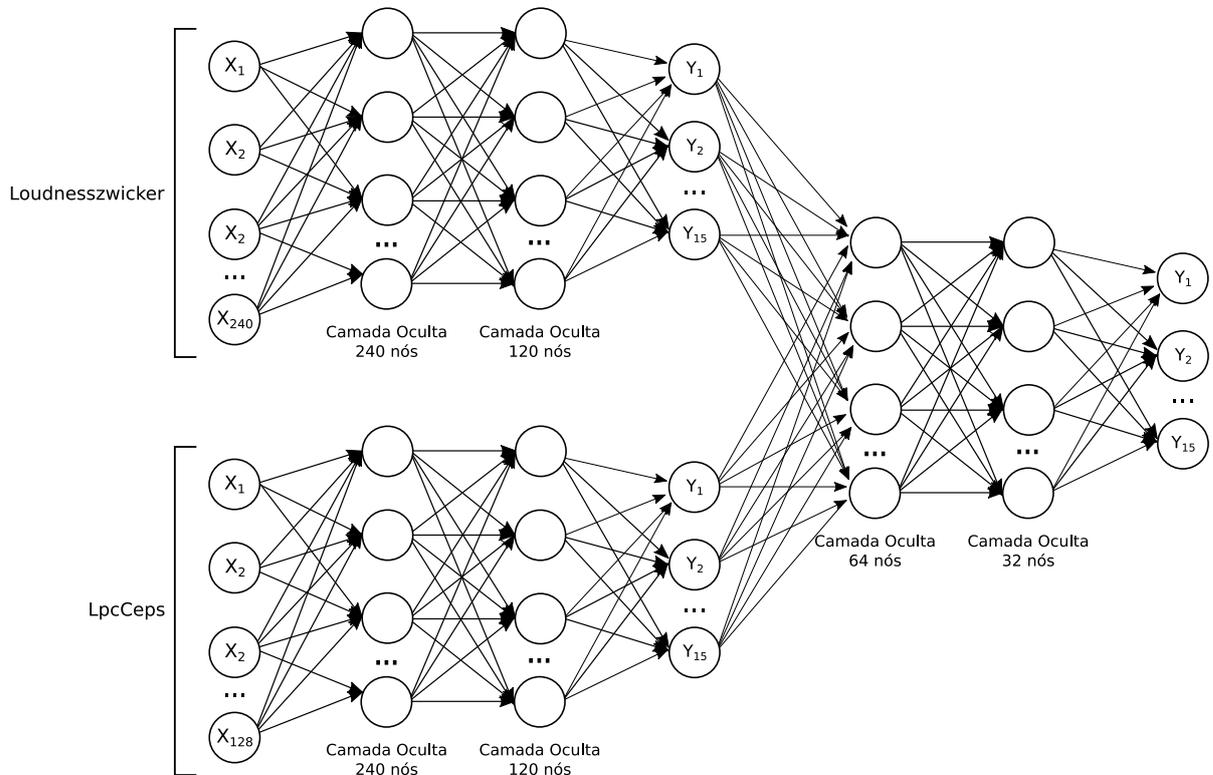


5.2 Uma arquitetura multirede

A combinação entre os vetores de entrada das *features loudnesszwicker* e *lpcCeps* não alcançou resultados melhores do que ambas individualmente. Desta forma, buscou-se explorar um modelo que mantém as duas redes neurais de formato 240x120x15 e combina suas saídas em uma terceira rede, como pode ser conferido na Figura 19. Nessa nova arquitetura, a saída dos 15 neurônios de cada uma das redes serve de entrada para uma terceira rede neural mais simples de formato 64x32x15, como pode ser visto à direita da imagem. Para o treinamento desta última, a base de dados é usada para fazer previsões nas duas redes anteriores, gerando os dois conjuntos de saída que são direcionados à nova rede. Por fim, esta é treinada com o resultado das previsões anteriores e os mesmos rótulos utilizados anteriormente.

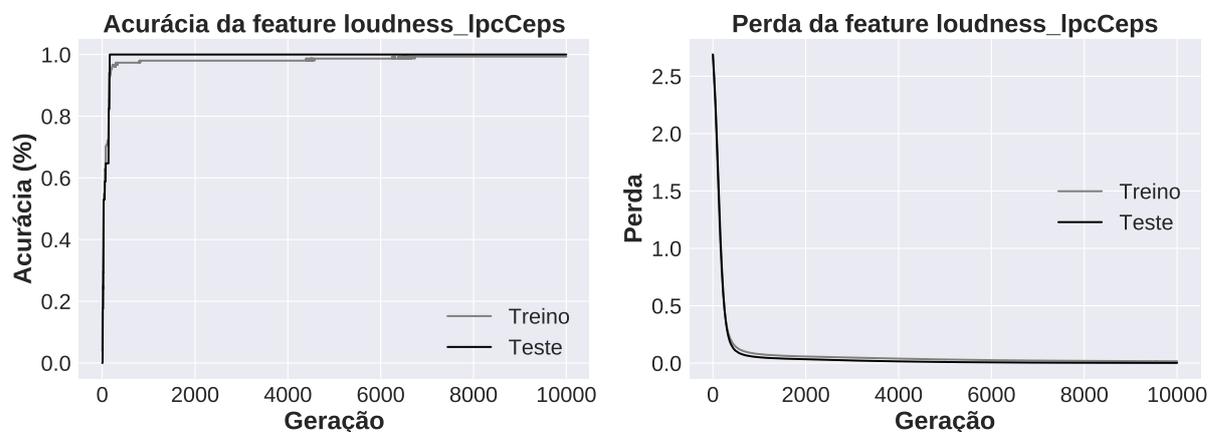
Como pode ser visto na Figura 20, a convergência da terceira DNN rapidamente converge para 100% de acurácia na base de treino. As linhas se aproximam tanto que dificulta a análise, mas pode-se confirmar que o desempenho da rede é satisfatório. Assim, a linha mais escura do gráfico da direita representa a acurácia na base de teste. Pode ser observado que os resultados na base de teste se mantêm melhores que a base de treino até a geração 7000. O gráfico que ilustra a evolução da perda durante o treinamento indica que a relação das previsões dadas pelas duas DNN's usadas como entrada são um bom

Figura 19 – Arquitetura multiredes que combina as saídas das RNA's anteriores em uma nova rede.



modelo. Essa confirmação era esperada, dado os resultados promissores das duas redes individualmente.

Figura 20 – Acurácia e perda durante o treinamento da arquitetura multiredes que combina três redes neurais, conforme a Figura 19.



Com a arquitetura apresentada na Figura 19, a acurácia do sistema alcançou 98,57% na base de testes. Um resultado consideravelmente satisfatório para o problema

de classificação em questão.

6 CONCLUSÃO

O presente trabalho abordou o problema de identificar pessoas através do som dos passos, empregando uma abordagem de aprendizado profundo. Para tanto, foi apresentada uma visão geral sobre o assunto que envolve classificação e redes neurais. Diante disso, buscou-se apresentar o funcionamento e as principais características de uma DNN. Em relação ao estado da arte, pôde-se perceber que os trabalhos se direcionam para a classificação por visão computacional, deixando muito a ser estudado na subárea de classificação através de som.

A base de dados utilizada foi a mesma apresentada em Carvalho e Rosa, a qual contém amostras de 15 indivíduos, uns com mais de um tipo de calçado, totalizando 210 amostras ao todo. Cada amostra sonora foi processada gerando um conjunto de 10 macro *features*, a saber: *envoliver*, *envwang*, *envwang2*, *gaitFreqmaxvalue*, *gaitFreqmaxvalue2*, *gaitFreqzwick*, *loudnesszwick*, *lpcCeps*, *mfccmalcom* e *mfccmalcom2*. Um dos percalces a serem superados foi o entendimento das *features*. Visto a dinamicidade em seus tamanhos e significados. A *loudnesszwick* por exemplo, consistiu em um vetor de 240 valores. Enquanto que a *mfccmalcom*, para cada onda sonora, gerou-se uma matriz de 20x227.

A abordagem proposta inicialmente consistiu em avaliar um modelo rede profunda com arquitetura 240x120x15 e verificar seu comportamento diante de cada uma das características de som. As *features* que resultaram em melhores classificações, de acordo com a medida de acurácia, incluem a *loudnesszwick*, com 93,33%; a *lpcCeps*, com 88,09% e a *mfccmalcom2*, com 87,66%. Deste estudo, foi percebido que as características *envwang*, *envwang2*, *envoliver* não permitiram a rede convergir ou resultaram em baixo desempenho, incluindo 78,76% para a *feature mfccmalcom*.

Uma vez identificadas as características com melhor desempenho, um último teste foi realizado com uma arquitetura combinada. Esta arquitetura combinada consistiu de um encadeamento de DNN's no formato 240x120x15, sendo uma para a *feature loudnesszwick*, outra para a *feature lpcCeps* e uma terceira que tem como entrada as saídas das duas anteriores, gerando, ao final do processo, uma acurácia de 98,57%. Com isso, concluiu-se que a presente proposta de trabalho obteve êxito, alcançando os objetivos estipulados no Capítulo 1.

Diante do exposto, sabe-se que as investigações científicas geralmente respondem a algumas perguntas, porém com o conhecimento adquirido ao longo do desenvolvimento faz emergir tantas outras. Brevemente, vislumbra-se como etapas para continuação desta pesquisa os seguintes itens:

1. avaliar novas arquiteturas e de novas combinações dos conjuntos de características;

2. aplicar a técnica de Análise de Componentes Principais na tentativa de minimizar a quantidade de características.

REFERÊNCIAS

- ARISTÓTELES. **On the Gait of Animals**. [S.l.: s.n.], 350 A.C.
- CARVALHO, R.; ROSA, P. Sistema de identificação para casa inteligente usando som. **IX Simpósio Brasileiro de Automação Inteligente, Brasília/DF**, 2009.
- CARVALHO, R. L. de; ROSA, P. F. F. Identification system for smart homes using footstep sounds. In: IEEE. **Industrial Electronics (ISIE), 2010 IEEE International Symposium on**. [S.l.], 2010. p. 1639–1644.
- COLLOTTA, M.; PAU, G. A solution based on bluetooth low energy for smart home energy management. **Energies**, Multidisciplinary Digital Publishing Institute, v. 8, n. 10, p. 11916–11938, 2015.
- CONNOR, P.; ROSS, A. Biometric recognition by gait: A survey of modalities and features. **Computer Vision and Image Understanding**, Elsevier, v. 167, p. 1–27, 2018.
- DING, D. et al. Sensor technology for smart homes. **Maturitas**, Elsevier, v. 69, n. 2, p. 131–136, 2011.
- DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern classification**. [S.l.]: John Wiley & Sons, 2012.
- GÉRON, A. **Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems**. [S.l.]: "O'Reilly Media, Inc.", 2017.
- GRUS, J. **Data science from scratch: first principles with python**. [S.l.]: "O'Reilly Media, Inc.", 2015.
- KHAN, M. S. et al. An unsupervised acoustic fall detection system using source separation for sound interference suppression. **Signal processing**, Elsevier, v. 110, p. 199–210, 2015.
- LARSEN, B. W. et al. Applying matching pursuit decomposition time-frequency processing to ugs footstep classification. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense XII**. [S.l.], 2013. v. 8711, p. 871104.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, Springer, v. 5, n. 4, p. 115–133, 1943.
- ORR, R. J.; ABOWD, G. D. The smart floor: A mechanism for natural user identification and tracking. In: ACM. **CHI'00 extended abstracts on Human factors in computing systems**. [S.l.], 2000. p. 275–276.

- ROSA, P. F. et al. A pervasive identification and adaptation system for the smart house. In: SPRINGER. **IFIP International Conference on Artificial Intelligence Applications and Innovations**. [S.l.], 2004. p. 141–154.
- SAMARASINGHE, S. **Neural networks for applied sciences and engineering: from fundamentals to complex pattern recognition**. [S.l.]: CRC Press, 2016.
- SAVIO, D.; LUDWIG, T. Smart carpet: A footstep tracking interface. In: IEEE. **Advanced Information Networking and Applications Workshops, 2007, AINAW'07. 21st International Conference on**. [S.l.], 2007. v. 2, p. 754–760.
- SHOJI, Y.; TAKASUKA, T.; YASUKAWA, H. Personal identification using footstep detection. In: IEEE. **Intelligent Signal Processing and Communication Systems, 2004. ISPACS 2004. Proceedings of 2004 International Symposium on**. [S.l.], 2004. p. 43–47.
- SINGH, M. S. et al. Transforming sensor data to the image domain for deep learning—an application to footstep detection. In: IEEE. **Neural Networks (IJCNN), 2017 International Joint Conference on**. [S.l.], 2017. p. 2665–2672.
- SMITH, M. **Neural networks for statistical modeling**. [S.l.]: Thomson Learning, 1993.
- YANG, J.; YANG, J.-y. Why can lda be performed in pca transformed space? **Pattern recognition**, Elsevier, v. 36, n. 2, p. 563–566, 2003.
- ZHOU, B. et al. The carpet knows: Identifying people in a smart environment from a single step. In: IEEE. **Pervasive Computing and Communications Workshops (PerCom Workshops), 2017 IEEE International Conference on**. [S.l.], 2017. p. 527–532.