



Programa Interdisciplinar de Pós-Graduação em
Computação Aplicada

Mestrado Acadêmico

Márcio Moura Leal

**SingApp: Um Modelo de Identificação de Língua de Sinais
Através de Captura de Movimento em Tempo Real**

São Leopoldo
2018

Márcio Moura Leal

**SINGAPP: UM MODELO DE IDENTIFICAÇÃO DE LÍNGUA DE SINAIS ATRAVÉS
DE CAPTURA DE MOVIMENTO EM TEMPO REAL**

Dissertação apresentada como requisito parcial
para a obtenção do título de Mestre pelo
Programa de Pós-Graduação em Computação
Aplicada da Universidade do Vale do Rio dos
Sinos — UNISINOS

Orientadora:
Prof^ª. Dr^ª. Marta B. Villamil

São Leopoldo
2018

L435s Leal, Márcio Moura.
SingApp: um modelo de identificação de língua de sinais através de captura de movimento em tempo real / por Márcio Moura Leal. -- São Leopoldo, 2018.

61 f. : il. (algumas color.) ; 30 cm.

Dissertação (mestrado) – Universidade do Vale do Rio dos Sinos, Programa Interdisciplinar de Pós-Graduação em Computação Aplicada, São Leopoldo, RS, 2018.

Orientação: Prof^a. Dr^a. Marta B. Villamil, Escola Politécnica.

1.Redes neurais (Computação). 2.Inteligência artificial.
3.Sistemas de reconhecimento de padrões. 4.Reconhecimento de padrões. 5.Visão por computador. 6.Detectores infravermelhos.
7.Língua de sinais. 8.Surdos – Meios de comunicação. I.Villamil, Marta B. . II.Título.

CDU 004.855.5
004.93'1:800.954

Catálogo na publicação:
Bibliotecária Carla Maria Goulart de Moraes – CRB 10/1252

Márcio Moura Leal

SingApp: Um Modelo de Identificação de Língua de Sinais Através de Captura de Movimento em Tempo Real

Dissertação apresentada à Universidade do Vale do Rio dos Sinos - Unisinos, como requisito parcial para obtenção do título de Mestre em Computação Aplicada.

Aprovado em 06/04/2018

BANCA EXAMINADORA

Prof. Dr. Rafael Piccin Torchelsen - UFPel

Prof. Dr. Sandro José Rigo - UNISINOS

Prof^ª. Dr^ª. Marta Becker Villamil (Orientadora)

Visto e permitido a impressão
São Leopoldo,

Prof. Dr. Rodrigo da Rosa Righi
Coordenador do PPG em Computação Aplicada

Dedico à minha avó, Nelsi, que partiu durante a elaboração desta dissertação.

Para o alto e avante!
— SUPERMAN

AGRADECIMENTOS

À minha orientadora Prof^a. Dr^a. Marta Villamil, interlocutora sensível, sempre interessada, atenciosa e paciente, pelo acolhimento e incentivo ao longo do percurso trilhado.

Aos professores que do PIPCA pelas aulas descontraídas, mas principalmente pelo conhecimento transmitido.

À minha família pelo suporte, se fazendo presente em todos os momentos da minha vida.

Ao Allan pelas dicas super importantes. Ao Anderson e Luiz que também fizeram parte deste projeto. Aos demais amigos e colegas que, nas conversas de corredor, foram importantes na elaboração desta dissertação.

Ao PIPCA, por me proporcionar várias trocas de bombonas d'água, possibilitando-me a inclusão de mais uma habilidade em meu currículo Lattes :)

*“Do you wish me a good morning,
or mean that it is a good morning
whether I want it or not;
or that you feel good this morning;
or that it is a morning to be good on?”.*
(Gandalf)

RESUMO

O reconhecimento das línguas de sinais visa permitir uma maior inserção social e digital das pessoas surdas através da interpretação da sua língua pelo computador. Este trabalho apresenta um modelo de reconhecimento de dois dos parâmetros globais das línguas de sinais, as configurações de mão e seus movimentos. Através da utilização de uma tecnologia de captura de infravermelho, a estrutura da mão foi reconstruída em um espaço tridimensional virtual e a Rede Neural Perceptron Multicamadas foi usada para fazer a classificação das configurações de mão e de seus movimentos. Além do método de reconhecimento de sinais, este trabalho visa disponibilizar um conjunto de dados representativos das condições do cotidiano, constituído por uma base de dados de configurações de mão e de captura de movimento validadas por profissionais fluentes em línguas de sinais. Foi usada como estudo de caso a Língua Brasileira de Sinais, a Libras, e obteve-se como resultados uma precisão de 99.8% e 86.7% de acertos das redes neurais que classificavam as configurações de mão e seus movimentos, respectivamente.

Palavras-chave: Infravermelho. Rede Neural Artificial. Língua de Sinais.

ABSTRACT

The sign language recognition aims to allow a greater social and digital insertion of deaf people through interpretation of your language by the computer. This work presents a recognition model of two global parameters of the sign languages, hand configurations and their movements. Through the usage of infrared capture technology we built the hand structure on a virtual three-dimensional space and the Multilayer Perceptron Neural Network was used to do the hand configuration and movements classifying. Beyond of method to recognize signs, this work aims to provide a set of representative data of the daily conditions, consisting of a database of hand configurations and motion capture validated by fluent professionals in sign languages. To this work, was used, as study case, the Brazilian Sign Language, Libras, and was obtained accuracy rates of 99.8% and 86.7% from neural networks classifying hand configurations and hand motion respectively.

Keywords: Infrared. Artificial Neural Network. Sign Language.

LISTA DE FIGURAS

Figura 1 – O <i>hardware</i> do aparelho	28
Figura 2 – A área de captura do aparelho	29
Figura 3 – O sistema de coordenadas cartesianas construído a partir do aparelho	30
Figura 4 – Exemplos de estruturas de dados entregues pelo Leap Motion.	31
Figura 5 – A representação gráfica de um neurônio artificial	32
Figura 6 – Rede Perceptron Multicamadas	33
Figura 7 – Aprendizado supervisionado	34
Figura 8 – Alfabeto manual da Libras	36
Figura 9 – Conjunto de sinais selecionados	44
Figura 10 – Exemplos de dados entregues pelo Leap Motion	45
Figura 11 – RNPM classificadora das configurações de mão	46
Figura 12 – Dados de entrada para a 2 ^a RNA.	47
Figura 13 – RNPM classificadora dos movimentos da mão	48
Figura 14 – Classificação de sinais com a mesma configuração de mão.	53

LISTA DE TABELAS

Tabela 1 – Frequências do espectro eletromagnético	27
Tabela 2 – Trabalhos relacionados ao termo ‘ <i>Sign Language Recognition</i> ’.	37
Tabela 3 – Tabela comparativa dos trabalhos relacionados	41
Tabela 4 – Correlação entre configurações de mão e sinais selecionados	45
Tabela 5 – Exemplos de dados usados na primeira RNA	46
Tabela 6 – Exemplo de dados usados no treinamento da segunda RNA	47
Tabela 7 – <i>Cross-Validation</i> das configurações de mão	51
Tabela 8 – Matriz de confusão das configurações de mão	52
Tabela 9 – <i>Cross-validation</i> do movimento	52
Tabela 10 – Matriz de confusão da classificação do movimento	52

LISTA DE SIGLAS

API	Application Programming Interface
ArSL	Língua Árabe de Sinais
ASL	Língua Americana de Sinais
DTW	<i>Dynamic Time Warping</i>
GHz	Gigahertz
HOG	Histograma de Gradientes Orientados
HSV	Matiz Saturação Valor
KSVM	Kernel Support Vector Machine
LED	Diodo Emissor de Luz
Libras	Língua Brasileira de Sinais
LMC	Leap Motion Controller
LSF	Língua de Sinais Francesa
MIZ	Momento Invariante de Zernique
MSL	Língua Mexicana de Sinais
IR	Infravermelho
INES	Instituto Nacional de Educação de Surdos
RGB	Vermelho Verde Azul
RGB-D	Vermelho Verde Azul Profundidade
RNA	Rede Neural Artificial
RNC	Rede Neural Convolutacional
RNPM	Rede Neural Perceptron Multicamadas
RGB	Vermelho Verde Azul
THz	Terahertz
WEKA	Waikato Environment for Knowledge Analysis
YIQ	Luminância Em-Fase Quadratura

SUMÁRIO

1	INTRODUÇÃO	23
1.1	Objetivos do trabalho	24
1.2	Contribuições do trabalho	24
1.3	Organização da dissertação	24
2	CONCEITOS BÁSICOS	27
2.1	O infravermelho	27
2.1.1	Infravermelho no espectro eletromagnético	27
2.1.2	O Leap Motion Controller	28
2.2	Redes neurais artificiais	28
2.2.1	A Rede Neural Perceptron Multicamadas	30
2.2.2	Validação de resultados de RNAs	32
2.3	A Libras	34
2.3.1	Histórico	34
2.3.2	Os parâmetros da Libras	35
3	TRABALHOS RELACIONADOS	37
3.1	Abordagens baseadas em dispositivos auxiliares	38
3.2	Abordagens baseadas em visão computacional	38
3.2.1	Utilização do Leap Motion Controller	39
3.3	Comparativo dos trabalhos relacionados	41
4	MÉTODO	43
4.1	Seleção dos sinais	43
4.2	Aquisição dos dados	43
4.3	Identificação das configurações de mão	44
4.4	Identificação do movimento	47
4.5	Tecnologias utilizadas	48
5	RESULTADOS	51
5.1	Classificação das configurações de mão	51
5.2	Classificação do movimento	51
5.3	Aplicação do modelo	52
6	CONCLUSÃO E TRABALHOS FUTUROS	55
	REFERÊNCIAS	57

1 INTRODUÇÃO

Segundo a Organização Mundial da Saúde, há cerca de 360 milhões de pessoas no mundo com perda auditiva que as incapacita de alguma maneira (World Health Organization, 2015). Apesar dos avanços no que diz respeito ao desenvolvimento de práticas de valorização e inclusão dos surdos (GUEDES, 2010; WITCHES, 2014), ainda há uma grande dificuldade de acesso por parte destes a serviços básicos (SOUZA; PORROZZI, 2009) e de lazer (FRANÇA et al., 2013).

Uma vez que esta inclusão do surdo na sociedade ouvinte ainda esbarra na falta de conhecimento do ouvinte a respeito das línguas de sinais (FLORES; BARBOSA; RIGO, 2012) ferramentas computacionais que ajudem na interação surdo-ouvinte trazem benesses a estas pessoas (RODRIGUES; ALVES, 2014).

Neste trabalho foi desenvolvido um reconhecimento de sinais das línguas de sinais afim de traduzi-los para a versão escrita da língua falada do país onde estas estão situadas. Diante da gama de línguas de sinais e línguas escritas existentes, tomou-se como campo para estudo de caso a Língua Brasileira de Sinais (Libras) e a versão escrita da Língua Portuguesa brasileira, e, ainda, diante da quantidade de parâmetros que as línguas de sinais utilizam, a presente abordagem representa o reconhecimento de uma parte destes parâmetros. Porém, esta abordagem pode ser tomada como ponto inicial para futuros trabalhos, na medida em que as técnicas empregadas, além da formalização de um conjunto sinais da Libras, podem ser estendidas e utilizadas em aplicações maiores e para outras línguas.

O presente trabalho relaciona-se ao reconhecimento de sinais das línguas de sinais, no entanto, este tipo de reconhecimento representa uma grande área da visão computacional e compreende uma grande quantidade de técnicas e métodos. Por isso, foi realizada uma revisão de uma parcela destes trabalhos, com ênfase naqueles voltados para as línguas de sinais.

Quanto a forma que a abordagem deste trabalho atua, esta se dá pelo reconhecimento de sinais provenientes da captura da estrutura da mão através de um aparelho específico para este objetivo, o *Leap Motion Controller* (LMC), portanto, tomou-se como parâmetros, a configuração das mãos e seus movimentos realizados. Outros parâmetros das línguas de sinais, como expressões faciais e corporais, não serão consideradas para este trabalho.

Outro ponto a ser ressaltado é que as línguas de sinais também compreendem um grande campo de estudo em termos linguísticos. Como o presente trabalho visa apresentar uma abordagem para o reconhecimento de sinais utilizando técnicas computacionais, tendo-se a Libras como campo de estudo, este restringiu-se a apresentar aspectos linguísticos relevantes para este reconhecimento. O modelo aqui proposto reconhece sinais provenientes da captura da estrutura e movimento das mãos. Outros parâmetros das línguas de sinais, como expressões faciais e corporais, não são consideradas neste trabalho.

Os estudos realizados a respeito das línguas de sinais restringiram-se ao conhecimento de sinais, seus parâmetros e ao entendimento de como o presente trabalho poderia ser aplicado na

prática, portanto, a sessão seguinte apresenta, com ênfase prática, os objetivos do trabalho.

1.1 Objetivos do trabalho

O presente trabalho tem como objetivo geral apresentar um modelo para o reconhecimento de sinais das línguas de sinais, assumindo como campo prático um subconjunto daqueles que fazem parte da Libras.

Esta abordagem apresenta objetivos específicos que vão da construção de uma base de dados até a realização de testes e validação dos resultados encontrados. Assim, os objetivos específicos são os seguintes:

- Revisar os principais métodos e trabalhos que visam reconhecer línguas de sinais;
- Analisar e selecionar os sinais pertencentes à Libras a serem reconhecidos;
- Criar uma base de dados de arquivos contendo informações de movimento e de configurações de mão capturados com o LMC;
- Analisar os resultados obtidos.

1.2 Contribuições do trabalho

O presente trabalho apresenta como contribuição principal **uma nova abordagem para reconhecimento de sinais que integra e analisa configurações de mãos e seus movimentos** para a tradução de sinais. Os trabalhos realizados até o momento buscaram traduzir sinais usando apenas a configuração de mão ou apenas o seu movimento. A integração proposta no modelo é fundamental para a tradução de língua de sinais automática visto que o significado de alguns sinais diferem entre si somente pela mudança do movimento realizado pelas mãos ou pela mudança de sua configuração.

Como contribuição secundária do trabalho realizado, identificou-se **a confecção de um novo dicionário digital de Libras** baseado em posições e movimentos de mão capturados. Os arquivos correspondem a sinais variados, utilizados no trabalho, validado por especialistas fluentes na língua. Salienta-se que não existe um dicionário digital de língua sinais que use captura de movimento na sua construção.

1.3 Organização da dissertação

Excluindo-se o capítulo de Introdução, o presente trabalho está subdividido em outros 5 capítulos, os quais são descritos a seguir.

O Capítulo 2 destina-se a revisão bibliográfica realizada para a construção da abordagem. São apresentadas as principais técnicas aplicadas no presente trabalho e uma breve explanação

a respeito das línguas de sinais.

O Capítulo 3 apresenta os principais trabalhos que tem como intuito fazer o reconhecimento e tradução das línguas de sinais.

O Capítulo 4, o método desenvolvido é apresentadas. Neste capítulo são apresentadas as etapas realizadas na construção do presente trabalho, justificando a opção pelas técnicas empregadas, desde o processo de aquisição dos dados de posição das mãos em um espaço tridimensional virtual até sua classificação como um sinal da Libras.

O Capítulo 5 apresenta os resultados obtidos com os experimentos realizados com o modelo desenvolvido;

Por fim, no Capítulo 6 são apresentadas as conclusões obtidas com o trabalho realizado, assim como as expectativas de trabalhos futuros que podem ser realizados.

2 CONCEITOS BÁSICOS

Como esta dissertação objetiva fazer a captura e a classificação das configurações das mãos e dos movimentos por elas realizados no espaço tridimensional através do infravermelho, faz-se necessária uma revisão a respeito desta técnica de captura de movimento e dos algoritmos de classificação de padrões.

Desta maneira, esta revisão da literatura foi dividida em 3 partes: a primeira relacionada ao infravermelho e ao LMC; a segunda parte a respeito de Redes Neurais; por fim, na última parte tem-se uma seção a respeito das línguas de sinais e da Libras.

2.1 O infravermelho

Esta seção tem por objetivo dar uma breve explanação a respeito do infravermelho, e como este pode ser utilizado na visão computacional através do LMC.

2.1.1 Infravermelho no espectro eletromagnético

A radiação infravermelha, ou simplesmente infravermelho (IR), é uma radiação eletromagnética com comprimentos de onda maiores do que aquelas da luz visível. Ela se estende desde o fim do espectro visível do vermelho em 700 nanômetros (nm) e frequência de 430 Terahertz (THz), a 1 milímetro (mm) e 300 Gigahertz (GHz) (LIEW, 2006). A relação do infravermelho com o restante do espectro eletromagnético é ilustrado na Tabela 1.

Tabela 1 – Frequências do espectro eletromagnético

Nome	Comprimento de onda (m)	Frequência (Hz)
Radiação Gama	< 0.01 nm	> 30 EHz
Raios X	0.01 nm – 10 nm	30 EHz – 30 PHz
Ultravioleta	10 nm – 400 nm	30 PHz – 790 THz
Luz Visível	400 nm – 700 nm	790 THz – 430 THz
Infravermelho	700 nm – 1 mm	430 THz – 300 GHz
Microondas	1 mm – 1 m	300 GHz – 300 MHz
Rádio	1 m – 100.000 km	300 MHz – 3 Hz

Fonte: (WEAST et al., 2011)

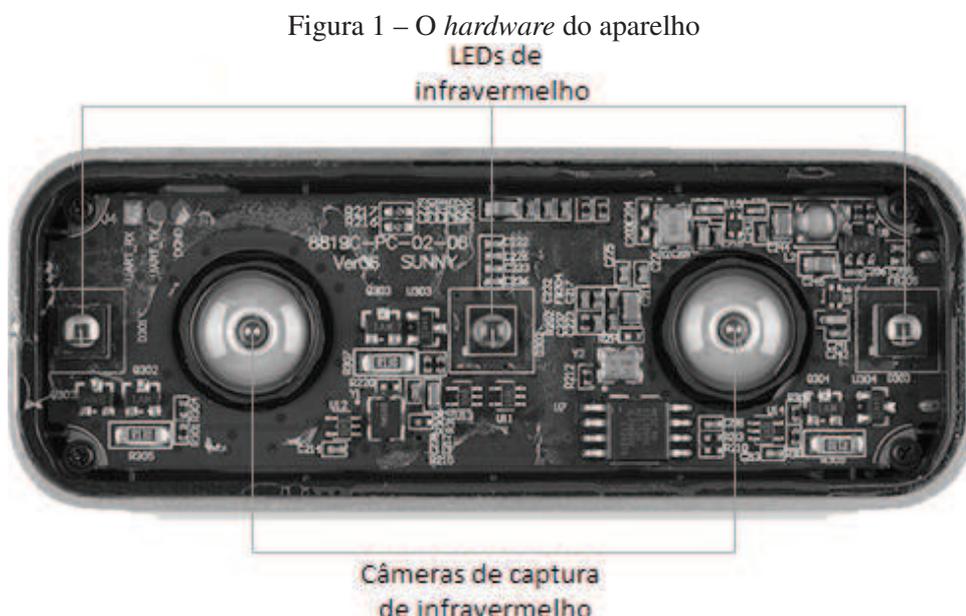
O infravermelho foi descoberto em 1800 pelo astrônomo Sir William Herschel, que descobriu um tipo invisível de radiação no espectro de energia menor que o da luz vermelha, por meio de seu efeito em um termômetro (ROWAN-ROBINSON, 2013).

A emissão de radiação infravermelha acontece por corpos que emitem calor (MCCREARY, 2004), portanto há uma gama de trabalhos que se utilizam desta característica. Sua aplicabilidade se estende pela indústria (OLIVEIRA et al., 2017), medicina (CUEVAS et al., 2017), entre outras. Os dispositivos de visão noturna, por exemplo, utilizam a ativação por proximidade do

infravermelho para permitir o rastreamento de pessoas ou animais (JIANG; DANIELL, 2004; HAN; SONG, 2016).

2.1.2 O Leap Motion Controller

O Leap Motion Controller é um sistema de reconhecimento e rastreamento de mãos (Leap Motion, 2014). O *hardware* do aparelho consiste de duas câmeras estereoscópicas e três diodos emissores de luz (LEDs) de infravermelho conforme ilustra a Figura 1. Trabalhar um espectro de luz tão específico proporciona a vantagem de filtrar objetos que recebam a luz nesse espectro, afinal, quanto menos tratamento a imagem precisar, mais rápido será o processamento.



Fonte: (COLGAN, 2015) (Editado pelo autor)

O dispositivo físico tem como objetivo unicamente fazer a captura das imagens em uma taxa de até 200 quadros por segundo em um campo de captura de 61 centímetros de raio. A Figura 2 ilustra a área de captura do aparelho.

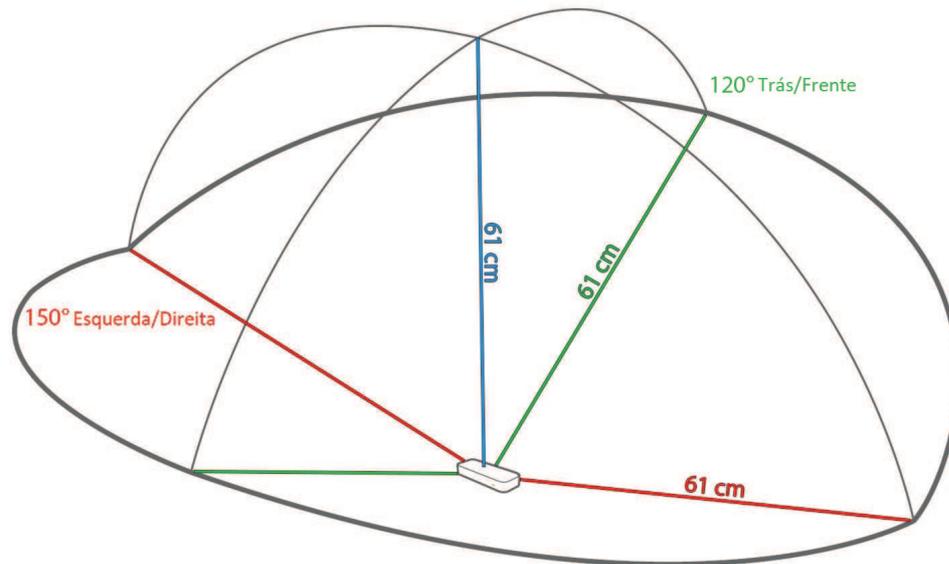
Uma vez que a imagem de infravermelho é capturada e transformada para escala de cinza, o *software* processa as imagens e constrói uma representação 3D das mãos, que estão em seu campo de captura, em um sistema de coordenadas cartesianas, onde a origem está centrada na parte superior do aparelho, tal sistema é ilustrado pela Figura 3.

O modelo virtual de mão fornece informações como posição e comprimento de cada osso da mão, vetor normal da palma, entre outras características, como ilustrado nas Figuras 4a e 4b.

2.2 Redes neurais artificiais

O estudo de redes neurais tem origem na análise de como o cérebro humano realiza com velocidade uma infinidade de tarefas complexas, como reconhecimento visual. O cérebro é

Figura 2 – A área de captura do aparelho



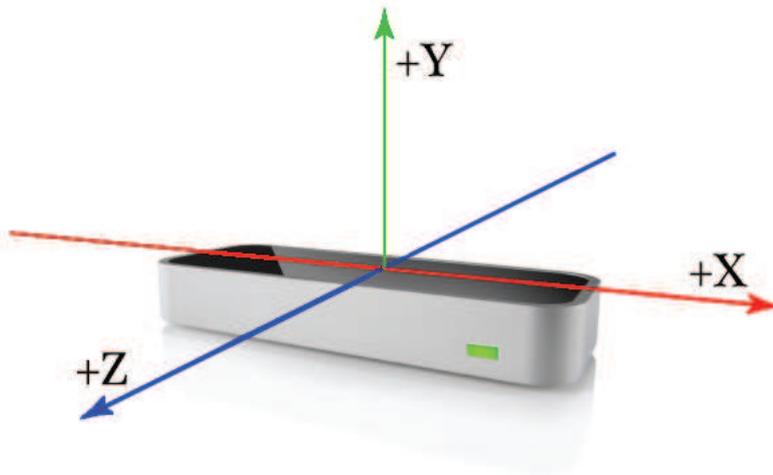
Fonte: (AUDY, 2013) (Editado pelo autor)

estruturado como um computador paralelo, complexo e não-linear, com capacidade de organizar suas estruturas básicas, os neurônios, para realizar os cálculos de forma rápida e paralela (HAYKIN, 2008). Cada um desses neurônios biológicos possuem complexidade e até mesmo velocidade similares a um microprocessador (DEMUTH et al., 2014).

As funções neurais biológicas são armazenadas nos neurônios e nas conexões entre eles. O processo de aprendizado é definido como a criação de novas conexões entre os neurônios e a modificação das conexões existente entre eles (DEMUTH et al., 2014). Baseado neste conhecimento foi modelado um conjunto de neurônios artificiais que agrupados constituem as chamadas redes neurais artificiais.

Um neurônio artificial é a unidade fundamental básica de uma rede neural artificial. Ele é definido por três elementos básicos: um conjunto de sinapses definidas como um peso, mais especificamente, um sinal de entrada x_j conectado ao neurônio k é multiplicado pelo peso da sinapse w_{kj} . O segundo elemento é um somador responsável pela adição do resultado da multiplicação dos sinais de entrada pelas sinapses do neurônio. Por fim, o neurônio possui uma função de ativação, que define a amplitude do sinal de saída a um valor finito (HAYKIN, 2008).

Figura 3 – O sistema de coordenadas cartesianas construído a partir do aparelho



Fonte: (Leap Motion, 2014)

Em termos matemáticos o neurônio pode ser descrito nas equações abaixo:

$$\sum_{j=1}^m w_{kj} x_j \quad (2.1a)$$

$$y_k = \varphi(u_k + b_k) \quad (2.1b)$$

onde x_1, x_2, \dots, x_m são os sinais de entrada, $w_{k1}, w_{k2}, \dots, w_{km}$ são os pesos sinápticos do neurônio k e b_k corresponde ao bias, responsável por realizar o deslocamento da função de ativação, definida por $\varphi(\cdot)$. A Figura 5 exemplifica de maneira gráfica a descrição de um neurônio artificial.

Para constituir uma rede neural, os neurônios são agrupados em estruturas denominadas camadas. Essas camadas são geralmente agrupadas definindo a profundidade da rede neural. O modo como essas camadas são agrupadas e a forma de aprendizado definem a arquitetura de uma rede neural artificial. Entre as várias arquiteturas de redes neurais, se destacam as Redes Neurais Perceptron e as Convolucionais.

2.2.1 A Rede Neural Perceptron Multicamadas

A rede neural do tipo perceptron é uma rede neural artificial utilizada em problemas de classificação e que apresenta sucesso mesmo na resolução de problemas complexos (HAYKIN, 2008).

Este classificador baseia-se no modelo de rede neural Perceptron de camada única proposta por (ROSENBLATT, 1960). Porém, devido as limitações do modelo de Rosenblatt na resolução de problemas não-lineares, esforços foram dirigidos para a aplicação das redes Perceptron Multicamadas e para o desenvolvimento de algoritmos de treinamento (ANDERSEN et al., 2001)

As redes Perceptron Multicamadas consistem em conjuntos de nós (unidades sensoriais)

Figura 4 – Exemplos de estruturas de dados entregues pelo Leap Motion.



Fonte: (Leap Motion, 2014)

dispostos em camadas da rede. Estas representam a camada de entrada, seguida por camadas intermediárias (ocultas), que por sua vez, são seguidas pela camada de saída da rede (HAYKIN, 2008). Assim, os sinais de entrada são propagados camada por camada até chegar à camada de saída, havendo conexões de uma camada somente com a camada imediatamente anterior e com a imediatamente posterior.

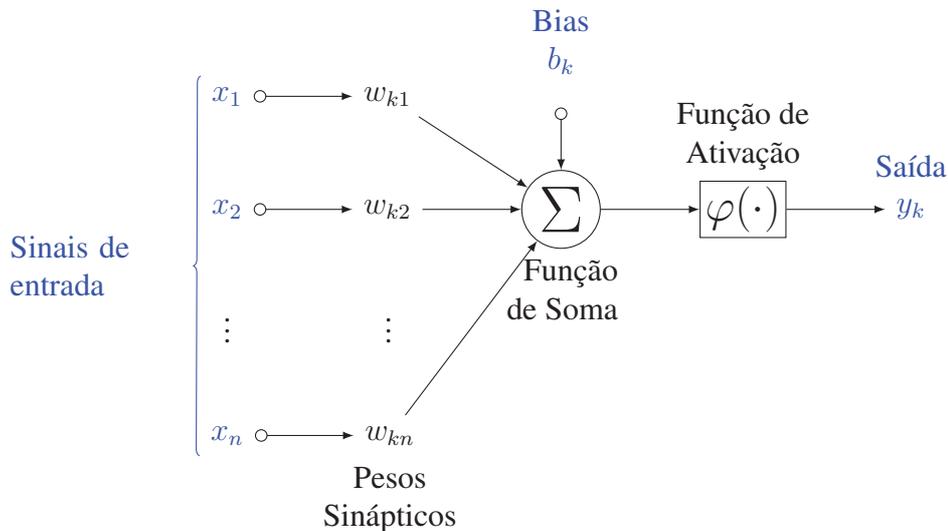
As camadas da rede Perceptron Multicamadas podem estar associadas a diferentes funções de ativação. A camada de entrada, por exemplo, está associada a uma função de ativação linear, enquanto nas camadas intermediárias, geralmente, usa-se funções não-lineares para que tenham a capacidade de distinguir padrões mais complexos (BASTOS, 2015).

A Figura 6 mostra uma típica rede Perceptron Multicamadas. Nota-se que ela apresenta apenas 3 camadas, dentre elas 1 única camada oculta. O sinal de entrada é representado pela variável N -dimensional x_p , assim como a saída é representada pela variável M -dimensional y_p . As variáveis w_{hi} e w_{oh} representam as matrizes de pesos que ponderam as conexões entre os neurônios. No caso da matriz w_{hi} pondera a conexão dos neurônios da camada de entrada com a camada intermediária e apresenta dimensões $N \times N_h$. Já a matriz w_{oh} , pondera as conexões entre os neurônios da camada intermediária e a camada de saída e apresenta dimensões $N_h \times M$.

O método mais comum para o treinamento do classificador Perceptron Multicamadas é o aprendizado supervisionado (BRAGA; CARVALHO; LUDERMIR, 2005). Neste tipo de treinamento, as entradas e as saídas desejadas são fornecidas, permitindo que, com base nestes valores, seja feito o ajuste do classificador de forma a relacionar estas entradas às suas saídas correspondentes.

O princípio do aprendizado supervisionado consiste na minimização da diferença entre a saída desejada e a encontrada pela rede para cada entrada do conjunto de treinamento. A mi-

Figura 5 – A representação gráfica de um neurônio artificial



Fonte: Traduzido de (HAYKIN, 2008)

minimização da diferença é incremental, proporcionando o ajuste dos pesos da rede neural e, conseqüentemente, levando a uma solução (BRAGA; CARVALHO; LUDERMIR, 2005).

Na Figura 7, um pequeno modelo esquemático pode ser visto. Este modelo aborda o funcionamento do aprendizado supervisionado para uma Rede Neural Artificial. Nesta, o Professor representa o conjunto de dados do treinamento e fornece à rede, representada por RNA, as entradas e saídas desejadas. Desta forma, o treinamento consiste na minimização do erro proveniente da diferença da saída fornecida pelo Professor e da saída fornecida pela rede. Esta minimização é incremental e é feita até que se satisfaça um critério de parada.

Dentre os métodos de aprendizado supervisionado, o *Backpropagation* é o mais utilizado em aplicações de redes neurais (JOHANSSON; DOWLA; GOODMAN, 1991). Este realiza a propagação do erro encontrado na camada de saída para as camadas internas da rede neural, buscando minimizar este erro pelo ajuste dos pesos em todas as camadas (HAYKIN, 2008).

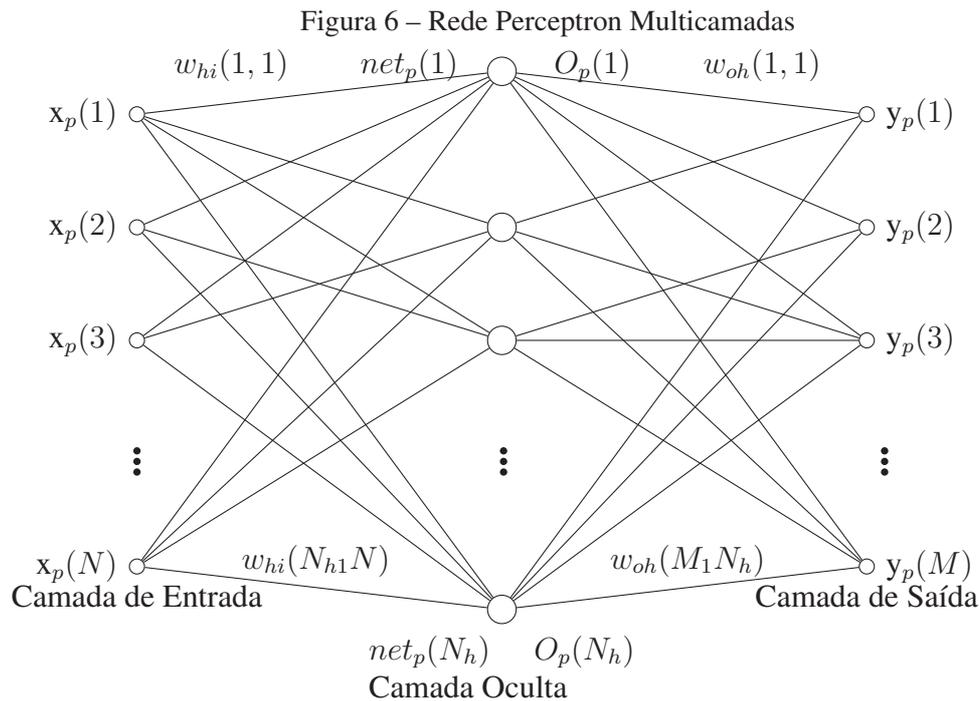
2.2.2 Validação de resultados de RNAs

Nesta subseção será apresentada a técnica utilizada para avaliar a precisão das classificações das redes neurais desenvolvidas, por meio dela foi possível comparar o modelo desenvolvido com os modelos desenvolvidos pelos trabalhos relacionados.

Dentre os métodos existentes, o *k-fold cross validation* foi escolhido e será apresentado na seção a seguir.

2.2.2.1 K-fold Cross Validation

K-fold Cross Validation é um dos métodos mais comuns para classificação de precisão, uma



Fonte: Traduzido de (WALTER; MICHAEL, 2005)

vez que ele permite mensurar o quão preciso um classificador pode trabalhar (SILVA, 2010).

Segundo (KOHAVI et al., 1995), neste método, um conjunto de dados D é particionado em k subconjuntos mutuamente excludentes (chamados de *folds*) D_1, D_2, \dots, D_k de tamanhos aproximadamente iguais. A rede é treinada e testada k vezes, onde, a cada iteração o conjunto D_k é reservado para testes e os demais são utilizados para treinamento.

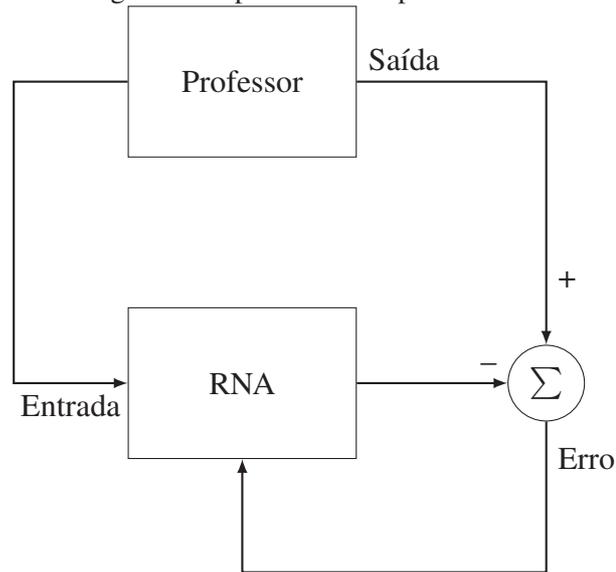
Ainda segundo (SILVA, 2010), *10-fold cross validation* é o método mais comum para este tipo de estimativa. Tal método provê uma série de informações a respeito da rede neural desenvolvida, a saber:

- **Estatística Kappa:** é uma medida de correlação entre medidas. Em (LANDIS; KOCH, 1977) se sugere a seguinte interpretação:

Valor Kappa	Interpretação
< 0	Sem correlação
0 – 0.19	Correlação fraca
0.2 – 0.39	Alguma correlação
0.4 – 0.59	Correlação moderada
0.6 – 0.79	Correlação substancial
0.8 – 1.00	Correlação quase perfeita

- **Erro absoluto médio:** mede o quanto, em média, as previsões se distanciaram dos dados efetivamente mensurados;
- **Erro quadrático médio:** também pode ser usado como uma medida de erro de previsão;

Figura 7 – Aprendizado supervisionado



Fonte: (BRAGA; CARVALHO; LUDERMIR, 2005)

- **Erro relativo absoluto:** é a média dos erros absolutos expressa em termos percentuais fornecendo uma indicação média do erro das previsões;
- **Erro relativo quadrado:** medida também expressa em termos percentuais em relação ao erro quadrático médio.

2.3 A Libras

Alguma contextualização faz-se necessária a respeito das línguas de sinais para melhor entendimento do presente trabalho. Portanto, nesta sessão, é feita uma breve contextualização histórica da Libras seguida de uma explanação a respeito dos elementos relevantes para o reconhecimento desta: os parâmetros.

2.3.1 Histórico

Uma língua natural é qualquer língua que tenha evoluído naturalmente nos seres humanos através do uso e repetição sem planejamento consciente ou premeditação (LYONS, 1991). As línguas de sinais são línguas naturais porque como as línguas orais surgiram espontaneamente da interação entre pessoas e porque devido à sua estrutura permitem a expressão de qualquer conceito - descritivo, emotivo, racional, literal, metafórico, concreto, abstrato - enfim, permitem a expressão de qualquer significado decorrente da necessidade comunicativa e expressiva do ser humano (BRITO, 1997).

2.3.1.1 O surdo no mundo

Segundo (GOLDFELD, 1997), a ideia que a sociedade fazia sobre os surdos, no decorrer da história, geralmente apresentava apenas aspectos negativos. Na antiguidade os surdos foram percebidos de formas variadas: com piedade e compaixão, como pessoas castigadas pelos deuses ou como pessoas enfeitiçadas, e por isso eram abandonadas ou sacrificadas.

Já nos séculos XVI, começaram a surgir as primeiras escolas para surdos. No século XVIII, em 1771, destaca-se a fundação d'A Instituição Nacional de Surdos Mudos, na França, por Abade Charles-Michel de l'Épée, instituição que mais tarde iria enviar professores para vários países do mundo, inclusive para o Brasil.

2.3.1.2 O surdo no Brasil

No Brasil, um marco importante para a história dos surdos, foi a chegada ao Brasil do professor Ernest Huet em 1855 a convite de Dom Pedro II (GOLDFELD, 1997). Este surdo francês trouxe o alfabeto manual francês e alguns sinais para o Brasil. Os surdos brasileiros, em contato com a Língua de Sinais Francesa (LSF), produziram a Língua de Sinais Brasileira (Libras), e a Figura 8 ilustra o atual alfabeto manual utilizado no Brasil. No ano seguinte, no dia 26 de setembro de 1857, foi fundado o Instituto dos Surdos-Mudos do Rio de Janeiro, e denominado o atual Instituto Nacional de Educação de Surdos (INES) (MONTEIRO, 2006).

Em 1911, o INES passou a seguir a tendência mundial, utilizando o oralismo puro em suas salas de aula. Todavia, o uso dos sinais permanece até 1957, momento em que a proibição é dada como oficial (ALBRES, 2010) durando até meados de 1980 (RAMOS, 2004).

O uso da Libras voltou a ganhar força nas décadas seguintes (BASTOS, 2015), chegando a virar língua oficial do país no início dos anos 2000 (BRASIL, 2002).

2.3.2 Os parâmetros da Libras

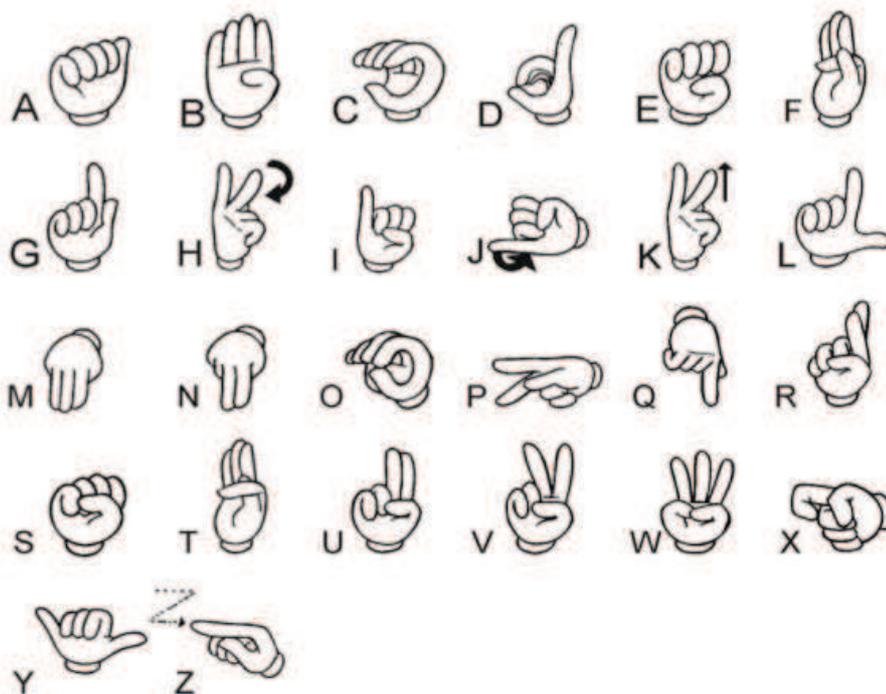
Segundo (BASTOS, 2015) e (LIEBER, 2015), a primeira formalização dos parâmetros necessários para identificação correta de um sinal foi feita por (STOKOE, 1960), tendo identificado 3 parâmetros principais:

- **Configuração de mão:** a forma assumida pela mão no momento da articulação de um determinado sinal, sendo que na Libras há 64 possíveis configurações de mão;
- **Ponto de Articulação ou Locação:** É a área do corpo na qual, ou próxima da qual, se articula o sinal;
- **Movimento:** É o movimento realizado pelas mãos do sinalizador no espaço.

Ainda segundo (LIEBER, 2015), (BATTISON, 1974) foi o primeiro trabalho a introduzir a

importância da orientação das palmas das mãos e (BAKER; PADDEN, 1978) a importância de expressões não manuais, ou seja, expressões faciais e corporais.

Figura 8 – Alfabeto manual da Libras



Fonte: (SILVA et al., 2017)

3 TRABALHOS RELACIONADOS

Nos últimos anos, observou-se um aumento na publicação de trabalhos relacionados ao reconhecimento sinais (KHAN; IBRAHEEM, 2012). Através de uma filtragem por décadas no portal do *IEEE Xplore* até o dia 20/02/2018, a Tabela 2 relaciona as últimas décadas e o número aproximado de trabalhos publicados e relacionados à *String* de busca ‘*Sign Language Recognition*’.

Tabela 2 – Trabalhos relacionados ao termo ‘*Sign Language Recognition*’.

Década	Número de Publicações
1981 - 1990	3
1991 - 2000	60
2001 - 2010	356
2011 - 2018	755

Fonte: Elaborado pelo autor

De forma a embasar a construção do método a ser empregada neste trabalho, foi feito um esforço no que diz respeito ao estudo de trabalhos relacionados ao reconhecimento de sinais e técnicas empregadas para este fim e, inclusive, a respeito da Libras, a qual foi escolhida como campo prático para a aplicação da abordagem.

O envolvimento computacional com as línguas de sinais remota à década de 80, quando a principal necessidade era a de otimização do envio de vídeos através de redes de computadores de longa distância (PARISH; SPERLING; LANDY, 1990).

A partir da década de 90, com o advento dos primeiros óculos de realidade virtual comerciais e das luvas eletrônicas que os acompanhavam, inicia-se uma busca por recursos de apoio à tradução das línguas de sinais para as línguas oralizadas, principalmente para sua forma escrita. Pode ser observada, então, sendo inicialmente feitas avaliações acerca do uso desta tecnologia, que apontavam dificuldades e restrições devido a questões técnicas associadas com os equipamentos e tecnologias utilizadas (QUAM, 1990). Ainda nesta década, aparecem os primeiros trabalhos que reconhecem sinais através de visão computacional como em (URAS; VERRI, 1994).

Atualmente os trabalhos de tradução das línguas de sinais podem ser divididos nos que se utilizam dispositivos auxiliares, como luvas eletrônicas, e nos que se utilizam de técnicas de processamento de imagens e visão computacional sem a necessidade de aparelhos acoplados ao corpo do usuário (BASTOS, 2015; FLORES; BARBOSA; RIGO, 2012). Por utilizar o LMC, o presente trabalho se encaixa no estado da arte, apresentado asseguir, junto aqueles trabalhos que se utilizam de visão computacional, atingindo um maior nível de interpretação, uma vez se utiliza de dois dos parâmetros globais da Libras, configurações de mão e movimentos de mão em combinação.

3.1 Abordagens baseadas em dispositivos auxiliares

No trabalho de (IGARI; FUKUMURA, 2014) o objetivo era a análise e o reconhecimento de 20 sinais com uma mão e 60 com duas mãos da Língua Japonesa de Sinais somente através dos movimentos feitos por estas. Na publicação, fez-se uma correlação de trajetórias através de pontos de passagens das mãos, ou seja, as mãos devem passar por determinados pontos no espaço para os sinais serem reconhecidos.

A comparação dos pontos adquiridos com o banco de dados é feita através de Programação Dinâmica e não reconhece configuração de mão. O sistema desenvolvido utiliza o histórico dos últimos 3 a 5 pontos no espaço 3D e utiliza um equipamento chamado *Libety* da empresa *POLHEMUS Inc.* e a taxa de reconhecimento ficou em 98%.

(NATESH et al., 2017) construíram um par de luvas de baixo custo (22 dólares norte americanos) específicas para este trabalho. Ele propõe a identificação de 45 sinais estáticos de 8 línguas de sinais diferentes. As luvas se utilizam de 17 sensores de flexão e contato em cada. Tais sensores geram 18 bits para cada mão, que são transformados nos correspondentes valores decimais. De posse destes 2 valores decimais a identificação do sinal é feita via busca em uma tabela de referências destes valores. O artigo apresenta uma taxa de 93.16% de precisão de reconhecimento.

3.2 Abordagens baseadas em visão computacional

(BASTOS, 2015) propôs a utilização de descritores para o reconhecimento de 40 sinais da Libras em imagens. No método, em um primeiro momento o autor utiliza uma RNPM para aplicação de uma máscara binária e consequente classificação de cada *pixel* como sendo de pele ou não. Tal rede neural deve ter como entrada um vetor de 9 posições referentes as informações de cor nos espaços de cores RGB (Vermelho, Azul e Verde), HSV (Matiz, Saturação e Valor) e YIQ (Luminância, Em-Fase e Quadratura). Tal método tem como objetivo destacar a presença da mão para o passo seguinte de extração de características.

O segundo passo do método proposto é a aplicação do Histograma de Gradientes Orientados (HOG) e de Momento Invariante de Zernike (MIZ) para extração de características que descrevessem as imagens.

Finalmente, objetivando a classificação das imagens, implementou-se uma segunda RNPM que recebia como parâmetros os dados computados no passo anterior e que atingiu uma taxa de 96.77% de reconhecimento.

(SAHA et al., 2016) propuseram reconhecer os 26 sinais do alfabeto manual da Língua Americana de Sinais (ASL) através de um novo conjunto de 11 características que alimentam uma Rede Neural MAdaline.

O primeiro passo é o pré-processamento das imagens para separar o que é mão do resto da imagem. Se faz isso através do uso dos espaços de cores RGB, HSV e YCbCr onde avalia-se

cada pixel dentro de um *threshold* para cada um dos espaços de cores.

A classificação é feita através da rede neural MAdaline se usa as seguintes características: Largura, comprimento, área e perímetro da mão na imagem já pré-processada. Ainda há outras 7 características calculadas a partir das 4 já citadas e que são nomeadas pelo autor como: Proporção na tela, circularidade, retangularidade, ralação do perímetro com comprimento e altura, relação do perímetro com a área, relação do comprimento com a área e relação da altura com a área.

Segundo o autor, a alta precisão de acertos, acima dos 90% justificariam o uso das características propostas, mas uma avaliação prévia seria necessária antes do seu uso, uma vez que são informações que podem ser redundantes podendo causar um *overhead* computacional desnecessário.

Um adendo importante, sobre este trabalho é que, apesar de o autor propor a identificação das 26 letras do alfabeto da ASL, este não correlaciona a identificação do movimento que envolvem as letras J e Z.

(KIM et al., 2017) capturaram as letras S, E, V, W, B e C da ASL utilizando um equipamento de transmissão de ondas de rádio. A recepção é feita por uma antena, amplificadores, um filtro passa-baixa e um osciloscópio. Segundo o autor, as formas das ondas mostram diferenças de amplitude e fase de acordo com as configurações de mão dos sinais, portanto estes dados foram utilizados como dados para uma Rede Neural Convolutiva atingindo uma taxa de reconhecimento acima de 90%.

(GARCÍA-BAUTISTA et al., 2017) utilizaram a câmera RGB-D (RGB com informação de profundidade) do Kinect para reconhecer 20 sinais da Língua Mexicana de Sinais (MSL). A proposta de método é a de obtenção da representação da trajetória da mão. O algoritmo DTW é usado para interpretar os sinais. Finalmente, se usa método it K-Fold Cross Validation para testes. Apresenta média de acerto de 98.57% para reconhecimento em tempo real. Limitação apresentada, pois o Kinect apresenta somente o esqueleto do corpo inteiro.

(PARIWAT; SERESANGTAKUL, 2017) dividiram o alfabeto de sinais tailandês em 3 grupos de acordo com movimento e combinações de caracteres. O trabalho concentrou-se na identificação do grupo de sinais estáticos. O trabalho envolve a captura de imagens coloridas de 280x288 de pessoas usando camisa preta em fundo azul e um processo em 8 estágios para segmentação da mão. Há um total de 87 características coletadas de cada imagem que são utilizadas em um algoritmo de *Support Vector Machine* para classificação. Ao final, chegou-se a uma precisão média de 91.2% de precisão na classificação dos sinais.

3.2.1 Utilização do Leap Motion Controller

(MOHANDÉS; ALIYU; DERICHE, 2014) fizeram um trabalho de comparação utilizando a Rede Neural Perceptron Multicamadas e o Classificador de Nave Bayes. No trabalho, foram selecionadas 12 das características fornecidos pelo Leap Motion, que incluem informações como

comprimento e largura dos dedos e posição do centro da palma. No trabalho, atingiu-se mais de 99% acerto com a Rede Perceptron e 98% com a Rede Bayesiana na classificação de 28 sinais do alfabeto estático da Língua Árabe de Sinais (ArSL);

(CHUAN; REGINA; GUARDINO, 2014) compararam os algoritmos de classificação *K-Nearest Neighbor* e *Support Vector Machine* (SVM) para classificar 26 letras da ASL. São utilizadas 9 informações das mãos para a identificação dos sinais, entre elas estão os ângulos entre cada dedo e o plano Z, distâncias entre as juntas dos dedos e o centro da palma, entre outros. Para o classificador *K-Nearest Neighbor*, atingiu-se 72.78% de precisão e 79.83% para o SVM. Apesar de o trabalho, inicialmente, se propor a reconhecer as 26 letras do alfabeto, não há menção ao J, única letra da ASL que, invariavelmente, precisa de movimento para ser reconhecida.

(ELONS et al., 2014) aplicaram a RNPM para o reconhecimento de 50 sinais ArSL. Como teste utilizou-se 2 tipos de dados de entrada, distância entre a posição de cada dedo e o Leap Motion, e a distância entre cada dedo e seu subsequente. Como resultado, atinge-se um máximo 88% de precisão para o segundo tipo de dados na classificação de sinais estáticos.

(FOK et al., 2015) fizeram uso de 2 Leap Motions e utilizaram do Modelo Oculto de Markov para a classificação dos números 0 a 9 da ASL, portanto somente sinais que não exigem movimento. Como dados de entrada do algoritmo, utilizou-se uma série de informações como a orientação de cada falange distal em relação ao sentido da palma, a distância de cada ponta de dedo para o ponto central da palma, a distância de cada ponta de dedo e a ponta do dedo subsequente e, por fim, o comprimento de cada dedo. Como resultado final, atingiu-se uma taxa média de reconhecimento de 93.14% com os 2 sensores.

(MAPARI; KHARAT, 2015) se utilizaram da distância euclidiana das pontas dos dedos para o centro da palma, a distância entre os dedos indicador e médio, indicador e anelar, indicador e mínimo para a identificação das Letras A à Z e dos números da Língua Indiana de Sinais, excluindo-se os números 0, 1, 2 e 7 por possuírem similaridades com outras configurações de mão. Dado o vetor de informações, fez-se duas tabelas comparativas,

(NAGLOT; KULKARNI, 2016) utilizaram a Rede Neural Perceptron Multicamadas para reconhecimento de 26 sinais do alfabeto da ASL. O trabalho utilizou como dados para classificação a distância entre pontas dos dedos a palma da mão e a distância entre dois dedos consecutivos, atingindo 96.15% de taxa de reconhecimento.

No sistema implementado por (ALMASRE; AL-NUAIM, 2016), utilizou-se o Leap Motion e o Kinect para extração de características. Utilizou-se o *Kernel Support Vector Machine* (KSVM) para classificação de 28 letras estáticas da Língua de Sinais Árabe. O sistema calcula 77 ângulos para as juntas das mãos e 26 ângulos a cada 2 ossos, totalizando 103 dados. Uso da PCA reduziu a dimensionalidade destes dados 103 para 36, que representaram 99% da variação. Como resultado, apresentou-se uma precisão de 86% para os dados de teste.

3.3 Comparativo dos trabalhos relacionados

A Tabela 3 abaixo sumariza alguns trabalhos relacionados à classificação de sinais da Libras, associando o número de características utilizadas para a identificação dos sinais por aquele trabalho, o método de reconhecimento utilizado, as taxas médias de reconhecimento atingidas por cada um dos métodos escolhidos e a quantidade de sinais que o trabalho se propôs a reconhecer.

Por haver nuances dos trabalhos, principalmente no que diz respeito ao tipo de dados que cada um utiliza, o que pode dificultar a comparação em termos de resultados, portanto na tabela abaixo chamamos a atenção principalmente para aqueles que se utilizam do LMC como fonte dos dados.

Tabela 3 – Tabela comparativa dos trabalhos relacionados

Autores	Número de Características	Método(s) de Reconhecimento	Taxa Média de Reconhecimento	Número de Sinais Reconhecidos	Configuração de Mão	Movimento
Utilizando Dispositivos Auxiliares						
(IGARI; FUKUMURA, 2014)	9 a 15	Histórico de pontos de passagem	98%	80	Não	Sim
(NATESH et al., 2017)	36	Tabela de referência de valores	93.16%	45	Sim	Não
Utilizando Visão Computacional						
(BASTOS, 2015)	2	RNPM	96.77%	40	Sim	Não
(SAHA et al., 2016)	11	MAdaline	> 90%	26	Sim	Não
(KIM et al., 2017)	2	RNPM	> 90%	6	Sim	Não
(GARCÍA-BAUTISTA et al., 2017)	60	DTW	99.1%	20	Não	Sim
(PARIWAT; SERESANGTAKUL, 2017)	87	SVM	91.2%	15	Sim	Não
Utilizando Leap Motion						
(MOHANDES; ALIYU; DERICHE, 2014)	12	Classificador Nave Bayes e RNPM	98% e 99%	28	Sim	Não
(CHUAN; REGINA; GUARDINO, 2014)	9	K-Nearest e SVM	72.78% e 79.83%	26	Sim	Não
(ELONS et al., 2014)	9	RNPM	88%	50	Sim	Não
(FOK et al., 2015)	19	Modelo Oculto de Markov	93.14%	10	Sim	Não
(MAPARI; KHARAT, 2015)	8	<i>Threshold</i> para Distância Euclidiana e Cosseno	88.39% e 90.32%	32	Sim	Não
(NAGLOT; KULKARNI, 2016)	9	RNPM	100%	36	Sim	Não
(ALMASRE; AL-NUAIM, 2016)	36	K SVM	86%	28	Sim	Não
(SIGNAPP, 2018)	21	RNMP	99.8% e 86.2%	26	Sim	Sim

4 MÉTODO

Este capítulo apresenta o método desenvolvido para a confecção do modelo e uma visão geral do protótipo desenvolvido. Este está subdividido do seguinte modo:

Seleção dos sinais – alguns sinais foram selecionados para a validação do modelo, bem como letras do alfabeto manual.

Aquisição dos dados – descreve os dados capturados, assim como seu processo de aquisição;

Identificação das configurações de mão – descreve o subconjunto das configurações de mão a serem identificadas, assim como o processo utilizado;

Identificação do movimento – apresenta o subconjunto de movimentos a serem identificados, bem como o processo utilizado;

Tecnologias utilizadas – enumera quais tecnologias foram utilizadas para implementar o modelo descrito.

4.1 Seleção dos sinais

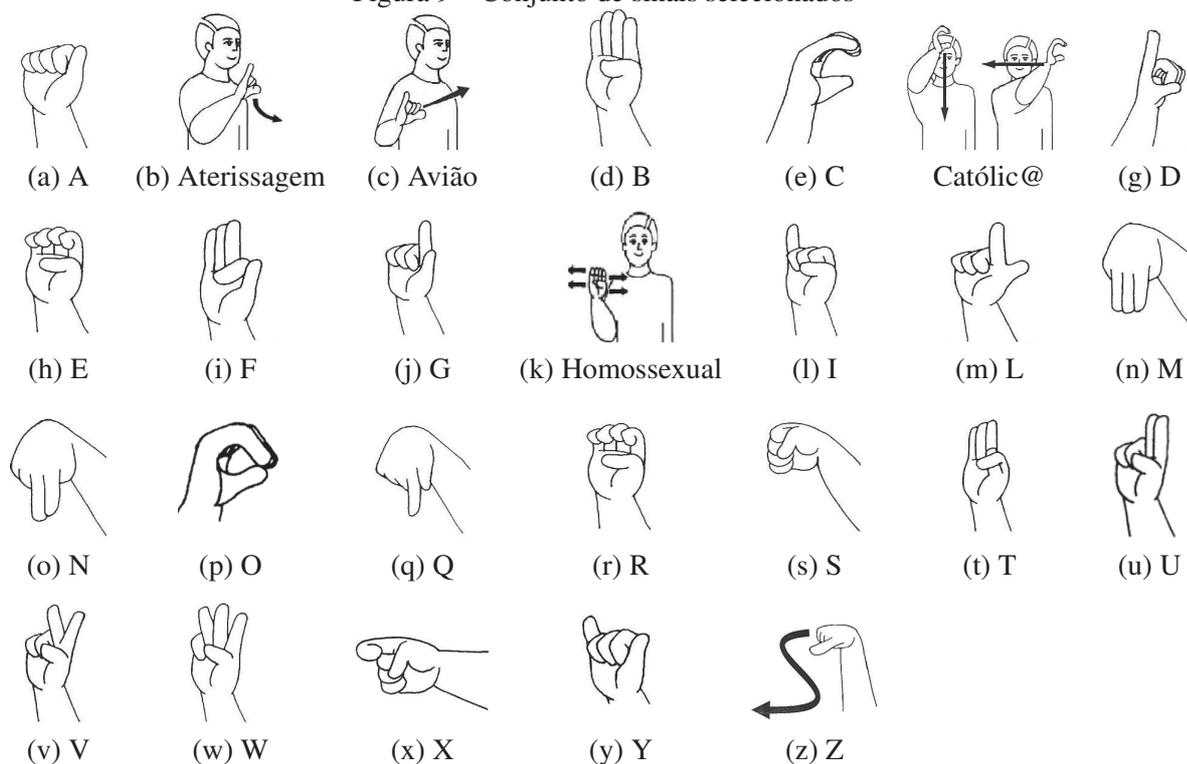
Nesta primeira sessão do método, são apresentados os 26 sinais selecionados para serem classificados. Tais sinais foram escolhidos afim de conter aqueles com e sem movimento. Também foram escolhidos conjuntos sinais que contivessem a mesma configuração de mão, podendo ser diferenciados pelo movimento. A Figura 9 apresenta os sinais selecionados.

4.2 Aquisição dos dados

A aquisição dos modelos tridimensionais das mãos é realizada utilizando o *Leap Motion Controller*, descrito na subseção 2.1.2. A motivação da seleção deste *hardware* em específico é atribuída à sua capacidade de capturar tais modelos em escala milimétrica de maneira precisa. Além disto existe a praticidade de acesso aos dados em tempo real, existindo, também, interfaces de programação para as principais linguagens do mercado. A *Application Programming Interface* (API) do Leap Motion API inclui diferentes características para reconhecimento da mão, dedos e ossos conforme abaixo:

- **Mão:** o modelo da mão fornece a informação sobre o seu tipo (esquerda ou direita), posição (a posição central da palma em milímetros), velocidade (em milímetro por segundo) entre outras características do braço a que o mão é anexada e dedos associados àquela mão.
- **Dedos:** características relacionadas aos dedos incluem direção (um vetor de direção), comprimento (em milímetros), largura, posição da ponta, velocidade da ponta e posição das falanges.

Figura 9 – Conjunto de sinais selecionados



Fonte: Adaptado de (CAPOVILLA, 2001)

- **Gestos:** quatro gestos básicos de movimento são reconhecidos pelo LMC, que são o movimento em círculo feito pelo indicador, toque em tecla, deslize e toque em tela.
- **Imagens:** imagens em escala de cinza são disponibilizadas pelo LMC. Quando o computador recebe os dados do *hardware*, o *software* do aparelho as trata aplicando algoritmos matemáticos. Primeiro trata-se o fundo, amenizado qualquer coisa que não seja de interesse, como objetos e luz ambiente. A imagem é então analisada e um modelo 3D da mão é criado.

Um exemplo de imagem e posição da mão no plano cartesiano 3D é mostrado na Figura 10.

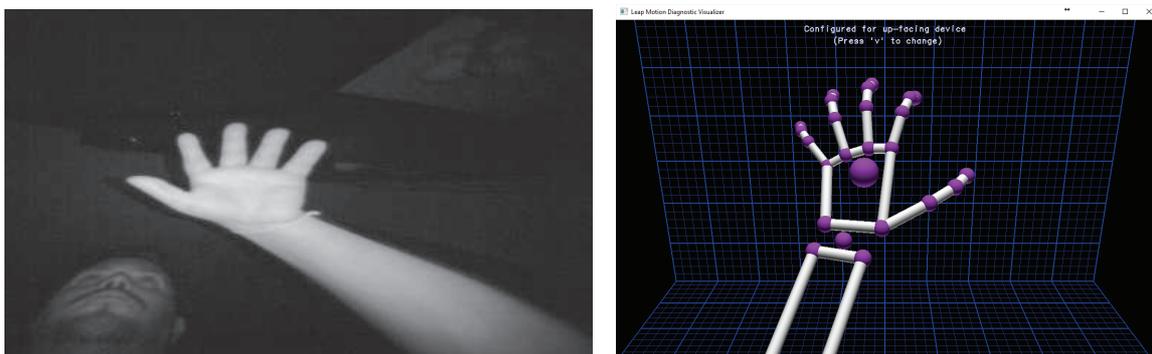
4.3 Identificação das configurações de mão

A identificação das configurações de mão se apresenta como um dos principais desafios da interpretação das línguas de sinais, portanto esta sessão explica o algoritmo desenvolvido para tal identificação.

Dentro do conjunto de 26 sinais selecionados, podem ser observadas um total de 19 configurações de mão, a Tabela 4 ilustra tal associação.

Com a análise do estado atual da arte, a abordagem de (NAGLOT; KULKARNI, 2016) teve destaque no que se refere a esta identificação e classificação. Deste modo, foi utilizado

Figura 10 – Exemplos de dados entregues pelo Leap Motion



(a) Exemplo de imagem entregue pelo Leap Motion

(b) Exemplo de posicionamento da mão no espaço 3D

Tabela 4 – Correlação entre configurações de mão e sinais selecionados

Configurações de mão	Sinal
Configuração em A	Letra A
Configuração em B	Letra B
Configuração em C	Letra C e Católico
Configuração em D	Letra D e Letra Z
Configuração em E	Letra E e Homossexual
Configuração em F	Letra F
Configuração em G	Letra G e Letra Q
Configuração em I	Letra I
Configuração em L	Letra L
Configuração em M	Letra M
Configuração em N	Letra N e Letra U
Configuração em O	Letra O
Configuração em R	Letra R
Configuração em S	Letra S
Configuração em T	Letra T
Configuração em V	Letra V
Configuração em W	Letra W
Configuração em X	Letra X
Configuração em Y	Letra Y, Avião e Aterrissagem

o modelo proposto pelos autores para implementação do recurso para esta dissertação. No trabalho, a classificação das configurações de mão é feita através de uma RNPM utilizando como dados de entrada da RNA:

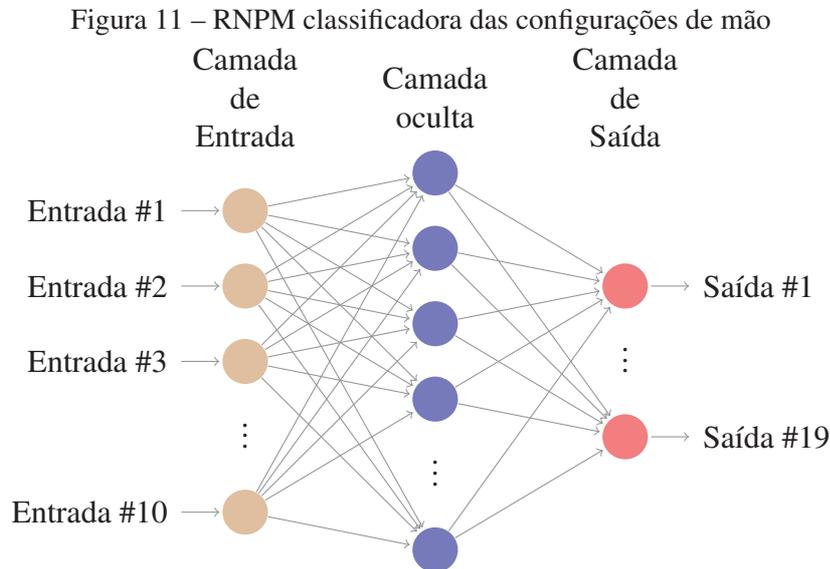
- As distâncias euclidianas entre a posição da ponta de cada dedo e a posição do centro palma; e
- As distâncias euclidianas entre a posição da ponta de dois dedos consecutivos.

O cálculo da distância euclidiana no espaço tridimensional entre dois pontos $P_1(x_1, y_1, z_1)$

e $P_2(x_2, y_2, z_2)$, utilizado para este trabalho afim de reimplementar o algoritmo proposto por (NAGLOT; KULKARNI, 2016) é descrito na Equação 4.1

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \quad (4.1)$$

A RNPM implementada para esta dissertação afim de classificar as configurações de mão é ilustrada pela Figura 11. Para facilitar o entendimento, as entradas numeradas de #1 à #5 representam as distâncias das pontas dos dedos até a posição do centro da palma, além disso, as entradas #6 à #10 representam as distâncias entre as pontas de dois dedos consecutivos, sendo a entrada #10 a distância do dedo mínimo para o polegar. Após o processamento da rede, as saídas #1 à #19 enumeram as configurações de mãos.



A rede neural desenvolvida se utiliza do método *backpropagation* para treinamento, tendo disponíveis 2000 amostras de cada uma das 19 configurações de mão para treinamento. A Tabela 5 ilustra um subconjunto dos dados utilizados.

Tabela 5 – Exemplos de dados usados na primeira RNA

Distância 1	Distância 2	Distância 3	Distância 4	Distância 5	Distância 6	Distância 7	Distância 8	Distância 9	Distância 10	Classe
59.890392	43.254395	39.30743	36.25928	37.18605	38.809265	12.588665	15.151179	15.975928	64.20929	A
51.456745	46.016586	40.978886	37.101067	80.83074	38.432358	14.711106	15.359165	86.46445	58.191963	IJ
78.26337	44.701645	41.855984	37.289246	81.97332	63.990616	13.022559	14.648802	79.50869	131.77003	Y

Para um melhor entendimento a respeito da Tabela 5, a coluna **Distância 1** contém as distâncias entre a ponta do polegar e o centro da palma até a coluna **Distância 5** que contém as distâncias entre a ponta do dedo mínimo e o centro da palma, enquanto as colunas intermediárias contém os dados dos demais dedos. A coluna **Distância 6** contém a distância entre o polegar e o indicador e a coluna **Distância 10** contém as medidas de distância entre o dedo mínimo e o polegar, enquanto as demais colunas contem as distâncias entre os demais dedos.

A fim de selecionar a melhor arquitetura da RNA, na camada escondida foram realizados testes onde esta camada continha entre 1 e 50 neurônios.

Por fim, a captura das configurações de mão ocorre em espaços de tempo e em tempo real. Outro dado importante, e que deve ser pontuado, é que a configuração de mão efetivamente selecionada ocorre por contagem simples dentro de uma janela de 300 *frames*, sendo que, tal janela ajuda a suprimir eventuais erros de classificação que possam ocorrer na transição de uma configuração de mão para outra.

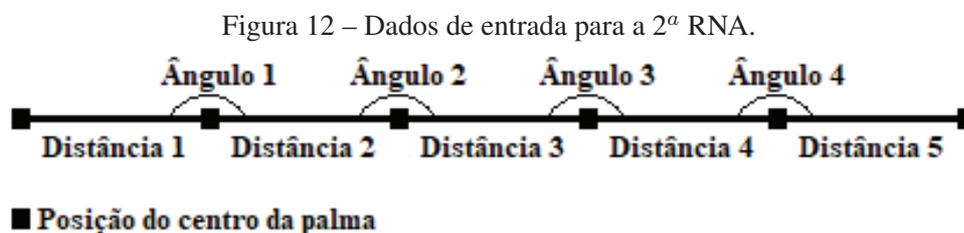
4.4 Identificação do movimento

O reconhecimento do movimento das mãos é chave para aplicações desde entretenimento a aplicações médicas (LOBO-PRAT et al., 2014). Por anos, o reconhecimento do movimento das mãos foi realizado com o auxílio de luvas eletrônicas (DIPIETRO; SABATINI; DARIO, 2008), mas o advento de técnicas de visão computacional criou inúmeras interfaces humano-computador como o reconhecimento do movimento das mãos pelas atuais *Smart TVs* e por consoles de *video games* como *Nintendo Wii* (CHIANG et al., 2009) e o *SONY Playstation Move* (PERL; VENDITTI; KAUFMANN, 2013).

Assim como ocorre com as configurações de mão, afim de reconhecer os seus movimentos, foi desenvolvida uma segunda rede neural que trabalha dentro de uma janela de 150 *frames*. Os dados de entrada utilizados para esta 2ª RNA são descritos como segue:

- Cálculo de distância euclidiana do centro da palma capturada a cada 30 *frames*; e
- Cálculo do ângulo do percurso do centro da palma a cada 60 *frames*.

portanto, contabilizando um total de 9 dados de entrada. A título de ilustração e melhor entendimento, a Figura 12 ilustra os dados que são calculados.

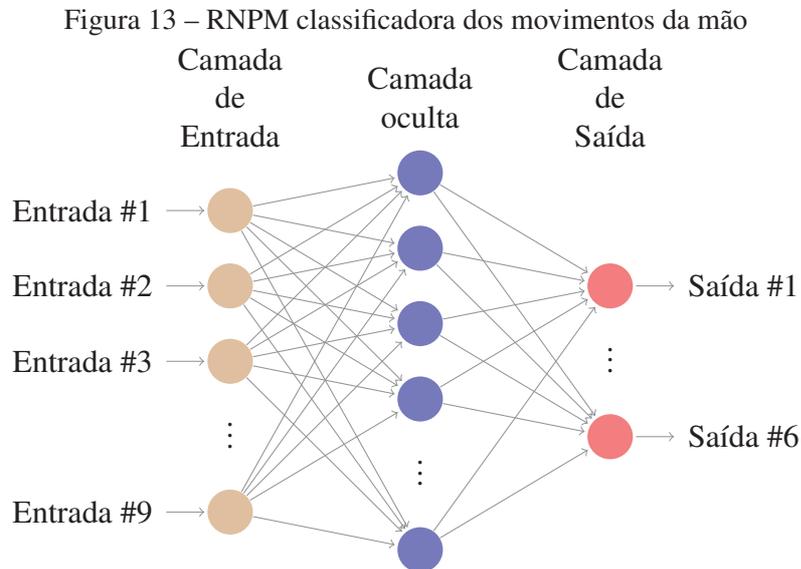


A Tabela 6 ilustra um subconjunto dos dados submetidos à RNA com objetivo de treinamento desta.

Tabela 6 – Exemplo de dados usados no treinamento da segunda RNA

Ângulo 1	Ângulo 2	Ângulo 3	Ângulo 4	Ângulo 5	Distância 1	Distância 2	Distância 3	Distância 4	Distância 5	Distância 6	Classe
69.327835	89.4853	45.924187	131.55725	141.0938	1.4434441	0.69450194	2.282888	3.2797399	1.0181031	1.7397951	Parada
174.69548	176.47647	53.393066	45.31341	162.48056	31.432396	76.47482	122.66641	209.02522	129.57028	171.79805	Católico
158.67671	161.37009	29.390408	125.8734	89.22113	57.168312	107.29819	92.43574	162.90576	26.764475	231.0731	Z

A Figura 13 ilustra a RNA desenvolvida a fim de reconhecer os movimentos das mãos, sendo as entradas #1 à #4 os ângulos calculados e as entradas #5 à #9 as distâncias.



Esta RNA possui 6 neurônios de saída, pois dentre os sinais selecionados, esta é a quantidade de sinais que possuem movimento, sendo relacionados da seguinte maneira:

- Saída #1 – Católico
- Saída #2 – Homossexual
- Saída #3 – Aterrissagem
- Saída #4 – Avião
- Saída #5 – Mão Parada
- Saída #6 – Movimento em Z

Assim como na rede neural utilizada para a classificação das configurações de mão, afim de selecionar a melhor arquitetura da RNA, também foram realizados testes para a camada escondida contendo entre 1 e 50 neurônios.

4.5 Tecnologias utilizadas

Para validar na prática o modelo proposto, foi desenvolvido um protótipo com algumas tecnologias computacionais conforme descritas a seguir:

- *Leap Motion Controller*: este foi o *hardware* escolhido pela facilidade em se capturar e trabalhar com as informações a respeito da estrutura virtual das mãos, tais informações fornecidas pelo aparelho em tempo real;

- **Java:** O LMC consegue trabalhar com as principais linguagens de programação do mercado, dentre elas a linguagem Java, escolhida para poder fazer a troca de informações com a tecnologia escolhida para a implementação das RNAs;
- *Waikato Environment for Knowledge Analysis (WEKA):* É um conjunto de algoritmos designados ao aprendizado de máquina, entre eles estão ferramentas de pré-processamento, classificação e visualização de dados, além de métodos de validação dos modelos desenvolvidos.

5 RESULTADOS

Nesse capítulo são apresentados os resultados obtidos nos experimentos com as camadas ocultas de ambas as redes neurais desenvolvidas, tanto para a classificação das configurações de mão quanto para a classificação dos movimentos, assim como a interface de *software* desenvolvida para a aplicação do modelo.

5.1 Classificação das configurações de mão

O primeiro experimento utiliza uma base de dados construída para tal. A base contém 114.000 amostras das 19 configurações de mão de 3 pessoas diferentes.

Os testes realizados utilizam entre 1 e 50 neurônios na camada oculta. A estrutura que obteve uma melhor classificação utiliza 15 neurônios.

Como pode ser lido na Tabela 7, a RNA que classifica configurações de mão teve um bom desempenho, tendo Estatística Kappa avaliado como 1, estabelecendo correlação quase perfeita segundo (LANDIS; KOCH, 1977). Já a tabela 8 mostra a matriz de confusão desta mesma RNA, mostrando que todas as 38.000 instâncias submetidas à classificação foram corretamente classificadas.

Tabela 7 – *Cross-Validation* das configurações de mão

Medida	Valor
Instâncias classificadas corretamente	113807 (99.8%)
Instâncias classificadas incorretamente	193 (0.2%)
Estatística Kappa	0.99
Erro absoluto médio	0.0004
Erro médio quadrado	0.012
Erro relativo absoluto	0.3804%
Erro relativo quadrado	5.3689%
Total de instâncias	114000

5.2 Classificação do movimento

O segundo experimento utiliza uma base de dados de movimentos construída contendo 300 amostras de cada uma dos 6 movimentos classificados. Tais amostras consistindo de 9 informações diferentes.

Os testes realizados utilizam entre 1 e 50 neurônios na camada oculta. A estrutura que obteve uma melhor classificação utiliza 26 neurônios.

Como pode ser lido na Tabela 9, a RNA que classifica os movimentos de mão teve um bom desempenho, apesar de abaixo da rede desenvolvida para as configurações de mão, podendo trabalhar com até 86.7% de precisão e estabelecendo Estatística Kappa em 0.84, ou seja, esta-

Tabela 8 – Matriz de confusão das configurações de mão

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	<— classificado como
5999	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	a = A
0	6000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	b = B
0	0	6000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	c = C
0	0	0	6000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	d = DZ
0	0	0	0	6000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	e = E
0	0	0	0	0	6000	0	0	0	0	0	0	0	0	0	0	0	0	0	f = F
0	0	0	0	0	0	6000	0	0	0	0	0	0	0	0	0	0	0	0	g = GQ
0	0	0	0	0	0	0	6000	0	0	0	0	0	0	0	0	0	0	0	h = I
0	0	0	0	0	0	0	0	6000	0	0	0	0	0	0	0	0	0	0	i = L
0	0	0	0	0	0	0	0	0	6000	0	0	0	0	0	0	0	0	0	j = M
0	0	0	0	0	0	0	0	0	0	6000	0	0	0	0	0	0	0	0	k = NU
0	0	0	0	0	0	0	0	0	0	0	6000	0	0	0	0	0	0	0	l = O
0	0	21	0	0	0	0	0	12	32	99	0	5808	0	0	28	0	0	0	m = R
0	0	0	0	0	0	0	0	0	0	0	0	0	6000	0	0	0	0	0	n = S
0	0	0	0	0	0	0	0	0	0	0	0	0	0	6000	0	0	0	0	o = T
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6000	0	0	0	p = V
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6000	0	0	q = W
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6000	0	r = X
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6000	s = Y

belecendo uma correlação substancial entre os dados de entrada e saída da RNA. Já a tabela 10 mostra a matriz de confusão desta segunda RNA, mostrando uma classificação um pouco mais esparsa das 1200 instâncias submetidas. Ressalta-se o resultado do movimento de mão parada, que obteve 99.8% de acerto.

Tabela 9 – Cross-validation do movimento

Medida	Valor
Instâncias classificadas corretamente	144 (86.7%)
Instâncias classificadas incorretamente	22 (13.3%)
Estatística Kappa	0.8404
Erro absoluto médio	0.0614
Erro médio quadrado	0.1977
Erro relativo absoluto	22.1426%
Erro relativo quadrado	53.1211%
Total de instâncias	166

Tabela 10 – Matriz de confusão da classificação do movimento

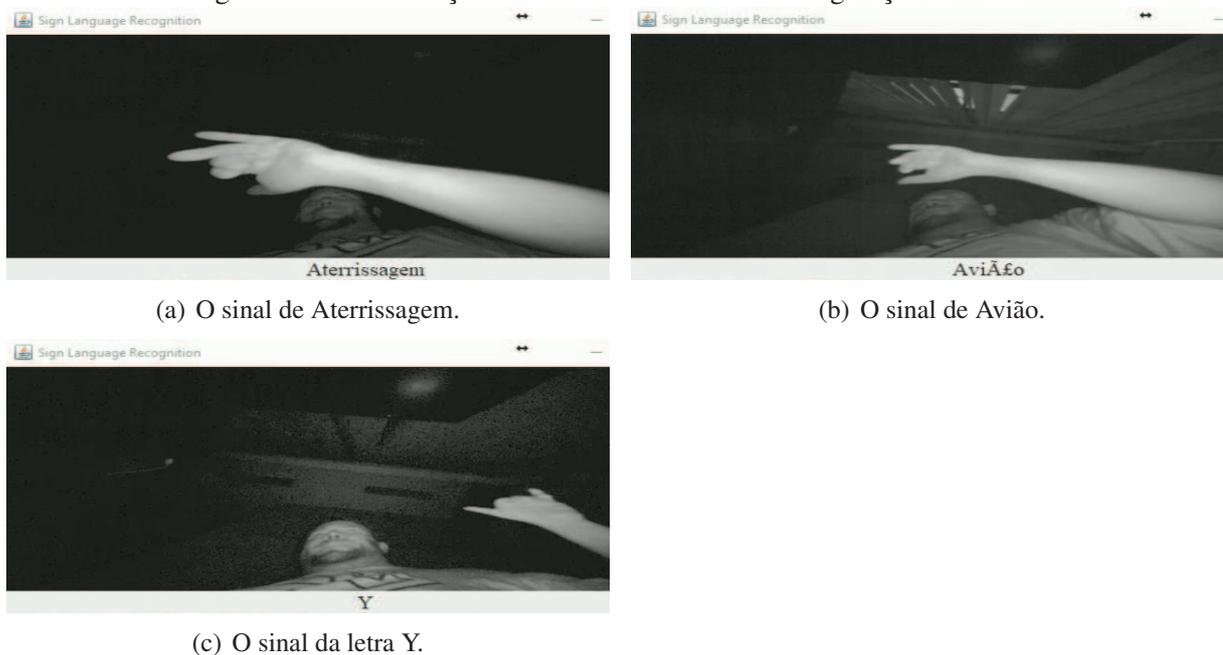
a	b	c	d	e	f	<— classificado como
23	0	0	0	0	7	a = Católico
0	20	0	0	3	0	b = Homossexual
0	0	20	3	0	0	c = Aterrissagem
0	0	0	30	0	0	d = Avião
0	2	0	0	28	0	e = Mão Parada
5	1	1	0	0	23	g = Z

5.3 Aplicação do modelo

A avaliação do modelo proposto foi realizada através do desenvolvimento de um protótipo. Foram avaliados alguns sinais que possuem a mesma configuração de mão mas significados

diferentes em função de seu movimento. Apesar deste aspecto não poder ser diferenciado de maneira satisfatória somente através de imagens, a Figura 14 exemplifica tal ganho.

Figura 14 – Classificação de sinais com a mesma configuração de mão.



A Figura 14(a) se refere ao sinal equivalente à palavra ‘Aterrissagem’, onde se tem a mesma configuração de mão da letra Y em adição ao movimento descendente da mão; a Figura 14(b) se refere ao sinal da palavra ‘Avião’, onde se tem a mesma configuração de mão da letra Y também, porém com movimento retilíneo e perpendicular ao tórax do sinalizador; e, por fim, a Figura 14(c) se refere ao sinal da letra ‘Y’ que, além da própria configuração de mão, têm-se a identificação que a mão está parada. Estes movimentos foram discernidos em tempo real pelo modelo e configuram o potencial de tradução para outros tipos de sinais que possuem a mesma configuração de mão e movimentos diferentes (ou vice-versa).

6 CONCLUSÃO E TRABALHOS FUTUROS

A tecnologia tem revolucionado o mundo das comunicações nos últimos anos, nesse sentido, ela pode fazer com que a sociedade seja um lugar mais acolhedor, principalmente em uma época em que se discute a inclusão das minorias. Entre estas minorias está a comunidade surda, que permanecem com certas dificuldades, uma vez que a sociedade ainda é majoritariamente construída para o mundo dos ouvintes. A partir da necessidade de, cada vez mais, incluir esta população surda, surge a inevitável interação através da comunicação entre surdos e ouvintes, portanto esta dissertação teve a oportunidade de realizar um estudo a respeito da língua utilizada pelos surdos, pois tal interação carece de conhecimento das línguas de sinais principalmente por parte dos ouvintes.

Uma vez que existe uma crescente quantidade de trabalhos em relação ao reconhecimento das línguas de sinais, ainda assim existe uma carência por ferramentas intuitivas que auxiliem na interação surdo-ouvinte e que auxiliem na disseminação destas línguas.

No que diz respeito ao trabalhos relacionados, isto é, trabalhos que consigam integrar a classificação de configurações de mão e seus movimentos, não foram encontrados trabalhos que o façam, há somente de maneira separadas, abrindo uma lacuna onde este trabalho se encaixa, pois o modelo apresentado já é capaz de reconhecer configurações de mão e os movimentos dos sinais.

Comparando o método proposto por esta dissertação com os métodos dos trabalhos relacionados, observou-se que as RNAs desenvolvidas se mostraram bastante satisfatórias no que diz respeito as taxas de classificação das configurações de mão, tendo taxas de erro pouco significativas, tendo atingido expressivos 99.8% de acerto. No que tange a classificação dos movimentos de mão, os resultados ainda trazem um grau mais significativo incerteza, atingindo taxas de 86.7% para os movimentos selecionados. Porém, o trabalho de pesquisa realizado abre espaço para o aprofundamento de estudos que busquem avançar na questão da classificação das configurações de mão e seus movimentos em tempo real, visto que o aparelho utilizado não conseguiu identificar de maneira satisfatória a configuração de mão utilizada para as letras H, K e P.

Como trabalhos futuros, o objetivo é dar seguimento à pesquisa, realizando novos experimentos e implementando novas técnicas que possam ajudar a suprir eventuais lacunas deixadas por este trabalho, assim como adicionar novas variáveis que possam dar suporte ao aumento dos sinais reconhecidos, dando uma maior amplitude ao trabalho desenvolvido. Visto que o modelo desenvolvido foi concebido no modelo de camadas (captura de dados, pré-processamento e classificação e visualização), isto facilita a inclusão de novos módulos de reconhecimento de padrões e, eventualmente, até a substituição por novas tecnologias que auxiliem a aumentar a precisão do que já foi desenvolvido.

Outro ponto que almejamos fortemente como trabalho futuro é realizar um número maior de testes com as redes neurais, principalmente àquela que diz respeito à classificação do mo-

vimento, afim de ter uma melhor avaliação da rede neural do movimento, uma vez que esta rede requer uma quantidade razoável de pré-processamento, os dados submetidos à rede neural ficam bastante reduzidos em comparação com a amostra capturada em tempo real.

Uma segunda questão que pode ser levada como trabalho futuro é a avaliação do *software* desenvolvido junto à comunidade surda, principalmente em virtude da relevância do que foi produzido para esta comunidade afim de avaliar sua coerência com suas necessidades e avaliar possíveis melhorias a serem acrescentadas no modelo.

REFERÊNCIAS

- ALBRES, N. d. A. Surdos & inclusão educacional. **Rio de Janeiro: Arara Azul**, [S.l.], 2010.
- ALMASRE, M. A.; AL-NUAIM, H. Recognizing arabic sign language gestures using depth sensors and a ksvm classifier. In: **COMPUTER SCIENCE AND ELECTRONIC ENGINEERING (CEEC)**, 2016. **Anais...** [S.l.: s.n.], 2016. p. 146–151.
- AUDY, A. **Leap motion** : le test complet ! Disponível em:
<<http://www.clubic.com/technologies-d-avenir/article-575170-1-leap-motion-test.html>>.
Acesso em: 5 de junho de 2017.
- BAKER, C.; PADDEN, C. Focusing on the nonmanual components of american sign language. **Understanding language through sign language research**, [S.l.], p. 27–57, 1978.
- BASTOS, I. L. O. **Reconhecimento de sinais da libras utilizando descritores de forma e redes neurais artificiais**. 2015. Dissertação (Mestrado em Ciência da Computação) — Instituto de Matemática, 2015.
- BATTISON, R. Phonological deletion in american sign language. **Sign language studies**, [S.l.], v. 5, n. 1, p. 1–19, 1974.
- BRAGA, A. d. P.; CARVALHO, A.; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. [S.l.]: Livros Técnicos e Científicos, 2005.
- BRASIL. Presidência da república. lei nº 10.436, de 24 de abril de 2002. dispõe sobre a língua brasileira de sinais - libras e dá outras providências. in: casa civil, subchefia para assuntos jurídicos. **Diário Oficial da União**, Brasília, DF, 2002. p. 23.
- BRITO, L. F. Língua brasileira de sinais - libras. **Educação Especial, Língua Brasileira de Sinais**, [S.l.], v. 3, 1997. nota.
- CAPOVILLA, F. C. **Dicionário enciclopédico ilustrado trilingue da língua de sinais brasileira: sinais de m a z**. [S.l.]: EdUSP, 2001. v. 2.
- CHIANG, P.-Y. et al. Wiistick: enhancing motion recognition capability for wii systems. In: **CIRCUITS AND SYSTEMS, 2009. ISCAS 2009. IEEE INTERNATIONAL SYMPOSIUM ON**, 2009. **Anais...** [S.l.: s.n.], 2009. p. 1445–1448.
- CHUAN, C. H.; REGINA, E.; GUARDINO, C. American sign language recognition using leap motion sensor. In: **INTERNATIONAL CONFERENCE ON MACHINE LEARNING AND APPLICATIONS**, 2014., 2014. **Anais...** [S.l.: s.n.], 2014. p. 541–544.
- COLGAN, A. **The alignment problem**: how to position cameras for augmented reality. Disponível em:
<<http://blog.leapmotion.com/alignment-problem-position-cameras-augmented-reality/>>.
Acesso em: 5 de junho de 2017.
- CUEVAS, I. F. et al. Infrared thermography for the detection of injury in sports medicine. In: **Application of infrared thermography in sports science**. [S.l.]: Springer, 2017. p. 81–109.
- DEMUTH, H. B. et al. **Neural network design**. [S.l.]: Martin Hagan, 2014.

DIPIETRO, L.; SABATINI, A. M.; DARIO, P. A survey of glove-based systems and their applications. **IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)**, [S.l.], v. 38, n. 4, p. 461–482, 2008.

ELONS, A. S. et al. Arabic sign language recognition using leap motion sensor. In: INTERNATIONAL CONFERENCE ON COMPUTER ENGINEERING SYSTEMS (ICCES), 2014., 2014. **Anais...** [S.l.: s.n.], 2014. p. 368–373.

FLORES, E. M.; BARBOSA, J. L. V.; RIGO, S. J. Um estudo de técnicas aplicadas ao reconhecimento da língua de sinais: novas possibilidades de inclusão digital. **RENOTE**, [S.l.], v. 10, n. 3, 2012.

FOK, K. Y. et al. A real-time asl recognition system using leap motion sensors. In: INTERNATIONAL CONFERENCE ON CYBER-ENABLED DISTRIBUTED COMPUTING AND KNOWLEDGE DISCOVERY, 2015., 2015. **Anais...** [S.l.: s.n.], 2015. p. 411–414.

FRANÇA, I. S. X. de et al. A relação entre atividades de lazer e níveis glicêmicos de adultos surdos. **Northeast Network Nursing Journal**, [S.l.], v. 14, n. 6, 2013.

GOLDFELD, M. **A criança surda: linguagem e cognição numa perspectiva sociointeracionista**. [S.l.]: Plexus Editora, 1997.

GUEDES, B. S. **Sobre surdos, bocas e mãos: saberes que constituem o currículo de fonoaudiologia**. 2010. Dissertação (Mestrado em Ciência da Computação) — Universidade do Vale do Rio do Sinos, 2010.

HAN, T. Y.; SONG, B. C. Night vision pedestrian detection based on adaptive preprocessing using near infrared camera. In: CONSUMER ELECTRONICS-ASIA (ICCE-ASIA), IEEE INTERNATIONAL CONFERENCE ON, 2016. **Anais...** [S.l.: s.n.], 2016. p. 1–3.

HAYKIN, S. O. **Neural networks and learning machines**. 3. ed. [S.l.]: Pearson, 2008.

IGARI, S.; FUKUMURA, N. Sign language word recognition using via-point information and correlation of they bimanual movements. In: INTERNATIONAL CONFERENCE OF ADVANCED INFORMATICS: CONCEPT, THEORY AND APPLICATION (ICAICTA), 2014., 2014. **Anais...** [S.l.: s.n.], 2014. p. 75–80.

JIANG, Q.; DANIELL, C. Recognition of human and animal movement using infrared video streams. In: IMAGE PROCESSING, 2004. ICIP'04. 2004 INTERNATIONAL CONFERENCE ON, 2004. **Anais...** [S.l.: s.n.], 2004. v. 2, p. 1265–1268.

JOHANSSON, E. M.; DOWLA, F. U.; GOODMAN, D. M. Backpropagation learning for multilayer feed-forward neural networks using the conjugate gradient method. **International Journal of Neural Systems**, [S.l.], v. 2, n. 04, p. 291–301, 1991.

KHAN, R. Z.; IBRAHEEM, N. A. Survey on gesture recognition for hand image postures. **Computer and Information Science Journal**, [S.l.], v. 5, p. 110–121, 2012.

KIM, S. Y. et al. A hand gesture recognition sensor using reflected impulses. **IEEE Sensors Journal**, [S.l.], v. 17, n. 10, p. 2975–2976, May 2017.

KOHAVI, R. et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: IJCAI, 1995. **Anais...** [S.l.: s.n.], 1995. v. 14, n. 2, p. 1137–1145.

LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. **biometrics**, [S.l.], p. 159–174, 1977.

Leap Motion. Disponível em: <<https://developer.leapmotion.com/>>. Acesso em: 9 de maio de 2017.

LIEBER, S. N. **Aspectos da constituição de uma criança surda pela fala do ouvinte**: entre traços e significantes. 2015. Dissertação (Mestrado em Ciência da Computação) — Pontifícia Universidade Católica de São Paulo, 2015.

LIEW, D. S. Electromagnetic waves. **Centre for Remote Imaging, Sensing and Processing**, [S.l.], p. 10–27, 2006.

LOBO-PRAT, J. et al. Non-invasive control interfaces for intention detection in active movement-assistive devices. **Journal of neuroengineering and rehabilitation**, [S.l.], v. 11, n. 1, p. 168, 2014.

LYONS, J. **Natural language and universal grammar**: volume 1: essays in linguistic theory. [S.l.]: Cambridge University Press, 1991. v. 1.

MAPARI, R. B.; KHARAT, G. Real time human pose recognition using leap motion sensor. In: IEEE INTERNATIONAL CONFERENCE ON RESEARCH IN COMPUTATIONAL INTELLIGENCE AND COMMUNICATION NETWORKS (ICRCICN), 2015., 2015. **Anais...** [S.l.: s.n.], 2015. p. 323–328.

MCCREARY, J. Infrared (ir) basics for digital photographers-capturing the unseen, u. **Digital Photography For What It's Worth**, [S.l.], 2004.

MOHANDÉS, M.; ALIYU, S.; DERICHE, M. Arabic sign language recognition using the leap motion controller. In: IEEE 23RD INTERNATIONAL SYMPOSIUM ON INDUSTRIAL ELECTRONICS (ISIE), 2014., 2014. **Anais...** [S.l.: s.n.], 2014. p. 960–965.

MONTEIRO, M. S. História dos movimentos dos surdos e o reconhecimento da libras no brasil. **ETD-Educação Temática Digital**, [S.l.], v. 7, n. 2, p. 295–305, 2006.

NAGLOT, D.; KULKARNI, M. Real time sign language recognition using the leap motion controller. In: INTERNATIONAL CONFERENCE ON INVENTIVE COMPUTATION TECHNOLOGIES (ICICT), 2016., 2016. **Anais...** [S.l.: s.n.], 2016. v. 3, p. 1–5.

NATESH, A. et al. Low-cost wireless intelligent two hand gesture recognition system. In: ANNUAL IEEE INTERNATIONAL SYSTEMS CONFERENCE (SYSCON), 2017., 2017. **Anais...** [S.l.: s.n.], 2017. p. 1–6.

OLIVEIRA, A. D. de et al. Comprehensive near infrared study of jatropa oil esterification with ethanol for biodiesel production. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, [S.l.], v. 170, p. 56–64, 2017.

PARISH, D. H.; SPERLING, G.; LANDY, M. S. Intelligent temporal subsampling of american sign language using event boundaries. **Journal of Experimental Psychology: Human Perception and Performance**, [S.l.], v. 16, n. 2, p. 282, 1990.

- PARIWAT, T.; SERESANGTAKUL, P. Thai finger-spelling sign language recognition using global and local features with svm. In: INTERNATIONAL CONFERENCE ON KNOWLEDGE AND SMART TECHNOLOGY (KST), 2017., 2017. **Anais...** [S.l.: s.n.], 2017. p. 116–120.
- PERL, T.; VENDITTI, B.; KAUFMANN, H. Ps move api: a cross-platform 6dof tracking framework. In: WORKSHOP ON OFF-THE-SHELF VIRTUAL REALITY, 2013. **Proceedings...** [S.l.: s.n.], 2013.
- QUAM, D. L. Gesture recognition with a dataglove. In: AEROSPACE AND ELECTRONICS CONFERENCE, 1990. NAECON 1990., PROCEEDINGS OF THE IEEE 1990 NATIONAL, 1990. **Anais...** [S.l.: s.n.], 1990. p. 755–760.
- RAMOS, C. R. Libras: a língua de sinais dos surdos brasileiros. **Revista virtual de cultura surda**, [S.l.], 2004.
- RODRIGUES, P. R.; ALVES, L. R. G. Criar e compartilhar games: novas possibilidades de letramento digital para crianças surdas. **RENOTE**, [S.l.], v. 12, n. 2, 2014.
- ROSENBLATT, F. Perceptual generalization over transformation groups. **Self-organizing systems**, [S.l.], p. 63–96, 1960.
- ROWAN-ROBINSON, M. **Night vision: exploring the infrared universe**. [S.l.]: Cambridge University Press, 2013.
- SAHA, S. et al. A novel approach to american sign language recognition using madaline neural network. In: IEEE SYMPOSIUM SERIES ON COMPUTATIONAL INTELLIGENCE (SSCI), 2016., 2016. **Anais...** [S.l.: s.n.], 2016. p. 1–6.
- SILVA, F. I. da et al. **Aprendendo libras como segunda língua: nível básico**. [S.l.]: Federal Instite from Santa Catarina, 2017.
- SILVA, R. S. **Previsão de venda de produtos em uma indústria de telecomunicação utilizando redes neurais artificiais**. 2010. Dissertação (Mestrado em Ciência da Computação) — Universidade do Vale do Rio do Sinos, 2010.
- SOUZA, M. T. de; PORROZZI, R. Ensino de libras para os profissionais de saúde: uma necessidade premente. **Revista Práxis**, [S.l.], v. 1, n. 2, 2009.
- STOKOE, W. C. Sign language structure: an outline of the visual communication systems of the american deaf. **Studies in Linguistics: Occasional Papers**, [S.l.], n. 8, 1960. Dept. of Anthropology and Linguistics, University of Buffalo.
- URAS, C.; VERRI, A. On the recognition of the alphabet of the sign language through size functions. In: IAPR INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, VOL. 3 - CONFERENCE C: SIGNAL PROCESSING (CAT. NO.94CH3440-5), 12., 1994. **Proceedings...** [S.l.: s.n.], 1994. v. 2, p. 334–338 vol.2.
- WALTER, H.; MICHAEL, T. Recent developments in multilayer perceptron neural networks. In: ANNUAL MEMPHIS AREA ENGINEERING AND SCIENCE CONFERENCE, 7., 2005. **Proceedings...** [S.l.: s.n.], 2005.
- WEAST, R. C. et al. **Crc handbook of chemistry and physics**. 92. ed. [S.l.]: CRC press Boca Raton, FL, 2011.

WITCHES, P. H. **A educação de surdos no estado novo**: práticas que constituem uma brasilidade surda. 2014. Dissertação (Mestrado em Ciência da Computação) — Universidade do Vale do Rio dos Sinos, 2014.

World Health Organization. **Deafness and hearing loss**. Disponível em: <<http://www.who.int/mediacentre/factsheets/fs300/en/>>. Acesso em: 18 de maio de 2017.