

**Universidade Federal do Maranhão**  
Centro de Ciências Exatas e Tecnológicas  
Faculdade de Engenharia Elétrica

Daniel de Matos Luna dos Santos

**Desenvolvimento de um Sistema de Reposicionamento para um  
Robô de Sondagem utilizando Redes Convolucionais**

São Luís  
2018

Daniel de Matos Luna dos Santos

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica, ao Programa de Pós-Graduação em Engenharia Elétrica, da Universidade Federal do Maranhão. Área de concentração: Processamento da Informação Biológica.

**Desenvolvimento de um Sistema de Reposicionamento para um Robô de  
Sondagem utilizando Redes Convolucionais**

Orientador: Prof. Dr. Allan Kardec Dualibe Barros Filho

Co-Orientador: Prof. Dr. André Borges Cavalcante

São Luís

2018

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).  
Núcleo Integrado de Bibliotecas/UFMA

de Matos Luna dos Santos, Daniel.  
Desenvolvimento de um Sistema de Reposicionamento para  
um Robô de Sondagem utilizando Redes Convolucionais /  
Daniel de Matos Luna dos Santos. - 2018.  
81 p.

Coorientador(a): André Borges Cavalcante.  
Orientador(a): Allan Kardec Duailibe Barros Filho.  
Dissertação (Mestrado) – Programa de Pós-graduação em  
Engenharia de Eletricidade/ccet, Universidade Federal do  
Maranhão, Faculdade de Engenharia Elétrica, 2018.

1. Automatização. 2. Banco de Dados. 3. Imagem. 4.  
Redes Convolucionais. 5. Reposicionamento. I. Borges  
Cavalcante, André. II. Kardec Duailibe Barros Filho,  
Allan. III. Título.

Daniel de Matos Luna dos Santos

**Desenvolvimento de um Sistema de Reposicionamento para um Robô de  
Sondagem utilizando Redes Convolucionais**

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica, ao Programa de Pós-Graduação em Engenharia Elétrica, da Universidade Federal do Maranhão. Área de concentração: Processamento da Informação Biológica.

Banca Examinadora:

---

Prof. Dr. Allan Kardec Duailibe Barros Filho (Orientador)  
Laboratório de Processamento da Informação Biológica da UFMA

---

Prof. Dr. Ewaldo Éder Carvalho Santana  
Faculdade de Engenharia Elétrica da UFMA

---

Prof. Dr. Raimundo Carlos Silvério Freire  
Universidade Federal de Campina Grande - UFCG - LIMC

---

Prof. Dr. André Borges Cavalcante (Co-Orientador)  
Faculdade de Engenharia Elétrica da UFMA

São Luís

2018

## **DEDICATÓRIA**

À Deus e à minha Família que fortemente me auxiliam nessa carreira.

## **AGRADECIMENTO**

Ao orientador Prof. Allan Kardec Barros pela oportunidade concedida e pela orientação no decorrer do mestrado.

Ao Prof. André Cavalcante pela Co-orientação e linha de pesquisa fornecidas durante o mestrado.

Ao Prof. Ewaldo Santana pelos ensinamentos e orientação concedida em Estatística e Processamento de Sinais.

Ao programa de Pós-Graduação em Engenharia de Eletricidade-UFMA.

Aos alunos do PIB (Laboratório de Processamento da Informação Biológica) pela carga de conhecimento compartilhada durante o mestrado.

Agradeço à CAPES pela bolsa de Mestrado concedida.

## RESUMO

As Redes Neuronais Convolucionais (do inglês *Convolutional Neural Networks-CNN*) têm sido utilizadas com sucesso para operações de alinhamento em plataformas autônomas, para ambientes cuja complexidade do cenário e os recursos de pré-processamento da imagem são fatores decisivos para o sucesso da classificação (atitudes de reposicionamento). O objetivo do presente estudo é desenvolver um sistema de aproximação autônomo com base na classificação de imagens por uma *CNN*. Os resultados obtidos mostram a *CNN* superior (Acurácia 82%) a um método que utiliza *Limiar de Decisão e Marcadores* (Acurácia 51,8%), desenvolvido para teste inicial do sistema de aproximação. A utilização da Rede Convolucional também implicou no desenvolvimento de uma ferramenta de geração do banco de dados em diferentes cenários nos quais foram gerados tanto sobre o modelo virtual do sistema mecânico, como sobre o modelo físico. Para a geração do banco de dados com base no modelo virtual e sua inserção diferentes cenários de operação foi utilizada uma técnica de processamento de imagem caracterizada como “Background Subtraction”, onde a partir de um limiar de controle, o objeto desejado (sistema de sondagem) foi extraído do “Background” (píxeis do cenário), sendo posteriormente inserido em outro “Background” (píxeis referentes ao cenário desejado), associado aos valores de ângulo do respectivo modelo. Na geração do banco de dados com o modelo físico foram utilizados sensores para aquisição de imagens e valores de ângulo de inclinação do sistema de sondagem. Os resultados finais obtidos contemplam uma ferramenta para geração de um banco de dados (aplicados à métodos de aprendizagem de máquina), e automatização do processo de reposicionamento.

Palavras-chave: Reposicionamento, Banco de Dados, Redes Convolucionais, Automatização, Imagem.

## ABSTRACT

Convolutional Neural Networks (*CNN*) has been successfully used for positioning operations on standalone platforms, for environments whose scenario complexity and image pre-processing capabilities are decisive factors for the success of the classification (repositioning attitudes). The objective of the present study is to develop an autonomous approximation system with the base classification of images by a *CNN*. The results show the superior *CNN* (accuracy 82%) to a method that uses Decision threshold and Markers (accuracy 51.8%), developed an initial test of the approach system. For the generation of the database based on the virtual model and its insertion different operating scenarios was used an image processing technique characterized as “*Background Subtraction*”, where from a control threshold, the desired object) was extracted from the “Background” (pixels of the scenario), and later inserted in another “*Background*” (pixels related to the desired scenario), associated with the angle values of the respective model. The final results obtained include a tool for generating a database (applied to machine learning methods) and automating the repositioning process.

Keywords: Repositioning, Database, Convolutional Networks, Automation, Image.

## LISTA DE FIGURAS

Figura 1	Diagrama do sistema de aproximação e cenários. Sistema mecânico caracterizado pela plataforma ( <i>Estrutura</i> ), sistema a ser reposicionado ( <i>Sonda</i> ) e seus possíveis ambientes de operação (Cenários de Operação) .....	16
Figura 2	Maquete virtual do sistema de posicionamento (esquerda); Modelo virtual em ambiente de operação (centro); Protótipo construído com implementação dos sensores de ângulo e de imagem (direita) .....	17
Figura 3	Robô de análise química do solo desenvolvido com uma sonda de 5 G.D.L (Graus de Liberdade). Fonte: Wet Chemistry Laboratory (NASA). ....	19
Figura 4	Modelo Virtual da Sonda: a) Modelo virtual com inclinação positiva (situação de levantamento); b) Modelo virtual com inclinação negativa (situação de reposicionamento). Fonte: Autor .....	22
Figura 5	Imagen da sonda virtual inserido em um cenário natural, no domínio espacial e no domínio da frequência: a) Inclinação positiva (situação de levantamento), com representação bidimensional e tridimensional na frequência; b) Inclinação negativa (situação de reposicionamento) e sua representação bidimensional e tridimensional na frequência. Fonte: Autor .	22
Figura 6	Imagen da sonda em comparação à uma imagem natural, no domínio espacial e no domínio da frequência: a) Imagem do sistema de sondagem em um cenário artificial (laboratório) com representação bidimensional e tridimensional na frequência; b) Imagem de um cenário Natural e sua representação bidimensional e tridimensional na frequência. Fonte: Autor .	23
Figura 7	O Robô (Plataforma Robótica), o sistema de sondagem (Sonda) e o seu ambiente de operação. Fonte: Autor.....	24
Figura 8	Ponto inicial de análise para o qual a sonda deve ser reposicionada. Fonte: Autor .....	24
Figura 9	Aproximação ao ponto desejado. Fonte: Autor .....	25
Figura 10	O inseto e seu aparelho lambedor, juntamente com seu sistema de visão. Fonte: Autor .....	27

Figura 11 Associação do sistema de visão e sondagem do inseto com os sistemas embarcados no robô. Fonte: Autor.....	27
Figura 12 Associação do sistema de visão e sondagem do inseto com os sistemas embarcados no robô. Fonte: Autor.....	28
Figura 13 Estrutura e fluxo de dados em uma RNA. Fonte: Adaptado de [1].....	31
Figura 14 Estrutura e fluxo de dados em uma RNA. Fonte: Adaptado de [1].....	33
Figura 15 Perceptron aplicado a mapeamento de sensores. Fonte: <a href="http://cse22-iiith.vlabs.ac.in/exp3/index.html">http://cse22-iiith.vlabs.ac.in/exp3/index.html</a> .....	35
Figura 16 Soluções encontradas por uma RNA tipo Perceptron para AND, OR e XOR. Fonte: Autor.....	36
Figura 17 Estrutura de aprendizado de uma RNA caracterizada como Adaline. Fonte: Autor.....	37
Figura 18 Imagem sendo processada por uma MLP. Fonte: Autor .....	39
Figura 19 Imagens associadas a dois tipos de classes diferentes. Fonte: Autor.....	40
Figura 20 Modo de treinamento de uma Rede Neural Convolucional. Fonte: Autor... .	42
Figura 21 Processo de convolução da imagem com um kernel tendo como saída um mapa de ativação. Fonte: Autor.....	43
Figura 22 Processo de formação do mapa de ativação. Fonte: Autor. ....	44
Figura 23 Processo de formação do mapa de ativação. Fonte: Autor. ....	45
Figura 24 Adição simétrica de zeros à matriz. Fonte: Autor. ....	45
Figura 25 Modelo Físico do Sistema de Sondagem. Fonte: Autor.....	47
Figura 26 Modelo Físico do Sistema de Sondagem. Fonte: Autor.....	48
Figura 27 Modelo Geométrico do Sistema de Sondagem. Fonte: Autor.....	48
Figura 28 Modelo Geométrico do Sistema de Sondagem (Correção das distâncias). Fonte: Autor.....	49
Figura 29 Triângulos do modelo geométrico no sistema de sondagem. Fonte: Autor. .	50
Figura 30 Triângulos do modelo geométrico no sistema de sondagem. Fonte: Autor. .	51
Figura 31 Modelo do Mecanismo de recolhimento e fornecimento do cabo. Fonte: Autor .....	54
Figura 32 Diagrama de medição do ângulo $\theta_{sonda(i)}$ . Fonte: Autor.....	56
Figura 33 Diagrama do Resistor Variável. Fonte: Autor.....	56

Figura 34 Gráfico do comprimento do cabo ( $L_{hip(j)}$ ) em função do ângulo de inclinação $\theta_{sonda}$ ) para os ângulos de $-45^\circ$ à $45^\circ$ . Fonte: Autor.....	58
Figura 35 Gráfico do comprimento $L_{hip(i)}$ do cabo em função da tensão $U_\theta$ variando entre 1 e 4,5 Volts. Fonte: Autor.....	58
Figura 36 Gráficos Relativos às aquisições do sinal de ângulo calculado a partir da medição de tensão no resistor variável. Operação de levantamento (a); Operação de levantamento e posterior abaixamento do sistema de sondagem (b). Fonte: Autor .....	59
Figura 37 Modo de Treinamento da Rede Neuronal Convolucional, a ser implementado no sistema de posicionamento. Fonte: Autor. ....	60
Figura 38 Ferramenta proposta para automatização do banco de dados de treinamento. Fonte: Autor. ....	61
Figura 39 Geração do banco de dados do sistema real de sondagem pela ferramenta proposta. Fonte: Autor. ....	63
Figura 40 Modelo virtual animado em plataforma virtual de desenvolvimento. Fonte: Autor .....	64
Figura 41 Interpolação bilinear aplicada a uma função $f(x,y)$ . Fonte:(Dourado.,2014).	65
Figura 42 Demonstração do “ <i>Chroma-Key</i> ” aplicado a cada quadro. Fonte: Autor ...	68
Figura 43 Gráfico de comparação entre o vetor de aprendizado gerado no teste real e o vetor de aprendizado gerado no modelo virtual. Fonte: Autor.....	68
Figura 44 Diagrama de funcionamento de sistema de aproximação. a) Método de aquisição e classificação para reposicionamento do sistema de sondagem ; b) Imagens das classes de inclinação e a descrição dos comandos de reposicionamento; c) Sistema Real de Aproximação. ....	69
Figura 45 Diagrama de blocos da <i>CNN</i> . Imagem do sistema mecânico real sendo convoluída com o filtros na camada convolucional <i>CONV</i> , retificada pela camada <i>RELU</i> e agrupada ou sub-amostrada na camada <i>POOL</i> , sendo gerado então um conjunto de valores caracterizado como <i>VETOR DE CARACTERÍSTICAS</i> inserido na rede <i>Multi-Layer-Perceptron</i> (MLP) tipo <i>Fully-Conected</i> . ....	70
Figura 46 Estrutura do Banco de Dados usado para o treinamento da rede Convolucional. ....	71

Figura 47 Diagrama dos Marcadores. a) Localização dos Marcadores; b) Área $A_1$ entre os marcadores que ultrapassa o limiar $\alpha$ (sonda distante do ponto desejado); c) Área $A_2$ entre os marcadores que não ultrapassa o limiar $\alpha$ (sonda estática no ponto desejado).....	72
Figura 48 Fluxograma do algoritmo de classificação. Fonte: Autor.....	73
Figura 49 Aquisição da imagem em formato RGB-8 bits. Fonte: Autor.....	75
Figura 50 Coordenadas de todos os píxeis de valor máximo de cada plano de cor. Fonte: Autor.....	75
Figura 51 Coordenadas de maior aglomeração dos píxeis de intensidade máxima. Fonte: Autor.....	76
Figura 52 Comparação das áreas para classes diferentes de posicionamento (1-esquerda e 0-direita). Fonte: Autor.....	76
Figura 53 Imagem real da sonda e dos marcadores (ao centro) e a câmera que irá captar a imagem dos marcadores (à direita). Fonte: Autor.....	77
Figura 54 Gráfico dos valores de área agrupados em dois vetores (pontos azuis e pontos pretos), cujos números (em vermelho) indicam: (1) Sonda no início da operação de içamento; (2) Posição cujos marcadores encontram-se nas extremidades da imagem; (3) Sonda fora da área contemplada pela imagem (situação de emergência); (4) Sonda estática na inclinação mínima (verificação da repetibilidade na classificação-pontos pretos).....	78
Figura 55 Geração do banco de dados do modelo virtual em cenário desejado. Fonte: Autor .....	80

## LISTA DE TABELAS

Tabela 1	Hiperparâmetros principais para configuração da <i>CNN</i> .....	46
Tabela 2	Características principais do sensor potenciométrico utilizado na medição de ângulo no hardware de aquisição de sinais. ....	59
Tabela 3	Características principais do sensor de imagem utilizado.....	62
Tabela 4	Comparação entre os dois classificadores pelas métricas de Acurácia, Pre- cisão e Sensitividade. Os verdadeiros-positivos correspondem à ( <i>tp</i> ), os verdadeiros-negativos à ( <i>tn</i> ), os falso-positivos à ( <i>fp</i> ) e falso-negativos à ( <i>fn</i> ).....	78

# SUMÁRIO

	<b>INTRODUÇÃO .....</b>	15
0.0.1	Motivação.....	15
0.1	Organização do Trabalho .....	17
1	<b>O PROBLEMA.....</b>	19
1.1	Posicionamento em Ambientes Hostis .....	19
1.2	Sistemas de Sondagem e Ambientes de Operação .....	21
1.3	Objetivos .....	24
1.4	Objetivos Específicos .....	25
2	<b>HIPÓTESES .....</b>	26
2.1	Inspiração Biológica.....	26
2.2	Extração de Padrões .....	28
3	<b>FUNDAMENTAÇÃO TEÓRICA .....</b>	30
3.1	Introdução à Redes Neuronais .....	30
3.2	Modelo Matemático da Rede MLP .....	31
3.3	Função de ativação.....	31
3.4	Função de custo-Problema de Otimização.....	33
3.5	Backpropagation .....	37
3.6	Problemas no Processamento de Imagens por Redes Clássicas .....	38
3.7	Redes Neuronais Convolucionais .....	40
4	<b>MODELO DO SISTEMA DE SONDAZEM.....</b>	47
5	<b>MÉTODOS IMPLEMENTADOS .....</b>	60
5.1	Geração do Banco de dados .....	60
5.2	Descrição da ferramenta para geração do Banco de Dados .....	61
5.2.1	Aquisição de sinais no modelo real .....	62
5.2.2	Modelo Virtual proposto para comparação .....	63
5.2.3	Método de Interpolação .....	64
5.2.4	Background Subtraction e Chroma-Key.....	66

5.2.5	Comparação dos Modelos Físico e Simulado .....	68
5.3	Materiais utilizados na Implementação Final .....	69
5.4	Método de classificação pela <i>CNN</i> .....	69
5.5	Método de classificação utilizando a CNN .....	70
5.5.1	Banco de Dados- Modelo Físico.....	71
5.6	Método de classificação utilizando Marcadores .....	72
6	<b>RESULTADOS .....</b>	75
	<b>CONCLUSÃO .....</b>	79
6.1	Contribuição Científica da Ferramenta para Geração do Banco de Dados ...	79
	<b>REFERÊNCIAS.....</b>	81

## INTRODUÇÃO

É crescente o uso de plataformas robóticas que possuem capacidade de readaptação à diferentes ambientes e aplicações, utilizando visão computacional. A readaptação dos algoritmos embarcados nestas plataformas se faz necessária devido à variação nas condições de luminosidade, nas formas e cores dos objetos de interesse, além das mudanças nas características do cenários que as cercam.

Umas das formas de contornar esses problemas é através da extração de características de uma imagem, como forma de reduzir a quantidade de informação a ser processada. Trabalhos que utilizam descritores de imagem como *contorno*, *cor* e *limiar* por Pereira (2014) [2], *Wavelets* por Haware (2007) [3], PCA (*Dependent Component Analysis*) por Xiao (2010) [4] ou ICA (*Independent Component Analysis*) por Liao (2008) [5], para redução de dimensionalidade ou extração de características. O posterior uso de algoritmos de aprendizado de máquina (ANN- *Artificial Neural Networks* ou sistemas *Neuro-Fuzzy*), tornam os algoritmos readaptáveis à diferentes aplicações. Algumas dessas técnicas [2] não são invariantes à translação de objetos na imagem, ruídos ou variações de luminosidade. Em contrapartida métodos utilizados nos trabalhos de Jozias et al (2006) [6] [7] fazem uso de distorção de cromaticidade e variação da distorção do brilho em imagens para segmentação em “*Foregrounds*”(Objetos) e “*Backgrounds*”(Cenários) estáticos, reduzindo os efeitos da variação de luminosidade, focos de luz e sombras nos cenários testados pelos autores.

No entanto, como mostrado por Goodfellow (2016) [8], Giusti et al (2014) [9] e Fares et al (2016) [10] as *CNN* (classe específica de Redes Neurais Profundas) possuem a capacidade de extração de características de cor e forma, invariantes à translação, ruídos ou variações de luminosidade, problema comum ao processamento de imagem [11].

### 0.0.1 Motivação

O desenvolvimento de uma plataforma robótica para fins de análise de sedimentos superficiais em solos (Páginas 16 e 17) gerou a necessidade de uma sistema de aproximação autônomo, que deveria realizar operações de aproximação da sonda (Figura 1) um objeto desejado, motivando o desenvolvimento de um sistema aproximação, para uso em diferentes ambientes de operação, que permitisse à adaptação em diferentes.

O processamento de imagem foi escolhido devido ao número de informações contidas na imagem adquirida (sistema de sondagem, ambiente de operação e ponto desejado para o reposicionamento). A utilização do aprendizado de máquina foi motivada devido a variação das estatísticas das imagens adquiridas (objeto artificial em meio a cenários naturais com variação de cor, forma e estatísticas dos pontos para os quais a sonda deveria ser reposicionada) [12].

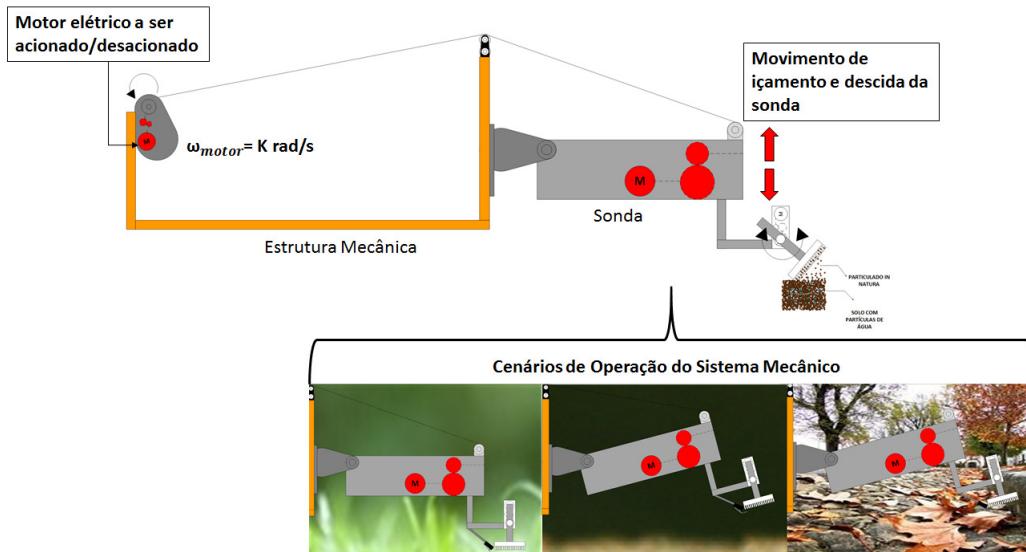


Figura 1 Diagrama do sistema de aproximação e cenários. Sistema mecânico caracterizado pela plataforma (*Estrutura*), sistema a ser reposicionado (*Sonda*) e seus possíveis ambientes de operação (Cenários de Operação).

Pode ser observado na Figura 1 a estrutura mecânica de suporte ao sistema de sondagem em uma estrutura basculante. Através do acionamento do motor elétrico, um sistema de recolhimento ou fornecimento de cabo realiza o processo de afastamento ou aproximação da extremidade do sistema de sondagem ao local desejado. Isso deve ocorrer de forma autônoma, em diferentes ambientes (Cenários de Operação).

O sistema de sondagem tratado no presente trabalho foi desenvolvido tanto em maquete virtual, demonstrado no ambiente de operação, como desenvolvido fisicamente. A partir do modelo virtual, foi realizada a prototipagem com implementação de sensores e algoritmos de controle de posicionamento que resultaram na necessidade de um posicionamento robusto à mudança do ambiente, autônomo, com base em um sensor de imagem.

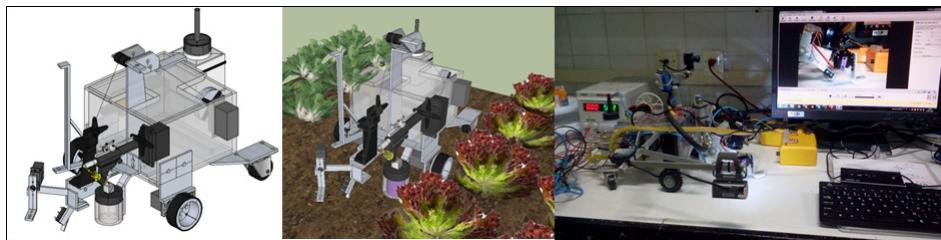


Figura 2 Maquete virtual do sistema de posicionamento (esquerda); Modelo virtual em ambiente de operação (centro); Protótipo construído com implementação dos sensores de ângulo e de imagem (direita)

Pode ser observado na parte esquerda da Figura 2 a representação tridimensional da plataforma juntamente com o sistema de sondagem modelados em software “*CAD*”(Computer Aided Design), posteriormente inserida em um de seus possíveis ambientes de operação (cenários simulado). Na extremidade direita da Figura 2, pode ser observado o modelo físico da plataforma e do sistema de sondagem com os sistemas de aquisição de imagem (instalado na parte lateral da sonda).

A partir deste protótipo surge a necessidade do posicionamento autônomo, para os diferentes cenários ou locais onde o sistema de sondagem deva operar, independentemente de qual objeto ou ponto para onde esta deva ser reposicionada. O problema do posicionamento autônomo será tratado em mais detalhes no capítulo 2.

### 0.1 Organização do Trabalho

Inicialmente no capítulo 1 tem-se a descrição do problema do posicionamento em ambientes naturais, juntamente com as estatísticas das imagens naturais e artificiais, assim como imagens artificiais inseridas em um cenário natural (representação no espaço e na frequência). A partir do problema do posicionamento autônomo do sistema de sondagem em ambientes naturais, são geradas hipóteses no Capítulo 2 com base na inspiração biológica do robô, e proposta a extração de padrões como técnica para tentar resolver o problema em questão. No capítulo 3 são fundamentadas as redes Neuronais clássicas e posteriormente o funcionamento das camadas convolucionais.

No capítulo 4 é exposto o modelo de posicionamento do sistema de sondagem, assim como as medições de ângulo de inclinação realizados através da implementação de um resistor variável em configuração de divisor de tensão. São modeladas as equações que regem o movimento do sistema de içamento, dependentes de variáveis elétricas do motor

de acionamento.

No capítulo 5 são expostos os métodos utilizados para geração do banco de dados, com base em um modelo de sondagem simulado e a geração do banco de dados com o modelo físico, implementados os *hardwares* necessários. É descrito o método da Rede Convolucional, comparado a um método de processamento de imagem que tem como base o processamento de imagem com o uso de marcadores de frequência específica, extraindo características geométricas e de cor.

Por fim são apresentados os resultados obtidos, realizada a comparação dos métodos e discutidos suas implementações.

## 1 O PROBLEMA

### 1.1 Posicionamento em Ambientes Hostis

O posicionamento autônomo de plataformas mecânicas em ambientes naturais é esforço que tem sido realizado para avanço nas áreas da agricultura, uso militar, aplicações em resgate, sondagem aeroespacial, cuja quantidade de variáveis de cenário e de pontos alvo tornam o processamento de informações uma tarefa complexa. Trabalhos como o de [9] sobre posicionamento autônomo de *Micro UAVs- Micro Unmanned Aerial Vehicles* (Microaeronaves não tripuladas) em florestas densas, ou plataformas como *R.C.A.L- Robotic Chemical Analisys Laboratory* (Laboratório Robótico de Análises Químicas), foram desenvolvidos tendo como um dos principais objetivos o posicionamento autônomo em ambientes naturais (Figura 3).

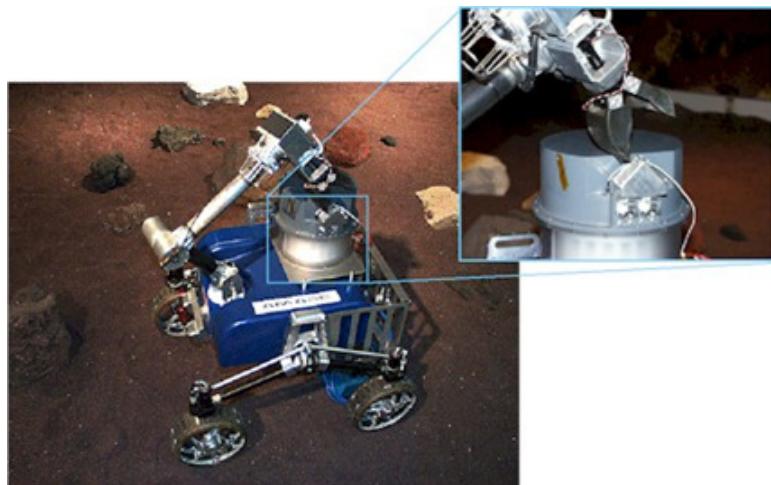


Figura 3 Robô de análise química do solo desenvolvido com uma sonda de 5 G.D.L (Graus de Liberdade). Fonte: Wet Chemistry Laboratory (NASA).

Pode ser observado na Figura 3 o braço robótico com 5 *DF* (*Degrees of Freedom*- Graus de Liberdade), extraíndo amostras em meio à rochas e solo arenoso. Para a sondagem autônoma, é necessário identificar o objeto, assim como sua forma (altura, largura e profundidade), em qual posição este se encontra em relação ao braço mecânico e por fim, formar a estratégia de reposicionamento.

Pode ser observado ainda na Figura 3 que, o sistema de análise propriamente dito ou *R.C.A.L* está instalado sobre um *SRR* (Sample Rover Return- Rover de Retorno de Amostra), desenvolvido para servir como plataforma de testes em análises químicas e

robô de treinamento. A captura de sedimentos superficiais do solo é realizada pelo braço robótico e deve ser levada até a entrada do sistema de análise (canto superior direito da Figura 3).

Fica evidente, devido ao grande número de dados a serem adquiridos, a necessidade de sistemas de visão computacional, para aquisição de imagens referentes a posição do braço e da posição do objeto, assim como outras características contidas no cenário.

Para o posicionamento autônomo em ambientes ou cenários naturais surgem problemas como a variação de luminosidade, de forma dos objetos, dos tipos de cenário (natural ou artificial), imperfeições no pavimento onde estão localizados o robô e o objeto (variações de cor, forma e textura dos alvos a serem alcançados), além dos problemas inerentes à aquisição de dados (ruído nas imagens adquiridas e vibrações).

Uma das formas de análise da imagem é a sua visualização na frequência, através da *DFT-Discrete Fourier Transform* (Transformada Discreta de Fourier) em duas dimensões. A equação 1.1 expressa a *DFT* em um domínio espacial é dada como:

$$f(K_x, K_y) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(\frac{k_x}{N} + \frac{K_y}{N})} \quad (1.1)$$

sendo  $(k_x, K_y)$  os valores espaciais de frequência,  $x, y$  a posição espacial do pixel na imagem,  $f(x, y)$  a imagem de entrada a ser decomposta,  $N$  a dimensão da imagem (considera-se uma imagem  $N \times N$ ) e  $j$  um número complexo do tipo  $j^2 = -1$ .

A *DFT* descrita na Eq. 1.1 de imagens naturais, possui como característica principal, picos mais suaves ou menos pronunciados em relação a imagens ou objetos artificiais, como pode ser confirmado no trabalho de [12]. Neste trabalho foram analisados 12.000 espectros de potência média de imagens, sendo 6.000 de imagens naturais e 6.000 imagens artificiais, observando que imagens naturais possuem, a curva de potência do sinal em função da frequência, seguindo a forma  $\frac{1}{f^\alpha}$  com  $\alpha \sim 2$  ou  $\alpha \sim 1$  (levando em conta a amplitude do espectro). Para observação da variação das formas da *DFT*, foi adquirida uma imagem da sonda extraída em uma ambiente artificial e uma imagem de um cenário natural (Figura 4). É possível observar nitidamente a mudança na forma do espectro.

A representação do espectro de potência parte do valor absoluto de  $f(K_x, K_y)$  (imagem no domínio da frequência), é dado por:

$$\|f(K_x, K_y)\|^2 = Re((f(K_x, K_y)))^2 + Im(f(K_x, K_y))^2 \quad (1.2)$$

sendo  $Re((f(K_x, K_y)))$  as componentes reais e  $Im(f(K_x, K_y))$  as componentes imaginárias da DFT.

Levando a definição do espectro de potência para coordenadas polares, seu valor esperado gera um modelo mais completo do espectro de potência segundo (Torralba et al, 2003) [12] dado por:

$$E[\|I(f, \theta)\|^2] \simeq \frac{A_s(\theta)}{f^{\alpha_{s(\theta)}}} \quad (1.3)$$

sendo  $A_s(\theta)$  a amplitude como um fator de escala para cada orientação  $\theta$  e  $\alpha_s(\theta)$  o expoente da frequência como um fator da orientação.

O modelo descrito pela equação 1.3, descrita no trabalho de (Torralba et al, 2003) [12] se torna necessário quando são considerados ambos espectro de potência média de imagens naturais e *man-made scenes images* (imagens de cenários construídos pelo homem ou objetos fabricados pelo mesmo). É possível observar portanto, para cada tipo de cenário ou objeto a ser localizado na imagem, que a forma ou orientação do espectro de frequências possuirá variações.

Para que o sistema de sondagem proposto consiga ser reposicionado autonomamente no conjunto de cenários cujas estatísticas são aproximadas pela Eq.1.3, o método que será desenvolvido deverá ser portanto robusto à variações de cenário e do ponto, objeto ou alvo para o qual o robô deverá ser reposicionado.

## 1.2 Sistemas de Sondagem e Ambientes de Operação

O problema de posicionamento de objetos artificiais em meio a cenários naturais se torna complexo pois contempla as estatísticas de ambos os tipos de imagens (imagens artificiais inseridas em cenários naturais e /ou objetos artificiais sendo reposicionados até pontos que possuem ambas as estatísticas). Para que isso pudesse ser observado, foi desenvolvido um modelo virtual do sistema de sondagem, sendo este posteriormente inserido em um cenário natural (Figura 4 e Figura 5).

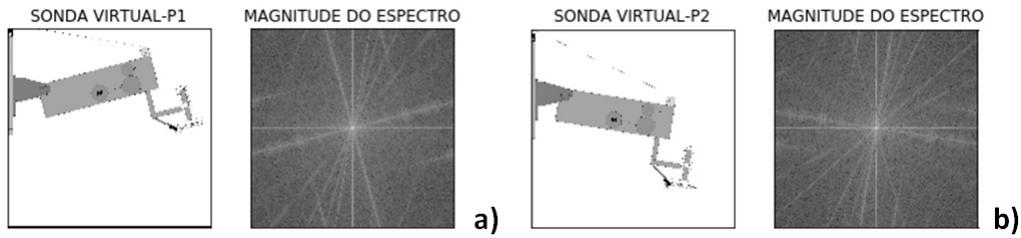


Figura 4 Modelo Virtual da Sonda: a) Modelo virtual com inclinação positiva (situação de levantamento); b) Modelo virtual com inclinação negativa (situação de reposicionamento).  
Fonte: Autor

Pode ser observado na Figura 4 o deslocamento de algumas frequências (Figura 4 a e b) com a mudança de inclinação do modelo virtual da sonda. O modelo virtual da sonda em um cenário ou “Background” natural e sua nova distribuição na frequência podem ser observados na Figura 5:

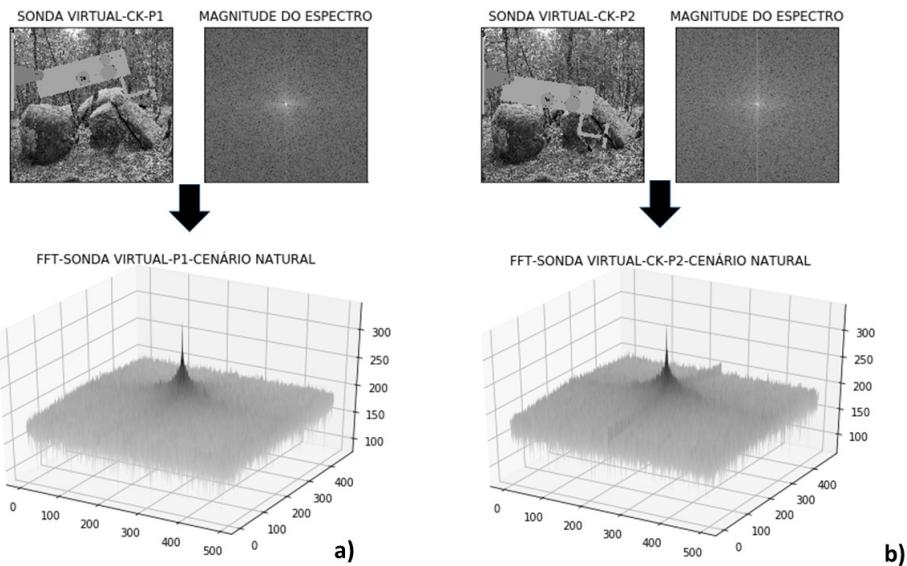


Figura 5 Imagem da sonda virtual inserido em um cenário natural, no domínio espacial e no domínio da frequência: a) Inclinação positiva (situação de levantamento), com representação bidimensional e tridimensional na frequência; b) Inclinação negativa (situação de reposicionamento) e sua representação bidimensional e tridimensional na frequência.  
Fonte: Autor

As imagens do sistema de sondagem real e de um cenário natural possuem as seguintes representações na frequência:

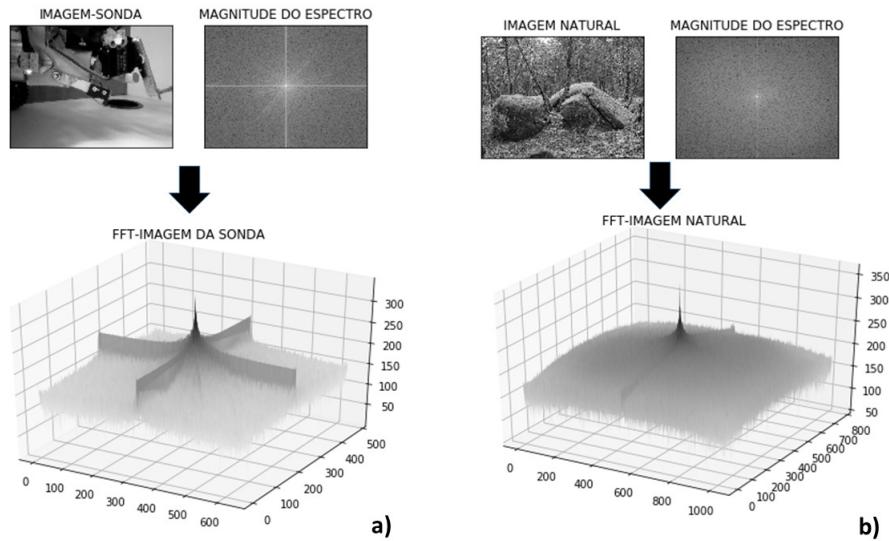


Figura 6 Imagem da sonda em comparação à uma imagem natural, no domínio espacial e no domínio da frequência: a) Imagem do sistema de sondagem em um cenário artificial (laboratório) com representação bidimensional e tridimensional na frequência; b) Imagem de um cenário Natural e sua representação bidimensional e tridimensional na frequência.  
Fonte: Autor

O sistema de sondagem (objeto artificial), localizado tanto em cenários cenários naturais (paisagem ao fundo nas imagens da parte superior da Figura 5 a e b), que deve ser reposicionado em direção à um ponto ou objeto desejado no cenário (objeto este podendo pertencer tanto a classe de objetos naturais como à de objetos *man-made*-fabricados ou modificados pelo ser humano). As suas respectivas representações na frequência mostram que a diferença entre imagens artificiais em cenários naturais (Figura 4) é mais pronunciada do que entre imagens da sonda, em diferentes posições, nos mesmos cenários.

O problema do posicionamento autônomo não envolve somente o tipo de cenário, mas também os objetos ou pontos-alvo para os quais a sonda deve ser reposicionada. As modelagens anteriores são também aplicadas ao processamento do objeto ou alvo desejado dentro da imagem. O problema em estudo pode ser melhor observado na Figura 7:

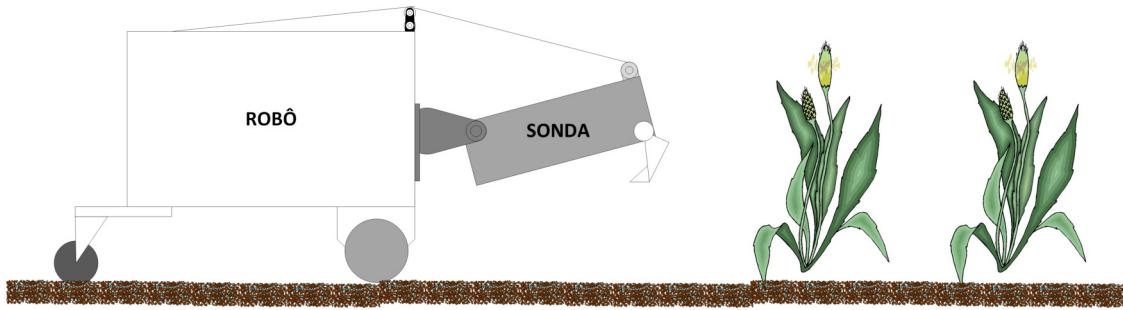


Figura 7 O Robô (Plataforma Robótica), o sistema de sondagem (Sonda) e o seu ambiente de operação. Fonte: Autor

A finalidade da aproximação autônoma foi gerada pois a plataforma robótica, assim como o robô apresentado na Figura 3, foram desenvolvidos para finalidade de sondagem ou análise superficial do solo. Isso deve ser feito de forma autônoma em um conjunto de ambientes possíveis, nas superfícies de objetos ou solos (pavimentos) onde o robô será operado. A representação do ponto de análise para onde o sistema deveria ser reposicionado pode ser observado na Figura 8:

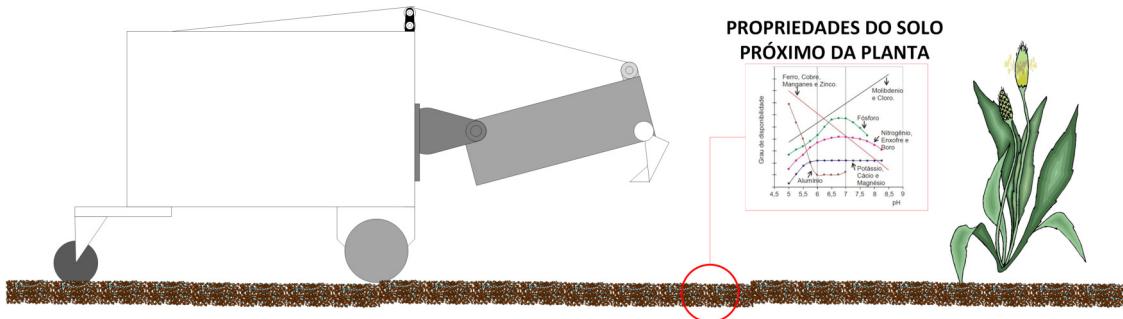


Figura 8 Ponto inicial de análise para o qual a sonda deve ser reposicionada. Fonte: Autor

A finalidade de análise do solo não será discutido no presente estudo, devido ao problema não se limitar somente à um tipo de finalidade de posicionamento. A finalidade para o qual o robô foi desenvolvido está sendo demonstrada somente para esclarecimentos quanto ao ponto para o qual o robô deverá ser reposicionado (características do ponto ou do objeto).

### 1.3 Objetivos

O objetivo principal do presente estudo é o reposicionamento de um sistema de sondagem a um ponto desejado em um conjunto de ambientes em que a mesma esteja operando. Isso pode ser melhor observado na Figura 9:

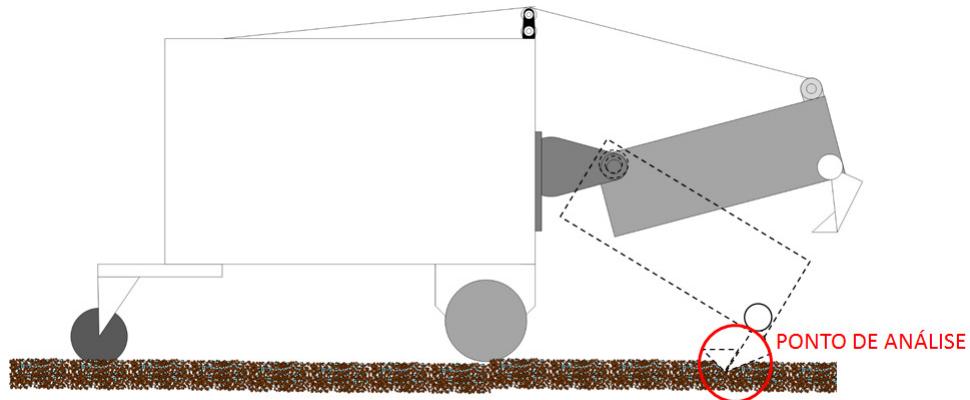


Figura 9 Aproximação ao ponto desejado. Fonte: Autor

Para alcançar o objetivo principal, é proposto o uso da imagem do sistema de sondagem (inserida no conjunto de cenários desejados), devido ao número de informações acerca do cenário e do ponto desejado (Ponto de Análise- Figura 9).

#### 1.4 Objetivos Específicos

1. Desenvolvimento dos *hardwares* de aquisição de sinais com posterior transferência de informação para o servidor de processamento;
2. Desenvolvimento do modelo de posicionamento;
3. Desenvolvimento do modelo virtual da sonda em diferentes cenários (montagem de um banco de dados);
4. Geração de um banco de dados com imagens do sistema de reposicionamento virtual em cenários naturais;
5. Geração de um banco de dados com imagens do sistema de reposicionamento real;
6. Desenvolvimento dos métodos para o processamento das imagens e reposicionamento autônomo com base nos dados gerados;
7. Desenvolvimento do hardware de ativação para o sistema mecânico real com base nos resultados do processamento da imagem adquirida;
8. Teste e análise dos resultados para refino do algoritmo de reposicionamento.

## 2 HIPÓTESES

### 2.1 Inspiração Biológica

Diante do problema da mudança de cenários em que o robô está inserido e mudança dos objetos para os quais o sistema de sondagem deve ser reposicionado, foi então pesquisado soluções encontradas pela natureza para o posicionamento e operações de sondagem realizadas em ambientes naturais.

Foi então encontrado nas abelhas do tipo *Honeybee* tanto o aparelho bucal do tipo lambedor, no qual foi inspirado o sistema de sondagem, quanto o sistema de visão, que levou ao uso de visão computacional, aprendizado de máquina (para extração de padrões e redução de dimensionalidade) e classificação com base na extração de características de cor localizadas em um ponto do espaço ou da imagem (classificador com base em limiar tratado no capítulo 6 deste trabalho).

As abelhas pertencem a ordem *Hymenoptera*, da superfamília *Apoidea*, do grupos do Apiformes, com mais de 20.000 espécies sendo uma das principais a *Apis mellifera*. O sistema de visão deste tipo de abelha tem sido o mais estudado entre os animais, ficando atrás somente dos primatas [13]. Através de experimentos de condicionamento por recompensa [13], foi evidenciado que as abelhas do tipo *Apis mellifera* possuem um sistema de visão colorido tricromático, que compartilha propriedades fundamentais com o sistema de visão dos humanos e primatas, como constância de cor(habilidade de reconhecimento de cor independente do nível de luminosidade à que os objetos estão expostos), oponência de cor (formação e compreensão da cor através de cores oponentes), segregação de cor (reconhecimento e separação de cor) e codificação de brilho [13].

O sistema de visão das abelhas é formado por um conjunto de partes menores, caracterizadas como omatídeos, agrupados para formar um sistema de visão composto, cuja imagem é formada semelhantemente à um mosaico.

A visão colorida das abelhas é caracterizada por 3 tipos de fotoreceptores, com pico de absorção nas partes UV, Azul e Verde no espectro. Os fotoreceptores são caracterizados como *S* (short-wavelength sensitive,  $\lambda_{max} = 344nm$ ), *M* (middle-wavelength sensitive,  $\lambda_{max} = 436nm$ ) e *L* (long-wavelength sensitive,  $\lambda_{max} = 544nm$ ). A quantidade de fotoreceptores pode variar em cada animal.

Após os sinais espectrais, mecanismos de pós-recepção irão realizar uma codificação

de aspectos cromáticos [13] e [14]. Os mecanismos de pós-recepção podem funcionar por oponência de cor (mecanismo onde a subtração dos estímulos resulta na informação de cor), mecanismos de codificação de aspectos cromáticos de estímulos coloridos e mecanismos não-oponentes de codificação de aspectos acromáticos (codificação de brilho). Essas características foram levadas em conta no processo de escolha do método de reposicionamento, obtendo ganhos conceituais (redução do número de Graus de Liberdade e uso de extração de características de cor para localização de pontos) no processo de reposicionamento através da imagem adquirida.

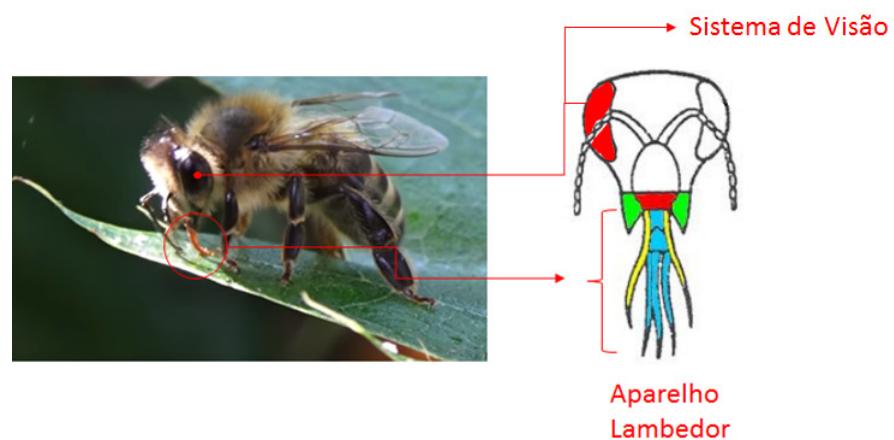


Figura 10 O inseto e seu aparelho lambedor, juntamente com seu sistema de visão. Fonte: Autor.

Para o reposicionamento autônomo é então proposta a seguinte relação dos sistemas de visão e aparelho lambedor do inseto para a automatização do reposicionamento:

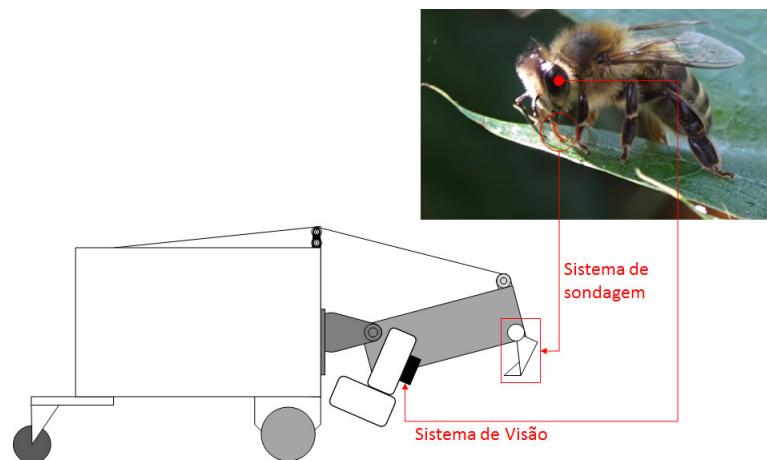


Figura 11 Associação do sistema de visão e sondagem do inseto com os sistemas embarcados no robô. Fonte: Autor.

O sistema proposto tem como princípio de funcionamento a imagem da sonda no

ambiente onde a mesma irá operar, juntamente com o ponto de análise para onde o sistema deve ser reposicionado. A hipótese é que com o processamento da imagem do sistema de sondagem inserido no cenário, estando contido o objeto ou ponto para onde a sonda deve ser reposicionada, pode-se inferir atitudes para o sistema de reposicionamento.

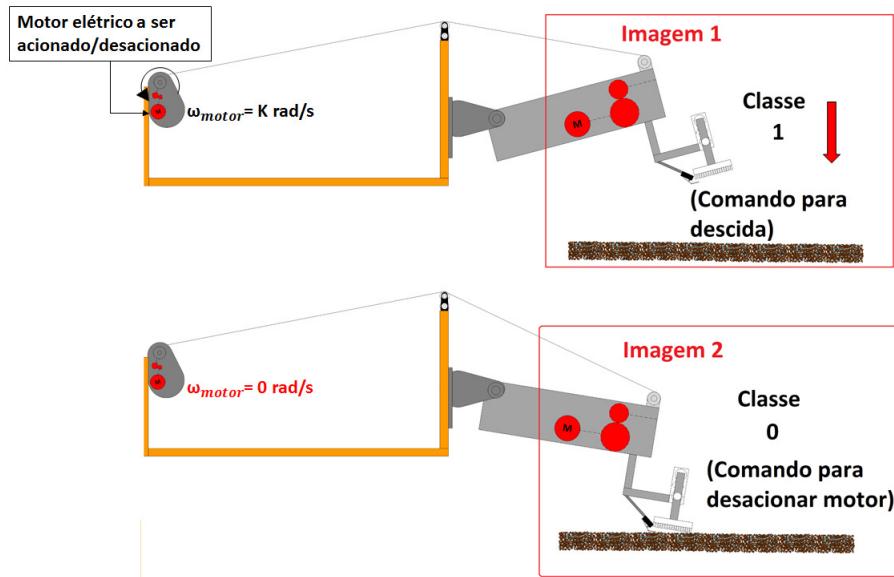


Figura 12 Associação do sistema de visão e sondagem do inseto com os sistemas embarcados no robô. Fonte: Autor.

A proposta para o método de reposicionamento (Figura 12) parte da hipótese de que, se medidos graus ou níveis de afastamento da extremidade da sonda ao objeto ou ponto desejado, pode-se gerar comandos para o reposicionamento da sonda. Para o presente estudo foram propostos dois níveis de grau de afastamento. O primeiro corresponde à classe 1 (grau de afastamento onde motor elétrico do sistema de sondagem será acionado para o reposicionamento da sonda). O segundo corresponde à classe 0 (grau de afastamento onde o motor elétrico deve ser desacionado, devido a proximidade com o ponto desejado).

## 2.2 Extração de Padrões

Com base no problema da inserção de uma imagem artificial (Sonda), em um conjunto de ambientes naturais (Cenários), assim como a mudança dos objetos ou pontos para os quais a sonda deve ser reposicionada (Ponto de Análise), tratados na Introdução do presente estudo, supõe-se que a extração de padrões ou características para a posterior classificação, consiga abstrair informações mais profundas da imagem, tornando mais

robusto o processo de classificação para o problema em questão.

Um padrão, em reconhecimento de imagens pode ser definido como qualquer elemento que pode ser definido quantitativamente, mesmo estando sujeito à variações [11]. A extração de características em geral é realizada por descritores, como descritores de forma, descritores de cor, de dimensões, de textura, entre outros, adequados a cada problema [11]. Para a variação de cenários, considerando um conjunto de ambientes naturais, pode-se utilizar os descritores citados, porém, para os objetos ou pontos que podem variar, propõe-se um método que aprenda as características do ponto ou objeto para onde a sonda deve ser reposicionada.

No trabalho de [9] foram utilizadas Redes Neurais Convolucionais para classificação de atitude de correção a ser tomada por um Micro-UAV- *Micro Unmanned Aerial Vehicle* em meio a trilhas abertas manualmente em florestas densas através da imagem adquirida por uma câmera disposta frontalmente na aeronave. O uso deste tipo de classificador foi justificado pela grande variação nos cenários (cenários exclusivamente naturais), cuja direção desejada tem como base trilhas sujeitas à intempéries (ambientes naturais modificados pelo homem sujeitos à processos de erosão, lixiviação, cobertura por folhagem da vegetação local, processos estes que mudam as estatísticas das imagens coletadas). Isso torna a navegação autônoma de plataformas robóticas uma tarefa complicada, cuja extração de características por descritores comuns não garante a correta classificação de orientação.

A semelhança dos problemas de navegação autônoma encontrados na literatura com o problema em estudo, assim como a sua resolução através do uso de Redes Neurais Convolucionais para obtenção de melhores performances de classificação motivam sua proposição como método de classificação principal proposto neste trabalho.

### 3 FUNDAMENTAÇÃO TEÓRICA

#### 3.1 Introdução à Redes Neuronais

As redes Neuronais Artificiais tem sua história iniciada quando um ramo da psicologia caracterizado como conexionismo, tenta descrever matematicamente a forma de aprendizado do cérebro animal, tendo como referência a estrutura básica da massa encefálica, o neurônio. O primeiro modelo artificial de um neurônio natural foi desenvolvido por Warren McCulloch e Walter Pitts em 1943.

Em 1949 Donald Hebb mostrou que a capacidade de aprendizado da rede pode ser aprimorada pelo ajuste dos pesos de entrada dos neurônios, propondo que o aprendizado dos neurônios poderia ser realizado pelo reforço das ligações sinápticas entre os neurônios excitados pela camada de entrada. Em 1958 Rosenblatt propôs uma rede (*Perceptron*) com base na estrutura do neurônios proposto por McCulloch e Pitts onde, através do ajuste dos pesos sinápticos ocorreria o aprendizado de certos tipos de padrões.

Nos trabalhos de Widrow e Hoff em 1960 foi proposta uma regra de aprendizado cujos os valores dos pesos seriam proporcionais a sua entrada, o erro entre o valor desejado na saída da rede e o valor apresentado pela mesma e uma constante (caracterizada como passo de aprendizado). As primeiras redes sofreram problemas quanto a classificação de padrões não-linearmente-separáveis, resolvido através do uso de redes *MLP* (*Multi-Layer Perceptron*) ou *Perceptron* de múltiplas camadas e do algoritmo de *backpropagation*.

As Redes Neuronais Artificiais tem se mostrado robustas à várias aplicações práticas como cancelamento de eco, reconhecimento de padrões, indo até extração de padrões de uma grande quantidade de imagens [1] [15] [16] [8]. A rede neuronal, pela sua constituição matemática, pode ser vista como um dispositivo de computação paralela, haja vista que um neurônio trata vários sinais paralelamente, gerando um efeito individual entradas-saída [16]. Uma RNA é formada pelos seguintes componentes mostrados na Figura 13:

A estrutura básica de um neurônio artificial observado na Figura 13 se resume à um vetor  $\mathbf{x}$  de entrada (sinal adquirido), combinado linearmente com um vetor de pesos  $\mathbf{w}^T$  (produto escalar). Após a combinação linear, o escalar  $y_{saída}$  é inserido em uma função de ativação (sigmóide, lógica, softmax, etc.), dimensionada para o problema em questão (em geral também utilizadas para limitar os valores de saída em uma faixa de valores desejada). A saída observada na Figura 13 possue valores  $S \in \mathbf{R}[0, 1]$ .

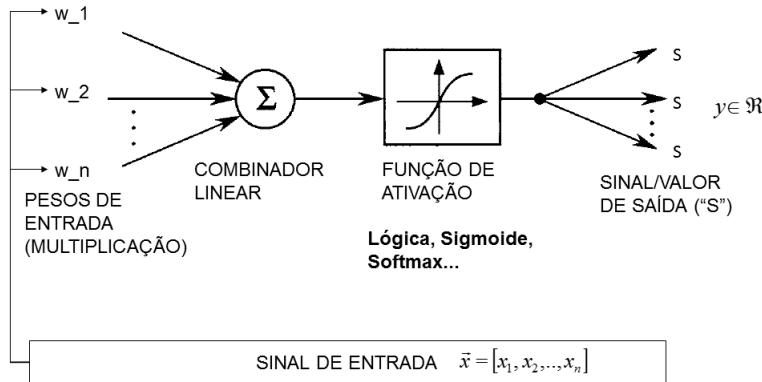


Figura 13 Estrutura e fluxo de dados em uma RNA. Fonte: Adaptado de [1]

As estruturas de rede dependerão da estrutura básica de cada neurônio. No presente trabalho serão tratadas as redes "ADALINE" e as Redes Perceptron-Multi-Camadas ou "Multi-Layer Perceptron-MLP".

### 3.2 Modelo Matemático da Rede MLP

A equação geral do Combinador Linear fica:

$$y_{saída} = \sum_{i=1}^n w_i x_i \quad (3.1)$$

sendo  $w_i$  os respectivos pesos do Combinador Linear e  $x_i$  os valores de entrada. Essa combinação pode assumir a forma do seguinte produto interno:

$$y_{saída} = \mathbf{x}^T \mathbf{w} \quad (3.2)$$

sendo  $\mathbf{x} = [x_1, \dots, x_n]^T$  o vetor com os valores de entrada e  $\mathbf{w}$  o vetor de pesos.

### 3.3 Função de ativação

A função de ativação, a ser utilizada neste trabalho é a função sigmóide, devido a sua flexibilidade quanto a mudança de parâmetros, gerando robustez a uma variedade maior de problemas. A função sigmoidal é representada a seguir:

$$S = \frac{1}{1 + e^{(-1)\lambda(y)}} \quad (3.3)$$

sendo  $S$  o valor de saída da função de ativação,  $\lambda$  um parâmetro ajustável e  $e$  o número de Euler.

A equação 3.3 caracteriza uma  $f_{(y)} = S$ , sendo  $y \in \mathbf{R}$  a saída do Combinador Linear, sendo  $\lambda$  uma constante que pode ser ajustável de acordo com a curva que se pretende obter. Outra característica importante da função sigmoidal está na simplicidade das derivadas, (devido a baixa complexidade de sua derivada e a existência dessa em ordens muito altas), tornam possíveis o algoritmo de retro-propagação ou “*backpropagation*” para atualização os pesos da rede neuronal artificial em suas várias camadas. A função sigmóide também pode ser modificada em outras partes, como numerador ou denominador, de acordo com a necessidade dos parâmetros de ativação. A função logística, a qual possui essa flexibilidade é mostrada na equação a seguir:

$$y = \frac{L}{1 + e^{-k(x-x_o)}} \quad (3.4)$$

sendo  $x_o$  o valor de  $x$  no ponto médio da curva sigmóide,  $L$  o valor máximo da curva e  $k$  a inclinação da curva,

Os valores máximos e mínimos da função são atingidos em  $x \in [+\infty, -\infty]$ .

A partir da função de saída da rede são comparados os valores desejados dos rótulos contidos no banco de dados.

O objetivo de uma rede neural, dado um mapa do tipo  $F : R^n \rightarrow R$  é aproximá-lo com o mínimo de erro, dado uma entrada  $x \in R^n$ . Para que seja realizado o treinamento são inseridos pares de treinamento do tipo  $(x_{d,1}, y_{d,1}), \dots, (x_{d,p}, y_{d,p})$  onde  $x_{d,1} \in \mathbf{R}^n$  e  $y_{d,1} \in \mathbf{R}$ . Desta forma o problema do aprendizado se resume à minimização da seguinte função:

$$J = \min \left[ \sum_{i=1}^p (y_{d,i} - x_{d,i}^n) \right] \quad (3.5)$$

Após o treinamento, uma rede neuronal adquire características de uma rede *feed-forward* (fluxo de dados em um único sentido não havendo retro-propagação- uso da matriz de pesos aprendida). Desta forma o neurônio seguinte recebe informação somente da camada anterior. A constituição de uma rede RNA *feed-forward* pode ser observada na Figura 14:

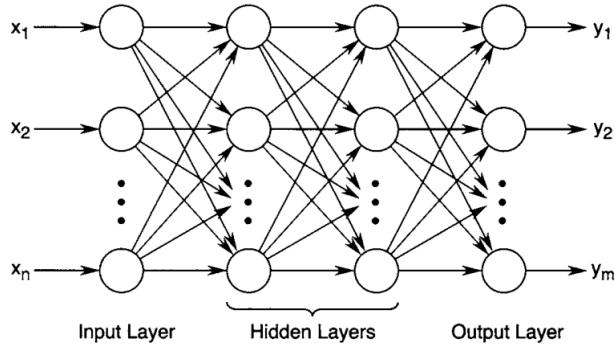


Figura 14 Estrutura e fluxo de dados em uma RNA. Fonte: Adaptado de [1]

Em uma rede Neuronal de múltiplas camadas, como pode ser observado na Figura 14, a primeira camada é definida como camada de entrada (input layer). A camada que possui a última combinação linear é caracterizada como camada de saída (output layer) e as camadas entre elas são caracterizadas como camadas escondidas (hidden layer).

O mapa formado pelos pesos de uma rede neuronal pode ser modificado pelo reajuste dos pesos na etapa de treinamento. A informação acerca do mapa está contida nos pesos sobre todos os neurônios (caracterizada como uma representação distribuída da função  $\mathbf{F}$ ). O mapeamento de uma função (muitas vezes desconhecida em um hiperplano) por uma rede neuronal, é caracterizado como um problema de aproximação, assim como o reconhecimento de padrões.

A capacidade de generalização de uma rede neuronal é verificada após o treinamento, onde uma característica intrínseca de um conjunto de dados é aprendida, de forma que, para valores de entrada que não estejam no conjunto de aprendizado, os valores de saída sejam muito próximos à um conjunto de valores desejados (predição), ou o valor de saída seja um número inteiro  $y_i \in \mathbb{N}$  correspondente à uma "Classe" desejada (Classificação). A forma de aprendizado das redes neurais utilizadas neste trabalho é do tipo supervisionado ou "Supervised Learning", ou seja, para cada entrada do conjunto de treinamento existe um valor desejado de saída (fazendo parte do mesmo conjunto).

### 3.4 Função de custo-Problema de Otimização

Considerando a forma matricial exposta anteriormente, o problema de otimização se resume à:

$$J = \min \|y_{d,i} - \mathbf{x}_{d,i}^T \mathbf{w}\| \quad (3.6)$$

São considerados dois tipos de problemas de otimização ( $p \leq n$  e  $p < n$  respectivamente). No primeiro caso tem-se o número de pares entrada-saída (conjunto de treinamento) menores ou iguais ao conjunto de pesos. Por conveniência é assumido que  $\mathbf{X}_d^T \in \mathbf{R}^n \rightarrow n = p$ . Para este caso há um número infinito de pontos que satisfazem a equação  $y_d = \mathbf{X}_d^T \mathbf{w}$ . Portanto, existem infinitas soluções de otimização, para o valor da função objetivo  $J = 0$  (Minimização do erro).

O método de otimização é definido de acordo com o problema. Em princípio para o problema de aprendizado, a equação dos pesos ótimos é dada por:

$$\mathbf{w}^* = (\mathbf{X}_d^T \mathbf{X}_d)^{-1} \mathbf{y}_d \quad (3.7)$$

sendo  $\mathbf{w}^*$  o vetor de pesos ótimos,  $\mathbf{X}_d$  o vetor de entrada e  $y_d$  o vetor de valores de saída desejados.

Uma forma eficiente de encontrar o vetor de pesos ótimos, proposta em [1] é o algoritmo de Kaczmarz's, que assume a seguinte forma:

$$\mathbf{w}^{k+1} = \mathbf{w}^{(k)} + \mu \frac{e_k \mathbf{x}_{d,R_{(k)+1}}}{\|\mathbf{x}_{d,R_{(k)+1}}\|^2} \quad (3.8)$$

sendo  $w^{(0)} = 0$ .

O erro para a  $k$ -ésima iteração:

$$e_k = y_{d,R_{(k)+1}} - \mathbf{x}_{d,R_{(k)+1}}^T \mathbf{w}^{(k)} \quad (3.9)$$

sendo que  $R_{(k)}$  o único inteiro em  $0, \dots, p-1$  que satisfaz  $k = lp + R_k$  para algum inteiro  $l$ , sendo  $R(k)$  é o resto da divisão  $\frac{k}{p}$ .

O algoritmo proposto na Eq.3.9 foi proposto para o treinamento de neurônios lineares por Widrow e Hoff, em 1986 época em que os estudos das redes neuronais foram retomados devido ao surgimento dos algoritmos de aprendizado.

As primeiras redes neuronais, com arquitetura tipo “Perceptron”, provindas de neurônios MCP-McCulloch e Pitts (primeira simplificação matemática de um neurônio biológico), que não apresentavam algoritmo de treinamento, levavam em sua estrutura um combinador linear e uma função de ativação caracterizada como portas de Limiar (“Threshold Gates”), onde era realizada uma comparação com um limiar pré-inicializado.

Essas redes ainda são utilizadas em problemas linearmente separáveis. Uma das

condições para os uso deste tipo de rede seria  $p < n$ , pois não ocorreria aprendizado caso o número de entradas fosse maior do que o número de neurônios  $p$ , logo toda lógica de aprendizado estaria relacionada à função de ativação. A estrutura de uma rede perceptron está descrita na Figura 15:

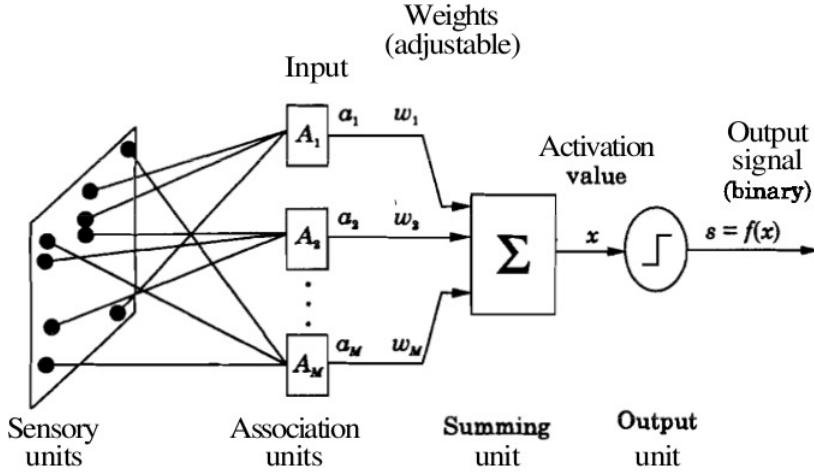


Figura 15 Perceptron aplicado a mapeamento de sensores. Fonte: <http://cse22-iiith.vlabs.ac.in/exp3/index.html>

Na Figura 15 podemos observar que inicialmente a estrutura rede perceptron não possui inicialmente algoritmo de aprendizado, sendo o número de  $n = p$ . As funções de limiar ou portas de limiar de ativação mostrada na figura anterior como etapa ligada a saída possuem a seguinte forma:

$$\begin{cases} 1 & se (\sum_{i=1}^n (w_i x_i) \geq \theta) \\ 0 & se (\sum_{i=1}^n (w_i x_i) < \theta) \end{cases}$$

Para entradas formadas por níveis lógicos, as funções de limiar correspondem à um pequeno subconjunto do total de funções booleanas possíveis. De acordo com o teorema de convergência do perceptron, se o problema for linear, o perceptron irá encontrar a solução em um intervalo finito de iterações.

Redes *Perceptron* de uma única camada podem ser implementadas para aprendizado das funções lógicas *AND* e *OR*, exceto para porta *XOR* (OU-Exclusivo), sendo necessária uma camada de pesos adicionais, juntamente com um limiar de ativação adicional.

As soluções encontradas pelo perceptron para as portas lógicas citadas anteriormente estão na Figura 16:

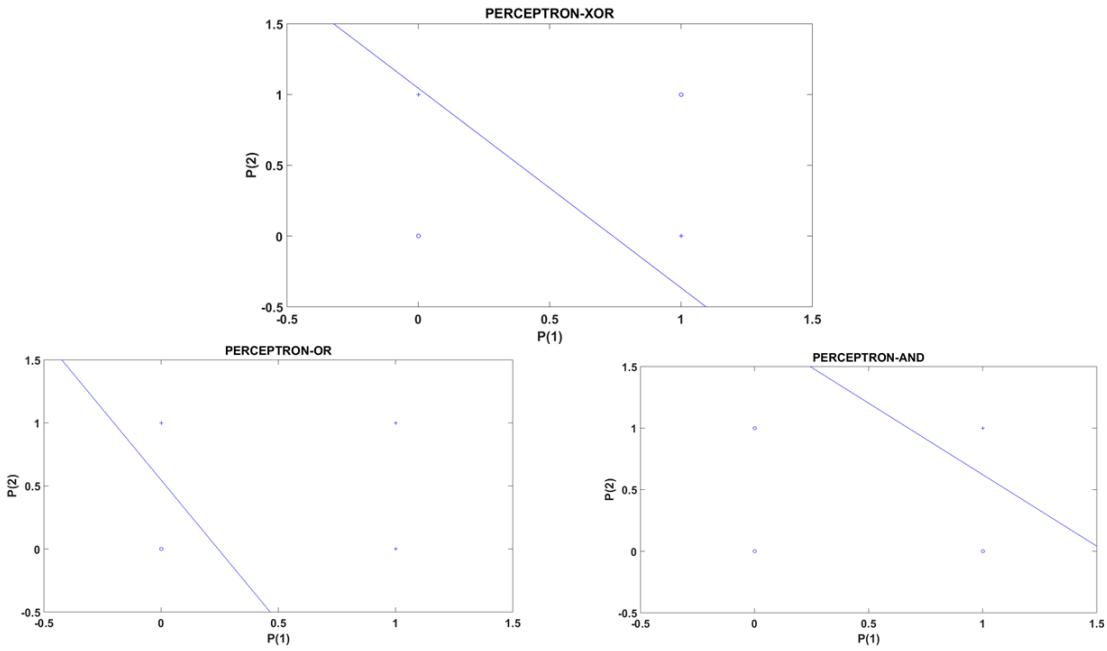


Figura 16 Soluções encontradas por uma RNA tipo Perceptron para AND, OR e XOR.  
Fonte: Autor.

É possível observar que, na implementação referente à porta *XOR* a Rede Perceptron não conseguiu convergir ou separar todos os dados requeridos. Desta forma foi necessário o desenvolvimento do algoritmo de *backpropagation* (retro-propagação) do erro, desenvolvido inicialmente como Regra Delta para o aprendizado dos pesos, caracterizado pela seguinte equação:

$$\mathbf{w}_{(n+1)} = \mathbf{w}_{(n)} + \eta \mathbf{x}_n e \quad (3.10)$$

sendo  $\eta$  o passo de aprendizado e  $e$  o erro entre a saída da rede e a saída do combinador linear e a saída desejada.

A partir da Regra Delta descrita na Eq. 3.10, é desenvolvido o algoritmo de aprendizado da rede. Para a implementação do algoritmo de treinamento do perceptron, são necessários alguns passos:

1. Inicializar  $\eta$  (valor inicial para o passo de aprendizado);
2. Inicializar vetor de pesos (valores aleatórios);
3. Aplicar a regra de atualização aos valores iniciais do vetor  $\mathbf{w}$ ;
4. Aplicar a regra de atualização dos pesos (Eq. 3.10) para todos os  $p$  pares  $(x^i, y_d^i)$  do conjunto de treinamento)  $\Gamma = (x^i, y_d^i)_{i=1}^p$ ;

5. Repetir o passo anterior até que  $e = 0$  para todos os  $p$  elementos de  $\Gamma$ .

É de suma importância se observar que devido a este problema, a pesquisa em RNA's ficou estagnada por quase 10 anos. A nova estrutura proposta pelo algoritmo de aprendizado é mostrado na Figura 17:

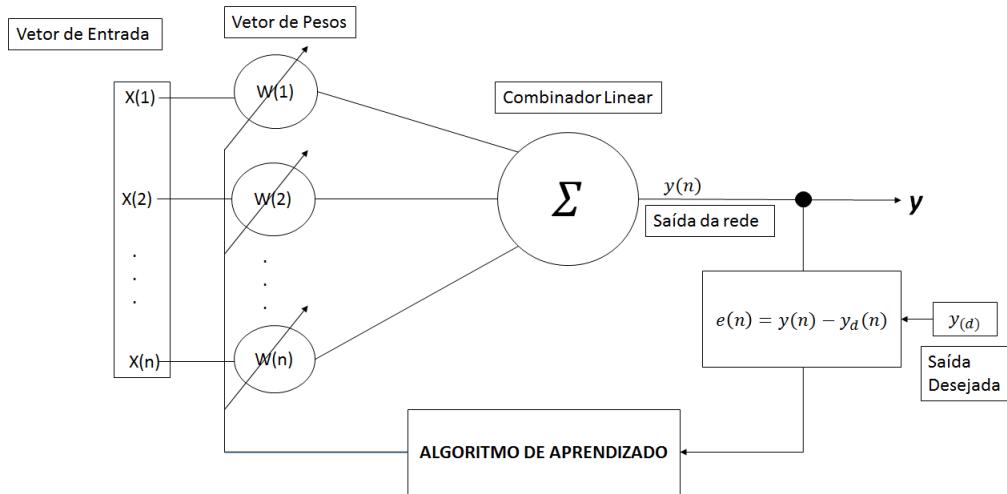


Figura 17 Estrutura de aprendizado de uma RNA caracterizada como Adaline. Fonte: Autor.

A figura acima demonstra um elemento linear adaptativo, sendo esta a rede neural que primeiro foi aplicado efetivamente na indústria para resolução de problemas mais complexos, onde a função de ativação, juntamente com o algoritmo de retro-propagação forma um sistema para caracterização de funções não lineares.

A estrutura básica e completa de um neurônio envolve ainda o *Bias* (Viés- Valor inicial inserido para garantir a ocorrência do aprendizado ou otimização do mesmo) e a inserção da combinação linear, adicionada do bias, à função de ativação ( $y = g(.)$  descrita na Eq.3.3). a estrutura final do neurônio é dado por:

$$y_{saída} = g(b + \sum_{i=1}^n w_i x_i) \quad (3.11)$$

sendo  $b$  o bias ou valor inicial e  $g(.)$  ou função composta  $g \circ a$  como resultado da sigmoide aplicada ao combinador linear adicionado do *bias*.

### 3.5 Backpropagation

O algoritmo de retropropagação é a maneira mais geral de aprendizado de uma rede. O objetivo do algoritmo é realizar o ajuste dos pesos com base no erro  $e$  o valor

desejado  $y_d$ .

O aprendizado de uma rede neuronal se resume à encontrar o mínimo de  $F : \mathbf{R}^n \rightarrow \mathbf{R}$  implícita no conjunto de treinamento  $\Gamma = (x^i, y_d^i)_{i=1}^p$ , através da minimização da função de erro descrita na Eq.3.6. O método em geral utilizado para a minimização é o “*Steepest Gradient Descent*” cujo gradiente da função de erro deve ser computado recursivamente, em cada iteração, como por entre as camadas da rede (redes de múltiplas camadas). O vetor gradiente da função de erro em relação aos pesos fica:

$$\nabla \mathbf{J} = \left( \frac{\partial J}{\partial w_1}, \frac{\partial J}{\partial w_2}, \dots, \frac{\partial J}{\partial w_n} \right) \quad (3.12)$$

sendo  $J$  a função de erro a ser minimizada observada na Eq. 3.6 e  $w_1, \dots, w_n$  o conjunto de pesos da rede neuronal. Para que ocorra o aprendizado, devem ser garantidos a continuidade e derivabilidade da função de erro. Através da função de ativação (função sigmoide observada na Eq. 3.3), pode-se garantir essas condições.

A função sigmoide em geral é escolhida como função de ativação das redes neurais devido a simplicidade da suas derivadas e comportamento ajustável, sendo uma aproximação da função de degrau, contínua e diferenciável. A função de atualização dos pesos com base no gradiente da função pode ser descrita como:

$$\Delta_{w_i} = -\gamma \frac{\partial J}{\partial w_i} \rightarrow i = 1, \dots, n_{pesos} \quad (3.13)$$

sendo  $\gamma$  o passo ou constante de aprendizado e  $n_{pesos}$  o número de pesos ( $n \geq n_{pesos}$ ).

Uma outra vantagem da função sigmoide é sua simetria, necessária à rede neuronal, buscando evitar problemas como *overfitting* (aprendizado limitado somente ao conjunto de treinamento, não havendo generalização) [12].

### 3.6 Problemas no Processamento de Imagens por Redes Clássicas

O número de informações contidas em uma imagem, em escala de cinza ou colorida geram uma forte tendência, em redes de arquitetura *Fully Connected* para o *overfitting* (sobreajuste devido à grande quantidade de informações e baixo número de dados) [8]. O *overfitting* torna aprendizado restrito ao conjunto de dados de treinamento, não alcançando capacidade de generalização. Características como cor, forma e aspectos, associadas com variância de posição, ângulo de rotação e variação das estatísticas do cenário

de aquisição da imagem, sobrecarregam esta arquitetura. O aumento exponencial do número de conexões pode ser observado na Figura 18:

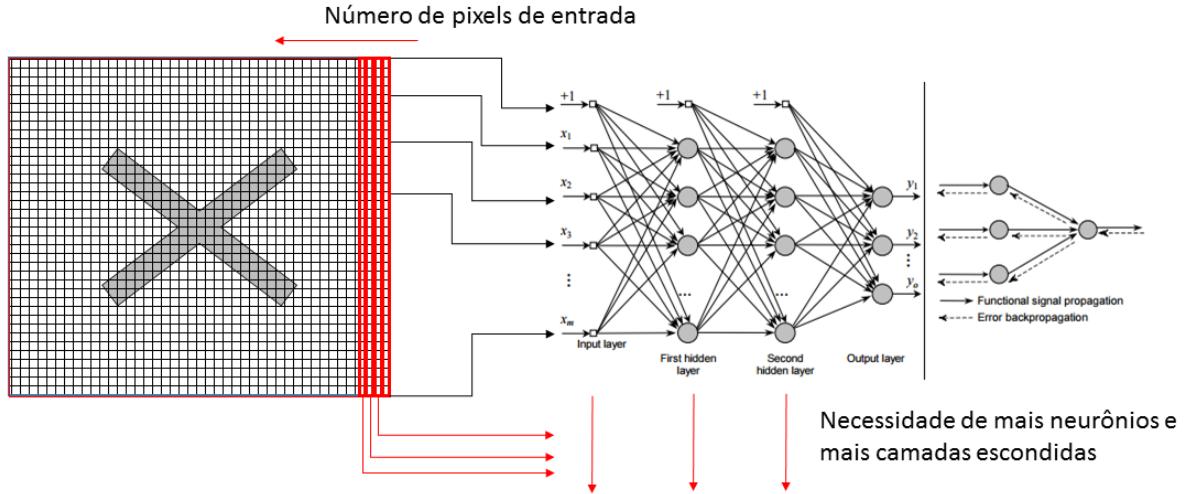


Figura 18 Imagem sendo processada por uma MLP. Fonte: Autor.

A dimensão do vetor de entrada na MLP, tendo como dado uma matriz (imagem em escala de cinza) MLP, caso fosse tratada por uma RNA clássica, considerando imagens  $n_{pixels} \times n_{pixels}$ ,  $\mathbf{x}_i$  teria a dimensionalidade observada na Eq. 3.13:

$$\mathbf{x}_i \in \mathbf{R}^{n^2_{(pixels)}} \quad (3.14)$$

sendo  $\mathbf{x}_i$  o vetor de entrada na MLP de forma que os sistemas

O número de conexões, considerando uma arquitetura tipo *Fully Connected*, seria aproximada pela Eq. 3.14:

$$n_{connections} \cong n_{(pixels)}^{(4 \times n_{hidden\ layers})} \quad (3.15)$$

sendo  $n_{connections}$  o número aproximado de conexões,  $n_{pixels}$  o número de linhas (dimensão da imagem, sendo esta quadrada) e  $n_{hidden\ layers}$  o número de camadas escondidas.

Para uma imagem RGB (*Red, Green, Blue*) a número de conexões se torna:

$$n_{connections} \cong 3 \times n_{(pixels)}^{(4 \times n_{hidden\ layers})} \quad (3.16)$$

Desta forma, a arquitetura proposta irá absorver várias características desnecessárias, o que aumenta o número de iterações, matrizes de entrada (imagens) e intensidade do erro e de aprendizado. As características ou padrões a serem extraídas em

uma imagem ou de um conjunto de imagens podem ser observados na Figura 19:

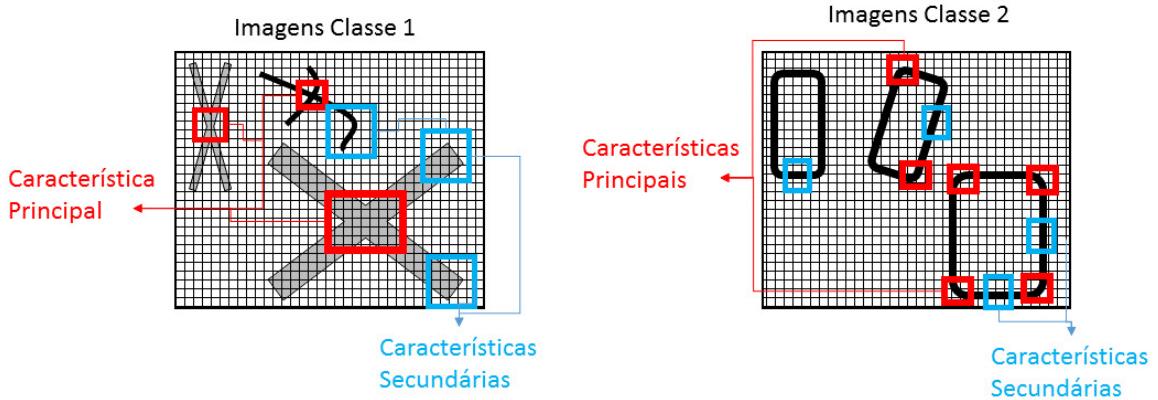


Figura 19 Imagens associadas a dois tipos de classes diferentes. Fonte: Autor.

Pode ser observado na Figura 19 que as imagens, tanto na classe 1 quanto da classe 2 possuem as características principais, as quais estarão presentes em todas as imagens do treinamento e as características secundárias (que possuem uma maior variação no conjunto de imagens, desnecessárias ao processo de classificação). Para uma imagem em escala de cinza o problema de classificação se resume inicialmente à um problema de redução de dimensionalidade do tipo:

$$\mathbf{R}^{n_{\text{pixels}}^2} \rightarrow \mathbf{R}^{v_{\text{features}}} \rightarrow \mathbf{R} \quad (3.17)$$

sendo  $v_{\text{features}}$  a dimensão do vetor de características  $v_{\text{features}}$  na primeira etapa de redução de dimensionalidade (realizada pelas camadas convolucionais vistas mais adiante no capítulo) e  $R$  o número real associado à classe da imagem (tarefa realizada pela arquitetura tipo *Fully Connected*).

Neste contexto, devido aos problemas de dimensionalidade e excesso de informação que um conjunto de imagens pode gerar em redes de arquiteturas clássicas, são necessárias arquiteturas mais avançadas, cuja extração de características seja um processo hierárquico robusto à variações nas estatísticas do cenário, translação e rotação dos padrões dentro da imagem.

### 3.7 Redes Neuronais Convolucionais

A extração de características de alta ordem para classificação de imagens, por redes neuronais, sem a necessidade de processamento prévio, tem sido desenvolvida com a

finalidade de aplicar uma ferramenta menos limitada à problemas considerados complexos (problemas onde ocorrem translações ou rotações do objeto a ser classificado na imagem, por exemplo). Redes Neuronais Convolucionais (classe específica de Redes Neuronais Profundas, voltadas ao processamento de uma quantidade extensiva de imagens) são utilizadas com sucesso para classificação supervisionada e não-supervisionada de imagens como no trabalho de (Xie et al, 2016). Posteriormente à análise dos resultados de (Hubel, 1981) sobre a forma como o córtex visual de animais respondia a estímulos com orientação espacial variável, foi observado que o cérebro dos animais realizava a extração de características das imagens de forma hierárquica, sendo os níveis primários responsáveis pelo reconhecimento e extração de características de cor e forma, sendo posteriormente processado um número reduzido de informações acerca da imagem por áreas mais profundas do cérebro [14].

As descobertas de Hubel geraram novas perspectivas aos algoritmos de aprendizado de máquina, cuja extração de características de um conjunto de imagens era uma atividade inviável para as redes neuronais tipo *Fully Connected*. O *Deep Learning* ou Aprendizado Profundo então surge como paradigma que dá origem as *DNN- Deep Neural Networks* ou redes cujas extração de características é realizada de forma hierárquica (forma de aprendizado descoberta nos trabalhos de Hubel).

Neste contexto, as *CNN- Convolutional Neural Networks* uma classe específica das *DNN* desenvolvidas com base no conceito de conectividade esparsa e conectividade local, onde campos receptivos (filtros responsáveis pelo processamento de partes de um dado de entrada) são utilizados para realizar a filtragem de cada parte de uma matriz (imagem, por exemplo), de forma que a rede consiga extrair informações de cor, forma, bordas, arestas (para o caso de imagens) de partes de imagens, tornando o aprendizado mais robusto ao “*overfitting*” (sobreajuste ou memorização do banco de dados ao invés da generalização, devido a alta dimensionalidade dos dados).

O conceito de conectividade local vem de encontro a forma das redes clássicas do tipo *MLP* (“*Multi-Layer-Perceptron*”) ou “*Fully Connected*” que possuem “Conectividade Global” (todos as unidades neuronais associadas a todas as partes do dado de entrada), não conseguindo explorar as características principais ou secundárias (conjunto de elementos de estatísticas semelhantes associáveis à um número em uma matriz ou um conjunto de matrizes), independentemente do local ou da posição angular nas matrizes ou imagens em

que estas características possam estar [8].

O processo de aprendizado profundo tem como base a forma de extração de características (camadas iniciais nas redes profundas), sendo a combinação destas, uma característica de maior nível (camadas finais das redes profundas).

A extração de características profundas de conjunto de imagens por uma Rede Neuronal Convolucional, pode ser visualizado na Figura 20:

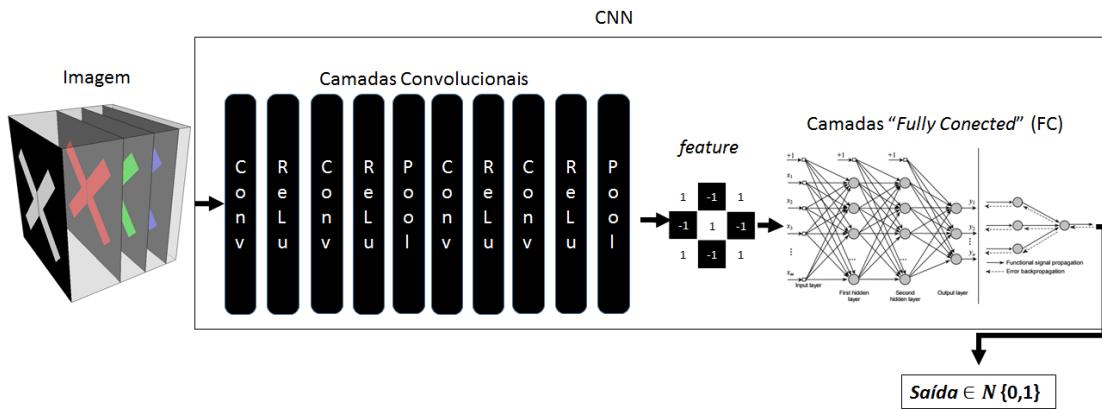


Figura 20 Modo de treinamento de uma Rede Neural Convolucional. Fonte: Autor.

Pode-se observar na Figura 20 as camadas *Conv*(Convolução com filtros de *Kernel* ou núcleo adaptativo), as camadas *ReLU*(Rectified Logic Unit) irão realizar a pré-ativação de acordo com os valores máximos de cada mapa de ativação (matrizes resultantes da camada *Conv*) e as camadas *Pool* que realizarão o agrupamento e subamostragem com base nos critérios descritos na Eq. 3.21. Após passar pelas estruturas *Conv*, *ReLU* e *Pool*(Figura 20), são extraídos da imagem vetores de características (caracterizados como *features* na Figura 20). Os vetores de características são inseridos em uma estrutura de rede do tipo *Fully Connected* para a classificação com base nas *features* extraídas.

O processo de extração de padrões inicia com a convolução da imagem com um filtro ou kernel que será posteriormente reajustado para a extração das melhores características, tendo como padrão de correção o rótulo ou valor esperado para a imagem de entrada. O processo de convolução ocorre segundo a Figura 21:

O modelo de convolução da imagem com o núcleo do filtro na camada convolucional (Figura 21) é dado por:

$$ma[i][j] = \frac{\sum_m \sum_n k_{i,j} m_{ch(i+m),(j+n)}}{\sum_n k_{(ch)(n)}} + rcc_{ch(i,j)}, \quad (3.18)$$

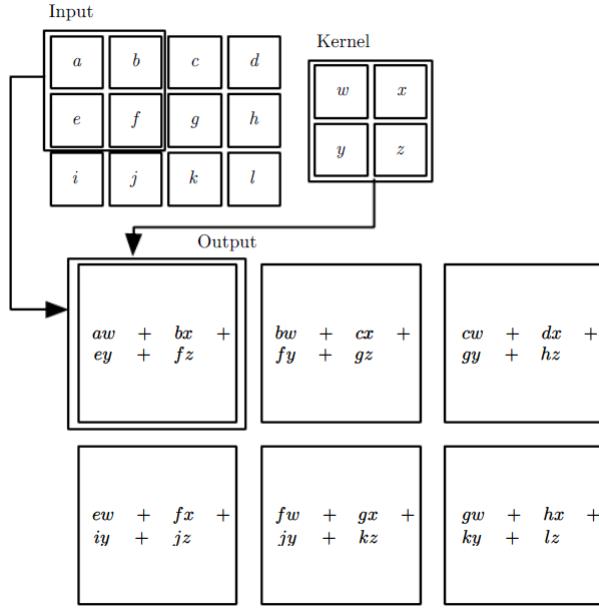


Figura 21 Processo de convolução da imagem com um kernel tendo como saída um mapa de ativação. Fonte: Autor.

sendo  $ma_{[i],[j]}$  os elementos que formam o *Mapa de Ativação MA* a matriz de saída da camada convolucional,  $m$  e  $n$  os índices dos elementos  $k_{(m,n)}$  do núcleo do filtro e  $im_{ch(i+m)(j+n)}$  os elementos do respectivo canal  $ch$  da imagem de entrada (Topo da Figura 21). Em seguida, sofrem uma combinação linear e são adicionadas ponderações  $rcc_{ch(i,j)}$  dos outros planos de cor à **MA**.

O processo de geração do mapa de ativação pode ser observado na Figura 22:

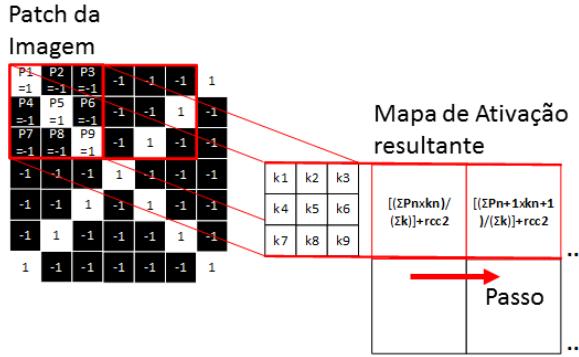


Figura 22 Processo de formação do mapa de ativação. Fonte: Autor.

O modelo de geração do mapa de ativação está descrito na Eq.3.19 como:

$$\mathbf{MA} = Red(\mathbf{CMC}) + Green(\mathbf{CMC}) + Blue(\mathbf{CMC}), \quad (3.19)$$

sendo a **CMC** a matriz formada pelos elementos de contribuição  $rcc_{ch}$  do respectivo plano de cor ou canal  $ch$ .

Após a geração dos mapas de ativação (Eq. 3.19), estes sofrerão uma retificação lógica (camada *ReLU-Rectified Logic Unit*), cuja função aplicada ao elemento é a função *sigmoide-logística* ou função de retificação, descrita para cada elemento do mapa de ativação **MA**, ou seja:

$$ma_{[i][j](ReLU)} = \ln(1 + e^{ma_{[i][j]}}), \quad (3.20)$$

Por fim, a matriz ou mapa de ativação retificado **MA<sub>ReLU</sub>** (Eq. 3.20), passa por um processo de agrupamento ou (*max-pooling*), dado por:

$$ma_{[i],[j](Pool)} = \max(ma_{[i],[j](ReLU)} : ma_{[i+m],[j+n](ReLU)}), \quad (3.21)$$

sendo  $ma_{[i],[j](Pool)}$  o respectivo elemento do mapa de agrupamento **MA<sub>pool</sub>**, gerado pelo valor máximo encontrado em cada sub-matriz formada pelos elementos  $ma_{[i],[j](ReLU)} : ma_{[i+m],[j+n](ReLU)}$ . As sub-matrizes selecionadas para extração dos seus valores máximos não são sobrepostas. O elemento  $ma_{[i],[j](ReLU)}$  é o primeiro da sub-matriz (contida no mapa de ativação **MA<sub>ReLU</sub>**) e  $ma_{[i+m],[j+n](ReLU)}$ , o último desta sub-matriz.

O agrupamento pode ser observado em maiores detalhes na Figura 23:

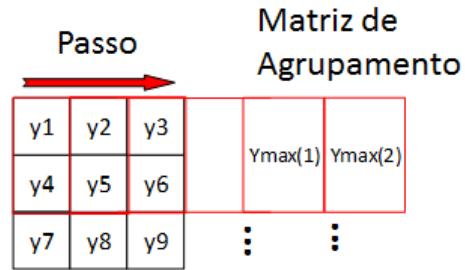


Figura 23 Processo de formação do mapa de ativação. Fonte: Autor.

Pode ser observado na camada de agrupamento que características principais são potencializadas, enquanto valores numéricos do mapa de ativação que não sejam de máximo valor são suprimidas (características secundárias), formando um novo mapa de ativação de menor dimensão, somente com as características mais expressivas.

O método de aprendizado da *CNN* é baseado no (*Steepest Gradient Descent-SGD*) ou *método da descida mais íngreme*.

É importante observar que, devido as dimensões da imagem e do “*kernel*” (núcleo do filtro ou das camadas convolucionais), as operações nas camadas de convolução (*Conv*) necessitam do deslocamento desse sobre as matrizes de entrada, não possuindo estas correta partição entre os deslocamento do “*kernel*” e as dimensões da matriz de entrada, logo, a adição simétrica de valores nulos às essas matrizes ocorre para fins de ajuste (ou devido à exigência de algumas arquiteturas de *CNN* quanto a dimensionalidade das matrizes entre camadas).

Zero Padding =1									
Zero Padding =0									
0	o	o	o	o	o	o	o	o	0
0	P1	P2	P3	P1	P2	P3	-1	-1	1
0	=1	=1	=1	=1	=1	=1	-1	-1	0
0	P4	P5	P6	-1	-1	1	-1	-1	0
0	=-1	=1	=-1	-1	-1	1	-1	-1	0
0	P7	P8	P9	-1	1	-1	-1	-1	0
0	=-1	=-1	=1	-1	1	-1	-1	-1	0
0	-1	-1	1	-1	1	-1	-1	-1	0
0	-1	-1	1	-1	1	-1	-1	-1	0
0	-1	1	-1	-1	-1	1	-1	-1	0
0	1	-1	-1	-1	-1	-1	-1	1	0
0	o	o	o	o	o	o	o	o	0

Figura 24 Adição simétrica de zeros à matriz. Fonte: Autor.

Desta forma os hiperparâmetros a serem configurados permitem observar, com base na dimensão dos dados de entrada, a dimensão da matriz de saída. Esses hiperparâmetros

estão descritos na Tabela 1:

Tabela 1 Hiperparâmetros principais para configuração da *CNN*.

[Hiperparâmetro]	[Descrição]
R	Dimensão geral do Campo Receptivo
S	Passo
P	Valor do <i>Zero Padding</i>
W	Largura do Campo Receptivo
H	Altura do Campo Receptivo
D	Profundidade do Campo Receptivo

A partir dos hiperparâmetros descritos na Tabela 1, a dimensão das matriz de saída nas camadas *Conv*, *ReLU* e *Pool* em relação a uma matriz de entrada, pode ser aproximada pela seguinte equação (apresentada em [15]):

$$W_{out}(i) \cong \frac{W_{in}(i) - R + 2P}{S} \quad (3.22)$$

sendo  $W_{out}(i)$  a largura ou dimensão da matriz de saída em função da dimensão  $W_{in}(i)$  da matriz de entrada (considerado também como dimensão dos campos receptivos).

A Equação 3.22 tem como função gerar a dimensão ou largura aproximada da matriz de saída entre uma camada e outra. Como informado pelos autores do trabalho de onde foi extraída a Eq. 3.22, esta resulta em um valor aproximado da dimensão de saída podendo apresentar variações em cada arquitetura específica de rede.

#### 4 MODELO DO SISTEMA DE SONDAGEM

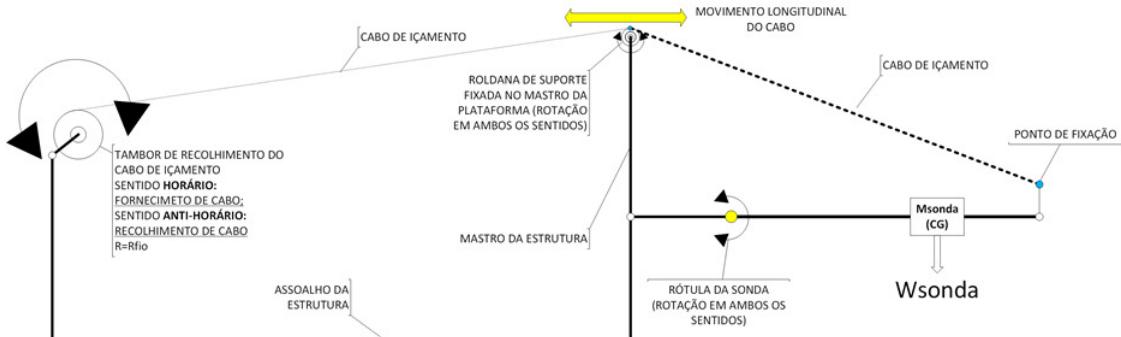


Figura 25 Modelo Físico do Sistema de Sondagem. Fonte: Autor.

Podem ser observado na Figura 25 o modelo geral do sistema de sondagem, composto por uma Plataforma Mecânica onde está fixado o sistema de sondagem em mecanismo basculante (rótula fixa na parte frontal da plataforma) com ponto de içamento na extremidade da sonda. O cabo de içamento é armazenado em um sistema de recolhimento caracterizado por um cilindro de armazenamento do cabo acoplado a um sistema de redução mecânica, acionado por um motor elétrico de corrente contínua (*DC-Direct Current*).

A partir do acionamento do motor elétrico em ambos os sentidos de rotação ocorre consequentemente o recolhimento ou fornecimento do cabo de içamento, resultando no movimento de subida e descida da extremidade frontal da sonda. A extremidade da sonda acoplada à rótula (Figura 25) executa um movimento de rotação, cujo sentido (horário ou anti-horário), depende do sentido de rotação do motor elétrico. A sonda tem o ponto de aplicação de seu peso no centro de massa, caracterizada como  $M_{sonda}(CG)$  na Figura 25.

Como base neste ponto de aplicação do peso, são calculados todos os esforços no mecanismo de içamento (esforço no cabo, esforços no ponto de içamento, esforços na estrutura de fixação e o torque necessário ao levantamento da extremidade da sonda). A movimentação da sonda assim como seus ângulos de inclinação estão representados pela Figura 26.

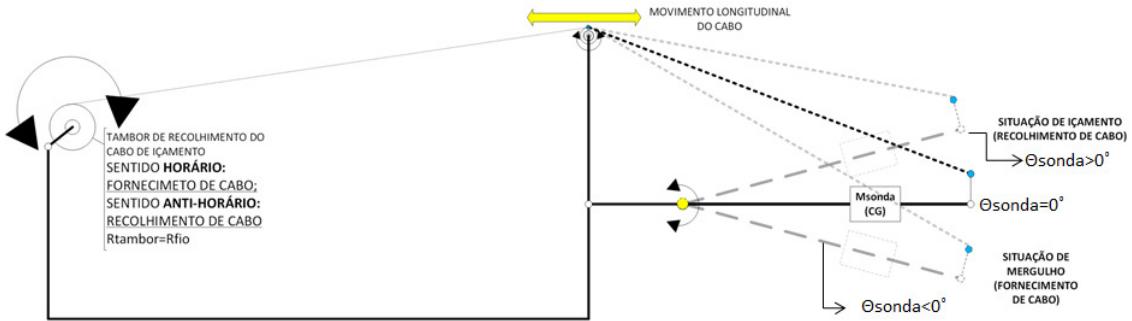


Figura 26 Modelo Físico do Sistema de Sondagem. Fonte: Autor.

Pode ser observado na Figura 26 os ângulos de inclinação do sistema de sondagem  $\theta_{sonda} > 0$  (situação de içamento da sonda- recolhimento do cabo),  $\theta_{sonda} \leq 0$  (situação de mergulho da sonda- fornecimento de cabo) e  $\theta_{sonda} = 0$  (sonda na posição “neutra”).

O modelo geométrico de movimentação tem como base o fornecimento e retirada de cabo de içamento, resultando em um triângulo imaginário com um dos lados formado pela distância da roldana de suporte e o ponto de içamento na extremidade frontal da sonda. Modelo geométrico pode ser melhor expresso na Figura 27:

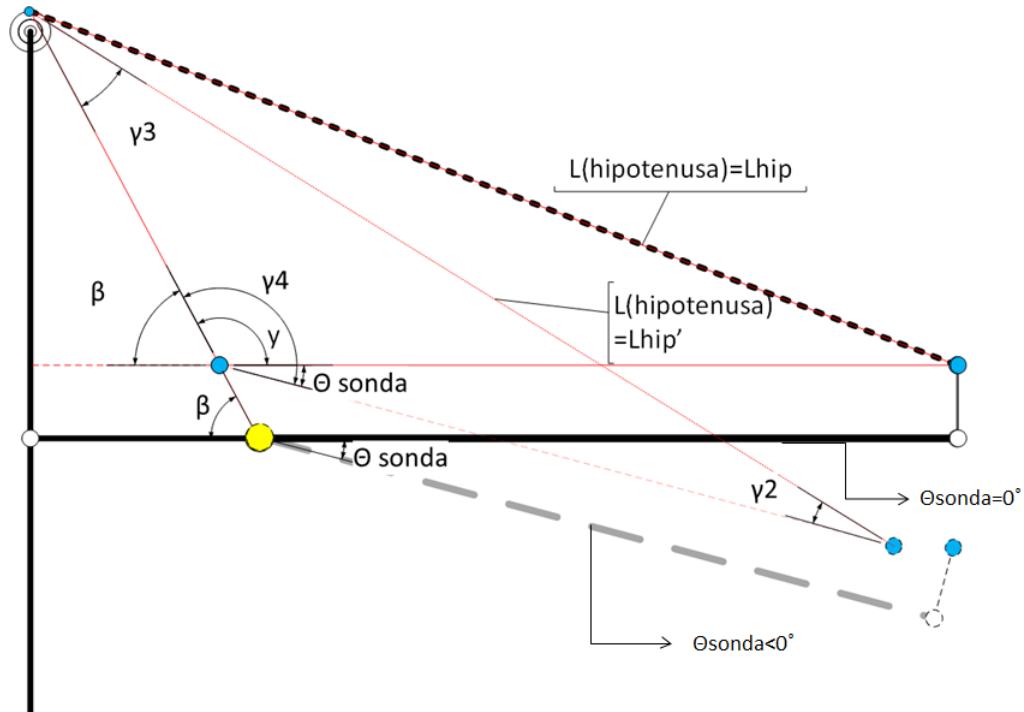


Figura 27 Modelo Geométrico do Sistema de Sondagem. Fonte: Autor.

Pode-se observar na Figura 27 alguns dos principais ângulos formados no sistema de sondagem  $\theta_{sonda}$ ,  $\beta$ ,  $\gamma$ ,  $\gamma_2$ ,  $\gamma_3$  e  $\gamma_4$ , que serão necessários ao processo de modelagem do movimento angular resultante da sonda. Na Figura 27 também é possível observar que

são formados triângulos, alguns destes deslocados do ponto de rotação (Rótula da Sonda na Figura 25) do ponto de içamento.

As novas distâncias formadas pelo deslocamento do ponto de içamento são mostrados na Figura 28:

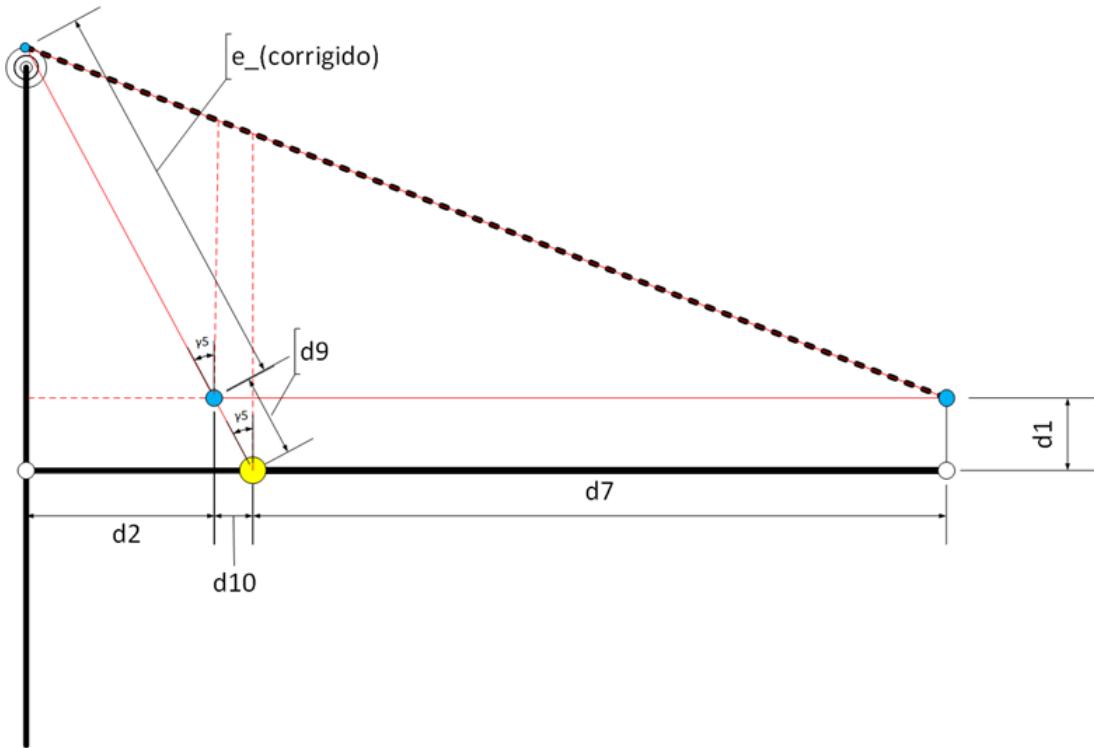


Figura 28 Modelo Geométrico do Sistema de Sondagem (Correção das distâncias). Fonte: Autor.

Podem ser observadas na Figura 28 as distâncias  $d1$ ,  $d9$ ,  $d10$ ,  $e_{corrigido}$  formadas pelo deslocamento do ponto de içamento e o ângulo  $\gamma_5$ . Os principais triângulos do modelo de posicionamento podem ser observados na Figura 29.

Os triângulos observados  $T1$ ,  $T2$ ,  $T3$ ,  $T4$  e  $T5$  são os principais no processo de modelagem do posicionamento. A equação geral que descreve o movimento da sonda em termos angulares:

$$\theta_{sonda} = \pi - \beta - \gamma_4 + \pi, \quad (4.1)$$

sendo  $\beta$  e  $\gamma_4$  ângulos observados na Figura 27 e ( $\theta_{sonda}$ ) o ângulo principal ou ângulo de inclinação do sistema de sondagem (Figura 27).

Pode ser observado na Figura 29 que o triângulo  $T3$  possue como maior lado o comprimento do cabo entre a roldana superior (Figura 25) e o ponto de fixação do

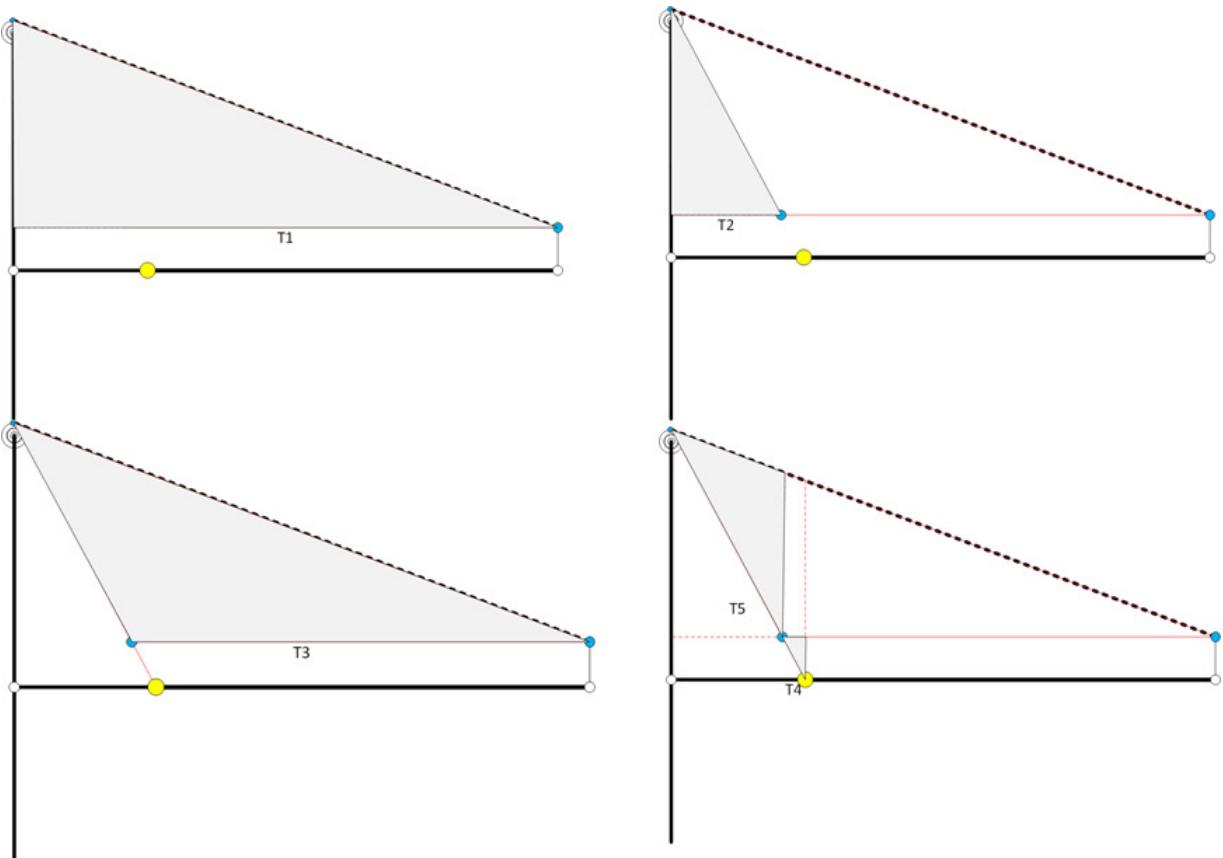


Figura 29 Triângulos do modelo geométrico no sistema de sondagem. Fonte: Autor.

mesmo na sonda (ponto de içamento), caracterizado como  $L_{(hip)}$ , isso para o instante  $t(s)$ . A partir desses parâmetros, pode ser formado o modelo dinâmico do sistema de posicionamento. O valor do último termo  $\pi$  radianos adicionado a Eq.4.1 é necessário devido ao cosseno do ângulo  $\gamma_4$  iniciar no quadrante cujo seu valor é negativo, o que representa uma troca de comportamento no modelo de posicionamento, invertendo sua variação. Desta forma, o quadrante que possui a mesma variação está à  $\pi$  radianos acima. Pode-se observar também observar que se, a distância  $d1 = 0$ , o triângulo  $T3$  será formado por algumas partes da sonda, resultando em limitações ao sistema de posicionamento, em situações onde  $\theta_{sonda} < 0$ , podendo ocorrer a trava do mecanismo ou excesso de carga no cabo devido à proximidade com o "PM"(Ponto Morto) do mecanismo de posicionamento. Devido à isso  $d1 > 0$  para todas as situações do presente estudo.

Os triângulos apresentados na Figura 29 tem como base as dimensões do sistema de posicionamento. As principais dimensões que dão origem ao modelo de posicionamento, juntamente com os esforços e torques nos principais pontos estão descritos na Figura 30:

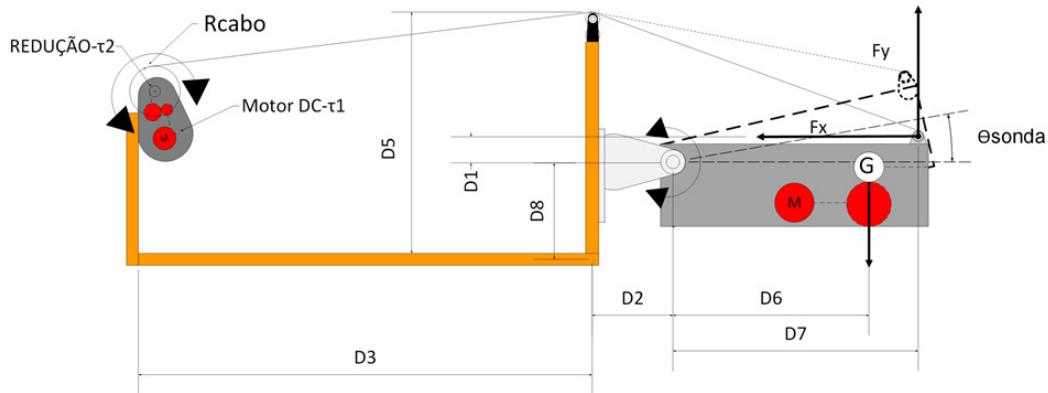


Figura 30 Triângulos do modelo geométrico no sistema de sondagem. Fonte: Autor.

Pode ser observado na Figura 30 as distâncias  $d_1$ ,  $d_2$ ,  $d_3$ ,  $d_5$ ,  $d_6$ ,  $d_7$ ,  $d_8$ , os esforços  $F_x$  e  $F_y$  e os torques do motor DC  $\tau_1$  e na saída do redutor  $\tau_2$ . A partir destes parâmetros, juntamente com a massa da sonda  $M_{sonda}$ , os modelos de posicionamento podem ser gerados, como pode ser observado na Eq. 4.2:

$$\beta = \tan^{-1}\left(\frac{d_5 - d_8 - d_1}{d_2}\right), \quad (4.2)$$

sendo  $d_5$ ,  $d_8$ ,  $d_1$  e  $d_2$  as distâncias do modelo encontradas na Figura 30.

O ângulo  $\gamma_4$  principal referente ao triângulo T3 (considera o ângulo  $\theta_{sonda}$ , cuja lei dos cossenos será aplicada com o objetivo de encontrar o comprimento do cabo de içamento no instante t. Quando  $\theta_{sonda} = 0$ , o ângulo  $\gamma$  pode ser medido e colocado como constante. Desta forma para encontrar a distância e temos a Eq. 4.3:

$$e^2 = d^2 + (d_5 - d_8)^2, \quad (4.3)$$

sendo  $e$ ,  $d_5$ ,  $d_8$

$$e_{corrigido} = e - d_9 \quad (4.4)$$

O lado adjacente vem a ser o comprimento da rótula  $d_7$  corrigida da seguinte forma:

$$d_7_{corrigido} = d_7 + d_{10}, \quad (4.5)$$

As distâncias  $d_9$  e  $d_{10}$  são obtidas a partir das seguintes deduções através das projeções do triângulo T4 pelo ângulo  $\gamma_5$  (Figura 29 e Figura 28) respectivamente:

$$d9 = \frac{d1}{\cos \gamma_5}, \quad (4.6)$$

sendo  $d9$ ,  $d1$  e  $\gamma_5$  as distâncias e o ângulo gerado pelo desvio do ponto de rótula. A distância  $d10$  é encontrada a partir da relação:

$$d10 = d1 \tan \gamma_5 \quad (4.7)$$

O ângulo  $\gamma_5$  é encontrado da seguinte forma:

$$\gamma_5 = \frac{\pi}{2} - \tan^{-1}\left(\frac{d5 - d8}{d2}\right) \quad (4.8)$$

Pela lei dos cossenos, tem-se:

$$(L_{hip'})^2 = e_{corrigido}^2 + d7_{corrigido}^2 - 2(e_{corrigido})(d7_{corrigido})(\cos \gamma_4), \quad (4.9)$$

sendo  $L_{hip'}$  o comprimento do cabo de içamento no instante  $t(s)$ .

Colocando o ângulo em função do comprimento do cabo, tem-se a Eq. 4.10:

$$-\cos^{-1}\left[\frac{e_{corrigido}^2 + d7_{corrigido}^2 - L_{hip'}^2}{(e_{corrigido})(d7_{corrigido})}\right] = \beta - 2\pi + \theta_{sonda}, \quad (4.10)$$

sendo o ângulo  $\theta_{sonda}$  o principal ângulo no processo de modelagem, tem-se:

$$\theta_{sonda} = 2\pi - \beta - \cos^{-1}\left[\frac{e_{corrigido}^2 + d7_{corrigido}^2 - L_{hip'}^2}{(e_{corrigido})(d7_{corrigido})}\right] \quad (4.11)$$

Com base no valor da variável  $L_{hip}$  pode-se desenvolver a estratégia de posicionamento (polaridade do sinal para o motor elétrico DC do mecanismo de posicionamento e por qual período o motor deverá permanecer acionado). O valor  $L_{hip}$  corresponde ao comprimento do cabo no instante inicial  $t(i)$ , considerado como comprimento de referência ou comprimento referente ao ângulo ( $\theta_{sonda(i)}$ ) e o comprimento do cabo  $L_{hip'}$  correspondente ao comprimento para o ângulo  $\theta_{sonda(j)}$  desejado para o qual a sonda deve ser reposicionada.

A estratégia de acionamento do sistema de posicionamento tem como início a entrada do ângulo para o qual a sonda deverá ser reposicionada. Com base na diferença entre os comprimentos  $L_{hip}$  e  $L_{hip'}$  tem-se o comprimento resultante  $L_{(e)}$ , que dá origem

as equações referentes ao tempo de acionamento do motor elétrico (Eq.4.12 e 4.19):

$$L_{hip'} - L_{hip} = L_{(e)} \quad (4.12)$$

O valor  $L_e$  terá seu valor absoluto  $\|L_{(e)}\|$ , juntamente com o seu valor real  $L_{(e)}$  utilizados no algoritmo de posicionamento. A estratégia para o acionamento do motor DC pode ser descrita da seguinte forma:

$$f_{acionamento} \begin{cases} \text{Se } \|L_e\| < L_{(hip-min)}, \text{ Motor DC desacionado, senão :} \\ \text{Se } L_{(e)} < 0, \text{ Motor DC acionado (sentido horário), senão :} \\ \text{Se } L_{(e)} > 0, \text{ Motor DC acionado (sentido anti-horário), senão :} \\ \text{Motor DC permanece desacionado} \end{cases}$$

O valor  $L_{hip-min} \geq 0$  é um limitante inferior para o acionamento do motor (abaixo deste valor o motor elétrico permanece inativo), devido a questões construtivas do mecanismo de posicionamento, sendo este incapaz de corrigir a posição angular ( $\theta_{sonda}$ ) de forma satisfatória.

De acordo com a função de rotação  $f_{acionamento}$  e com as variáveis eletromecânicas implícitas no sistemas de posicionamento, o tempo de acionamento para o motor elétrico  $\Delta t_{(acionamento)}$  fica:

$$\Delta t_{(acionamento)} = f(\|L_{(e)}\|, \omega_{motor}), \quad (4.13)$$

sendo  $\omega_{motor}$  a velocidade angular do motor elétrico e  $\|L_{(e)}\|$  o valor absoluto do comprimento do cabo de içamento (a ser fornecido) necessário ao alcance do ângulo  $\theta_{sonda}$  desejado  $\theta_{sonda(j)}$ .

O valor da velocidade angular do motor elétrico  $\omega_{motor}$  será mantida constante durante o processo de posicionamento. A partir dessa condição, pode-se estimar a velocidade angular do mecanismo de recolhimento (Figura 31) ou  $\omega_{mecanismo}$ :

$$\omega_{mecanismo} = f(\%Duty\ Cycle) \quad (4.14)$$

sendo *Duty Cycle* a porcentagem do sinal (onda quadrada) que é responsável pelo aciona-

mento do motor elétrico em nível máximo de tensão. A maior porcentagem do *Duty Cycle* resulta em um aumento linear da velocidade angular  $\omega_{motor}$  e por consequência  $\omega_{mecanismo}$ .

O modelo geral do mecanismo de recolhimento do cabo tem como objetivo fornecer uma equação do tempo de acionamento do motor elétrico em função da distância angular  $\Delta\theta = \theta_{(j)} - \theta_{(i)}$  a ser percorrida pela sonda. O modelo geral do mecanismo de posicionamento poder ser observado na Figura 31:

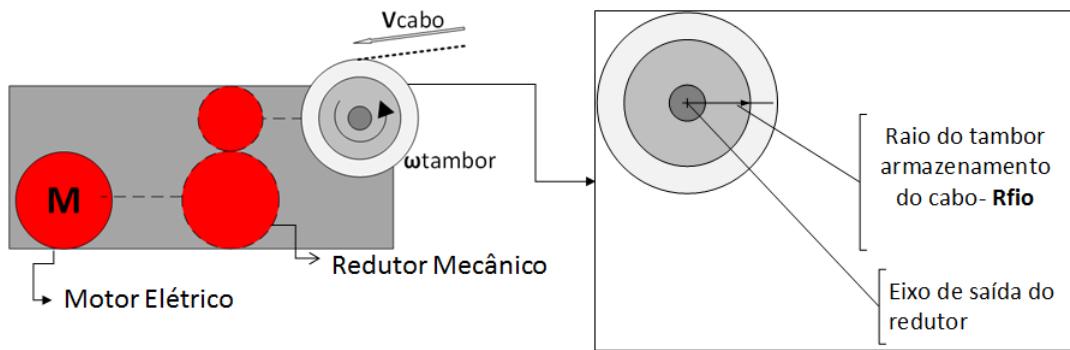


Figura 31 Modelo do Mecanismo de recolhimento e fornecimento do cabo. Fonte: Autor.

Pode ser observado no diagrama na Figura 31 o motor elétrico acoplado à um sistema de redução mecânica onde a velocidade angular de entrada ( $\omega_{motor}$ ) será reduzida por um fator caracterizado como “Relação de Transmissão” ( $rtr$ ). Como resultado, obtém-se no eixo de saída do redutor mecânico, uma velocidade angular ( $\omega_{mecanismo}$ ) associado à um torque (em N.m.) ampliado pelo mesmo fator de redução. Observa-se ainda na Figura 31 o tambor de armazenamento do cabo, com um raio médio caracterizado por  $R_{tambor}$ . A velocidade média linear do sistema de recolhimento ou fornecimento do cabo de içamento  $V_{cabo}$  é obtida por:

$$V_{cabo} = \omega_{mecanismo} R_{cabo} \quad (4.15)$$

Com base nas variáveis expostas, a equação que relaciona a velocidade angular do motor ( $\omega_{motor}$ ) com a velocidade angular do mecanismo ( $\omega_{mecanismo}$ ) é dada por:

$$\omega_{mecanismo} = \frac{\omega_{motor}}{rtr} \quad (4.16)$$

sendo  $rtr$  a relação de transmissão do redutor mecânico ( $rtr > 0$ ).

A partir da velocidade angular do mecanismo, a equação do comprimento  $\Delta S_{(cabo)}$  é dada por:

$$\Delta S_{(cabo)} = (\omega_{mecanismo})(R_{tambor})(\Delta_{acionamento}), \quad (4.17)$$

sendo  $R_{tambor}$  o raio médio do tambor de armazenamento do cabo (considerado constante na operação de posicionamento).

O valor absoluto  $\Delta S_{(cabo)}$  é numericamente igual ao valor  $\|L(e)\|$ , sendo este o desajuste a ser corrigido. Desta forma, a equação do tempo de acionamento em função do desajuste  $\|L_{(e)}\|$  fica:

$$\Delta t_{(acionamento)} = \frac{\|L_{(e)}\|}{(\omega_{tambor})(R_{tambor})} \quad (4.18)$$

O comprimentos dos cabos  $L_{hip'}$  e  $L_{hip}$  que formam a diferença  $\|L_{(e)}\|$  serão respectivamente tratados como  $L_{hip(j)}$  e  $L_{hip(i)}$ . O comprimento do cabo na versão  $L_{hip(i)}$  e  $L_{hip(j)}$ :

$$L_{hip(i)} = \sqrt{e^2 + f^2 - 2ef \cos 2\pi - \beta - \theta_{sonda(i)}}, \quad (4.19)$$

sendo a posição angular  $\theta_{sonda(i)}$ .

A medida  $\theta_{sonda(i)}$  é obtida por um sensor resistivo que será descrito na Figura 32 e Figura 33. O comprimento do cabo para o ângulo desejado  $\theta_{sonda(j)}$  é dado por:

$$L_{hip(j)} = \sqrt{e^2 + f^2 - 2ef \cos 2\pi - \beta - \theta_{sonda(j)}} \quad (4.20)$$

sendo  $L_{hip(j)}$  o comprimento do cabo na posição angular desejada  $\theta_{sonda(j)}$ .

A sensor utilizado na medição do ângulo é caracterizado por um resistor variável angular, cujo circuito implementado para pode ser observado na Figura 32:

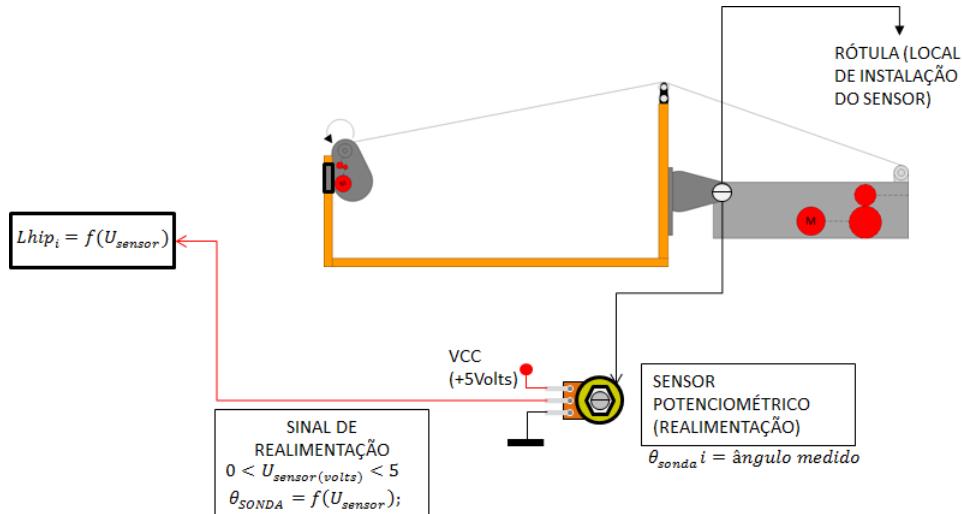


Figura 32 Diagrama de medição do ângulo  $\theta_{sonda(i)}$ . Fonte: Autor.

Pode ser observado na Figura 32 o sensor resistivo (comercialmente conhecido como sensor “Potenciométrico” ou “Poteniômetro”) instalado no ponto de rotação (rótula) do sistema de posicionamento. ângulo  $\theta_{sonda(i)}$  é medido através da leitura de tensão de saída no terminal central do resistor variável (divisor de tensão). O sinal de tensão proveniente do sensor tem sua faixa de variação entre 0 e 5 Volts (tensão de alimentação do divisor de tensão), sendo  $\theta_{sonda}$  função da tensão  $U_{sensor}$ . O modelo do sensor pode ser observado com mais detalhes na Figura 33:

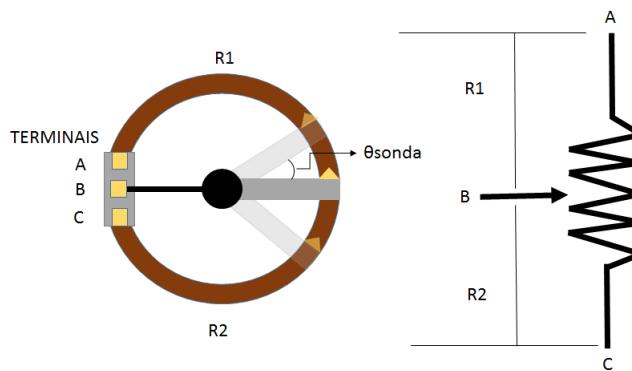


Figura 33 Diagrama do Resistor Variável. Fonte: Autor.

O resistor observado na Figura 33 possui uma trilha circular de carbono sobre um substrato isolante. A medida que a alavancinha de contato tem sua extremidade deslocada ao longo da trilha, o valor de resistência terá seu valor reduzido ou ampliado, mesmo comportamento reportado à tensão no terminal central  $B$  (Figura 33) (Alimentação entrando nos terminais  $A$  e  $C$  na Figura 33). A equação que relaciona a tensão de operação com o

ângulo de inclinação  $\theta_{sonda}$  é dada por:

$$U_\theta = \frac{(Vcc)(\theta_{sonda})}{\theta_{varredura}} + \frac{Vcc}{2}, \quad (4.21)$$

sendo  $U_\theta$  a tensão na no terminal central do divisor de tensão,  $\theta_{varredura}$  o ângulo total de varredura do resistor variável,  $\theta_{sonda}$  o ângulo  $\theta_{sonda(i)}$  e  $Vcc$  o valor da tensão de alimentação ou da fonte.

O valor do ângulo de varredura para fins de estudo é de  $\frac{3\pi}{2} rad$  ( $270^\circ$ ), porém pode assumir diferentes valores, dependendo dos fatores construtivos do sensor. O ângulo  $\theta_{(i)}$  pode ser reformulado em função da tensão pela Eq. 4.22:

$$\theta_{sonda(i)} = \left(\frac{U_\theta}{Vcc} - \frac{1}{2}\right)\theta_{varredura} \quad (4.22)$$

Desta forma, a equação que relaciona o comprimento do cabo  $L_{hip(i)}$  em função da tensão medida nos terminais é dado por:

$$L_{hip(i)} = \sqrt{e^2 + f^2 - 2ef \cos \left[ \left(\frac{U_\theta}{Vcc} - \frac{1}{2}\right)\theta_{varredura} + 2\pi - \beta \right]}, \quad (4.23)$$

sendo  $U_\theta$  a tensão medida no terminal central do resistor e  $\theta_{varredura}$  assumido como  $\frac{3\pi}{2}$ .

Os resultados a cerca das medições do modelo de posicionamento estão mostradas no gráfico da Figura 34:

Pode ser observado na Figura 34 os comprimentos do cabo no instante  $j$  para os ângulos-limite na situação de mergulho da sonda ( $\theta_{sonda} = -45^\circ$ ) e na situação de içamento ( $\theta_{sonda} = 45^\circ$ ). O comprimento do cabo possui seu limite inferior, de acordo com as medidas em aproximadamente 0,12 metros (120mm) enquanto seu limite superior encontra-se em aproximadamente 0,17 metros (170mm).

O comprimento do cabo  $L_{hip(i)}$  calculado em função da tensão  $U_\theta$  medida no resistor variável em configuração de divisor de tensão segue os mesmos parâmetros, onde a faixa varia entre 1 e 4,5 Volts (tensões menores ou maiores do que os limites estipulados) não correspondem ao comportamento real do sistema de posicionamento. Cálculo do comprimento  $L_{hip(j)}$  em função da tensão  $U_\theta$ , com base na Eq. 4.23 pode ser observada na Figura 35:

Pode-se observar que o comprimento  $L_{hip(i)}$  segue uma curva não linear. Deve-se ressaltar que o sensor utilizado possui comportamento linear, ou seja possui a tensão  $U_\theta$

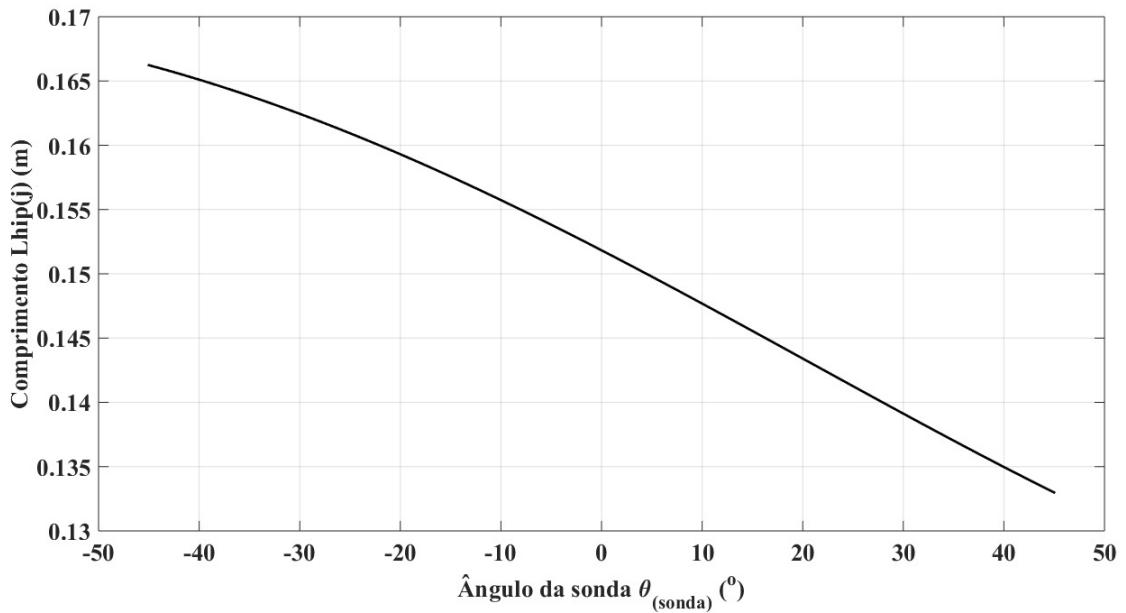


Figura 34 Gráfico do comprimento do cabo ( $L_{hip(j)}$ ) em função do ângulo de inclinação  $\theta_{sonda}$ ) para os ângulos de  $-45^{\circ}$  à  $45^{\circ}$ . Fonte: Autor.

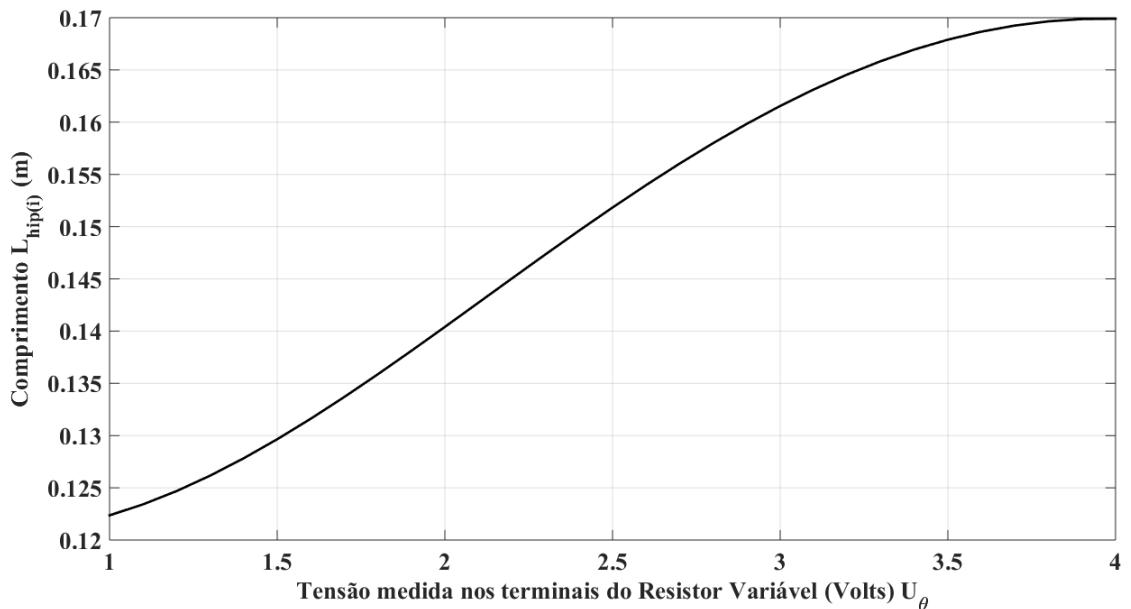


Figura 35 Gráfico do comprimento  $L_{hip(i)}$  do cabo em função da tensão  $U_{\theta}$  variando entre 1 e 4,5 Volts. Fonte: Autor.

como função linear do ângulo de inclinação  $\theta_{sonda}$ . Desta forma a não linearidade observada principalmente próximo aos valores máximo e mínimo são decorrentes do modelo.

Foram realizadas medições de ângulos em operações de levantamento da sonda  $\theta_{sonda}$  sempre crescente e em uma operação de levantamento e posterior abaixamento do sistema de sondagem. Essas aquisições podem ser melhor observadas na Figura 36:

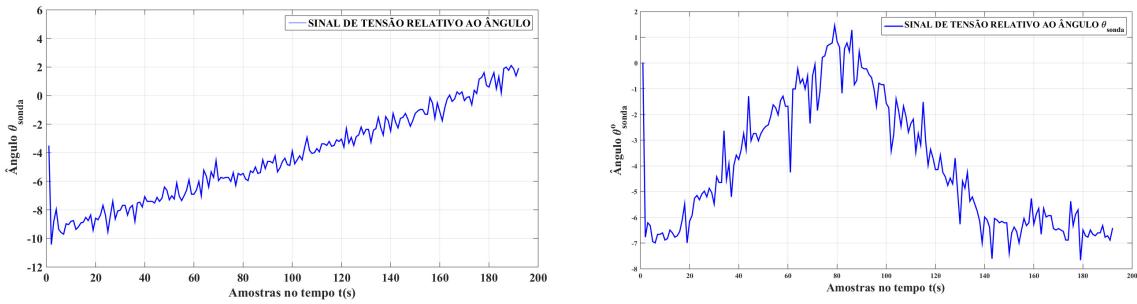


Figura 36 Gráficos Relativos às aquisições do sinal de ângulo calculado a partir da medição de tensão no resistor variável. Operação de levantamento (a); Operação de levantamento e posterior abaixamento do sistema de sondagem (b). Fonte: Autor.

Pode ser observado na Figura 36 (a) a operação de levantamento tendo inclinação inicial aproximada de  $-10^\circ$  e finalizando a operação em aproximadamente  $2^\circ$ . Na Figura 36 (b) pode-se observar que a operação inicia com o levantamento da sonda, partindo de um ângulo aproximado de  $-7^\circ$  indo até o pico máximo de aproximadamente  $2^\circ$  e retornando para posição de mergulho do sistema de sondagem em aproximadamente  $-7^\circ$ . Existem erros de medição associados à fatores construtivos, vibrações e atrito não tratados no escopo deste trabalho.

As características construtivas do sensor estão descritas na Tabela 2:

Tabela 2 Características principais do sensor potenciométrico utilizado na medição de ângulo no hardware de aquisição de sinais.

Característica	Unidade	Valor
Constante Angular	$K$ ( $Ohm/rad$ )	2.3873e+03
Valor de Resistência	$R$ (Ohms)	10e+03
Delta de Varredura	$\theta_{max}$ (rad)	4.1888
Material	Resistor de trilha de carbono	

As operações de posicionamento (levantamento e levantamento com posterior abaixamento) foram realizadas, cada uma, em um período aproximado de 96s. À uma frequência de amostragem de 2Hz, foram coletadas 192 amostras de tensão que foram posteriormente transformadas em valores de ângulo de inclinação  $\theta_{sonda}$ .

## 5 MÉTODOS IMPLEMENTADOS

### 5.1 Geração do Banco de dados

Para que haja o treinamento da Rede Convolucional, é necessário um vasto banco de dados [9] [15] [8]. Para o problema em questão, houve necessidade de desenvolvimento de uma ferramenta automática para aquisição e rotulação dos dados (imagens associadas aos respectivos ângulos de inclinação).

Inicialmente, para o desenvolvimento da ferramenta, foi analisado o processo de aquisição de dados, além de como seria processo de aprendizado da *CNN* (descritos no capítulo 3 deste trabalho) (Figura 37). A partir dessa análise foi desenvolvido um modelo de aquisição dos dados e posterior treinamento descrito pela Figura 37:

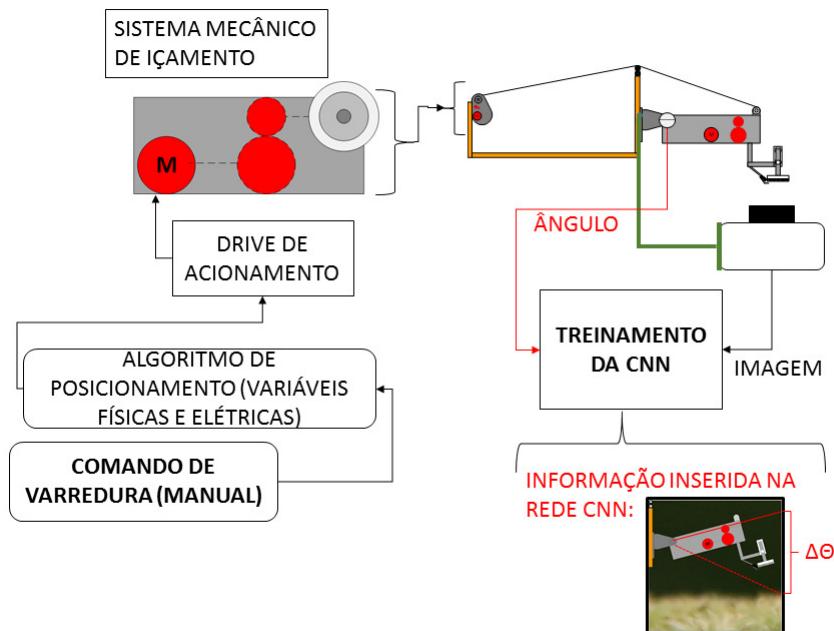


Figura 37 Modo de Treinamento da Rede Neuronal Convolucional, a ser implementado no sistema de posicionamento. Fonte: Autor.

A representação da posição angular (imagem vista no canto direito inferior da Figura 37) e seu respectivo  $\Delta_\theta$ , necessários ao posicionamento desejado, serão inseridos na entrada da *CNN* onde deverão ser extraídas características de inclinação e posteriormente comparação ao rótulo ( $\Delta_\theta$ ) da respectiva imagem.

O cálculo do ( $\Delta_\theta$ ) é realizado através da medição feita pelo sensor de ângulo (Figura

33), além de parâmetros anteriormente inicializados (ex: considera-se o menor valor de ângulo medido pelo sensor no processo de treinamento como o ângulo desejado para o qual a sonda deverá ser posicionada).

É importante observar que complexidade da imagem ou o tamanho da mesma está diretamente relacionada ao número de camadas convolucionais necessárias, assim como ao número de dados de imagem para o treinamento [17].

## 5.2 Descrição da ferramenta para geração do Banco de Dados

O ferramenta para geração do banco de dados tem como base de funcionamento a extração inicial de quadros ou imagens do vídeo (sensor de imagem com tecnologia *CMOS*), com resolução aproximada de dois Mega Pixels- 2MP), juntamente com aquisição dos valores de ângulo, medidos por um *hardware* de aquisição de dados baseado na plataforma Arduíno. Estes valores são posteriormente transferidos pela porta serial à um servidor (cuja taxa de transferência de dados é configurada em 115.200 baud- cento e quinze mil e duzentos bits por segundo), digitalizados por um conversor analógico- digital com resolução de 10 (dez) bits. A medida em que os valores são adquiridos ocorre sua concatenação e posterior reescalonamento com base no número de *frames* (imagens).

A estrutura do algoritmo para geração automática do banco de dados, será mostrada na Figura 38:

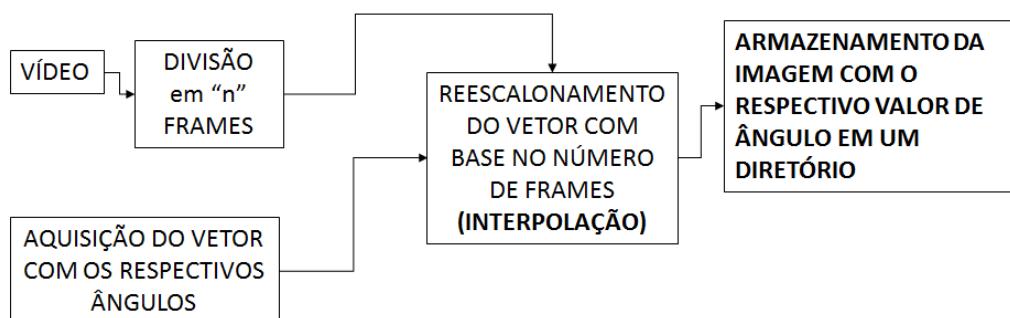


Figura 38 Ferramenta proposta para automatização do banco de dados de treinamento.  
Fonte: Autor.

Pode ser observado na Figura 38 necessita de um vídeo em formato *AVI- Audio Video Interleave* (Formato de Armazenamento de Audio e Vídeo) do sistema de posi-

onamento em funcionamento (no presente trabalho foi gerado um vídeo com a operação de içamento), associado com os valores adquiridos pelo sensor de ângulo implementado.

O modelo do algoritmo proposto na Figura 38 tem como finalidade reduzir o tempo de implementação de uma rede neuronal convolucional em sistemas físicos reais, haja vista que o banco de dados será criado automaticamente. O reescalonamento do vetor (uma das tarefas realizadas pelo algoritmo-Figura 38) é necessário para que haja bijeção entre o quadro (imagem) e o vetor de ângulo (mecanismo utilizado-interpolação).

### 5.2.1 Aquisição de sinais no modelo real

A aquisição do ângulo de inclinação da sonda (Figura 32, Figura 33, Figura 34 e Figura 35) foi realizada através de um resistor variável (Figura 32 e Figura 33). Para aquisição da imagem foi utilizado um sensor de imagem tipo CMOS-Complementary Metal Oxide Semiconductor (Semicondutor de óxido metálico Complementar) cujos parâmetros estão organizados na Tabela 3:

Tabela 3 Características principais do sensor de imagem utilizado

Característica	Unidade	Valor
Taxa de Quadros (utilizada no projeto)	(FPS)	25
Resolução	1600x1200 (pixels)	1920000.00

O sensor de imagem foi implementado na parte lateral direita da plataforma robótica (orientação da câmera voltada para o sistema de sondagem).

O vídeo adquirido pelo sensor CMOS, assim como o sinal do sensor potenciométrico (Tabela 2 e Tabela 3), foram extraídos sem nenhuma filtragem prévia, tentando manter o sinal mais próximo possível dos valores adquiridos em cada *frame*. O funcionamento desta ferramenta pode ser observada (processo de aquisição e geração automática do banco de dados no modelo real) na Figura 39:

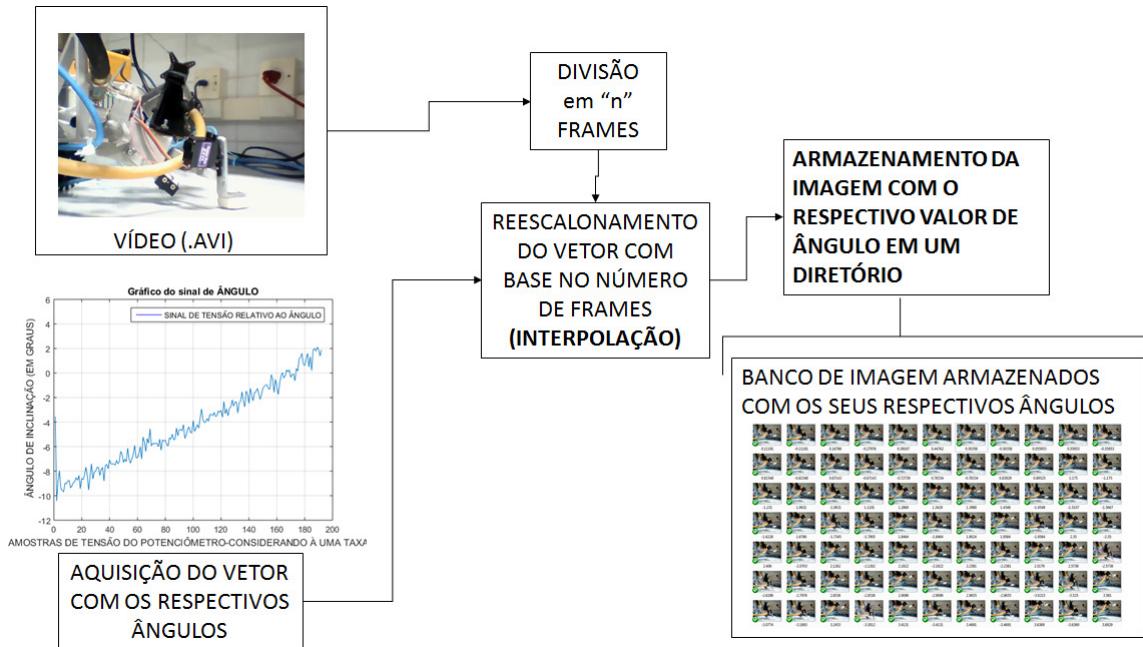


Figura 39 Geração do banco de dados do sistema real de sondagem pela ferramenta proposta. Fonte: Autor.

Como pode ser observado na Figura 39 o sinal do sensor de ângulo absorve as imperfeições de construção e instalação do sensor utilizado, tornando ruidoso o sinal de inclinação adquirido (sinal ruidoso visto graficamente no canto inferior esquerdo da Figura 39). Houve então, a necessidade da montagem de um modelo animado de maior simplicidade quanto à forma e cenário de operação, com movimentos correspondentes a um modelo matemático previamente desenvolvido.

### 5.2.2 Modelo Virtual proposto para comparação

O modelo virtual simplificado foi desenvolvido no software *Algodoo*<sup>®</sup> (*software* livre de engenharia física, criado por Emil Ernerfeldt, nomeado inicialmente como “Phun”), liberado em 17 de Dezembro de 2007, sofrendo atualizações desde então. O modelo criado neste software leva em consideração valores reais da plataforma como massa, gravidade, resistência do ar, atrito, proximidade com os pontos de mudança de velocidade do mecanismo (caracterizados tecnicamente como “Pontos Mortos” do mecanismo), tornando-o o mais próximo da realidade. As posições da extremidade da sonda são armazenadas em um arquivo do tipo *CSV- Comma Separated Values* (arquivo cujos valores são separados por vírgula), sendo posteriormente extraídos ferramenta proposta e reescaladas para se adequar a ordem dos *frames*. O modelo virtual em simulação no *software* pode ser

observado na Figura 40:

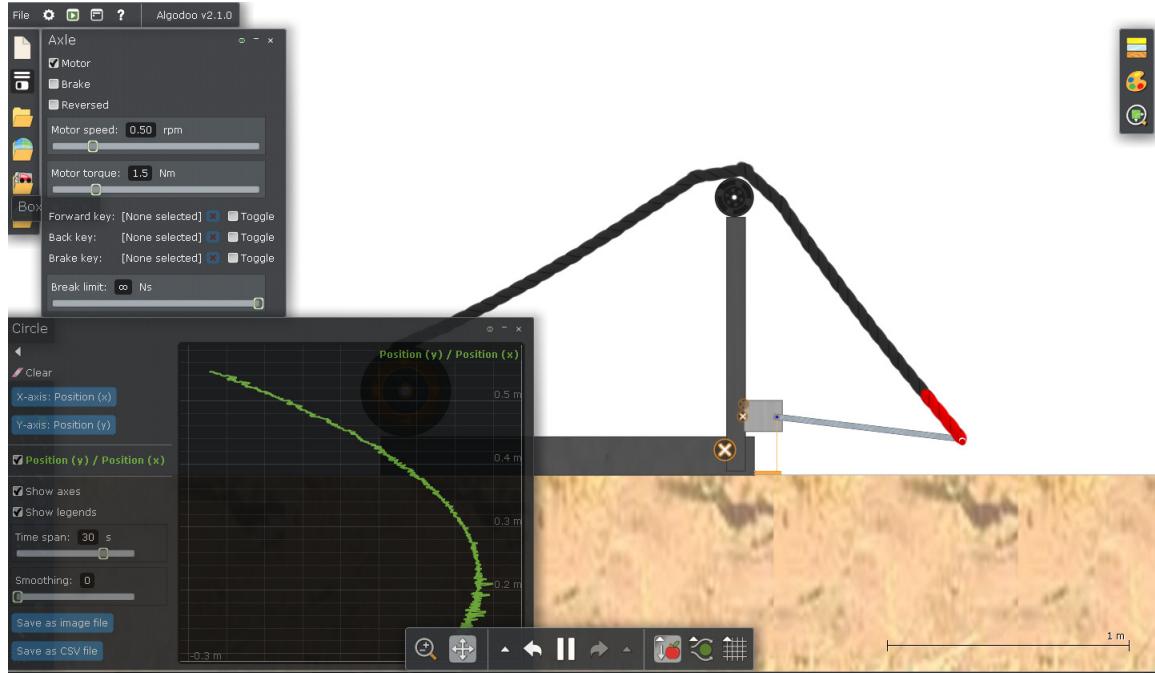


Figura 40 Modelo virtual animado em plataforma virtual de desenvolvimento. Fonte: Autor.

A equação que transforma os valores de posição do ponto de extremidade (circunferência vermelha na extremidade da sonda na Figura 40) é representada por:

$$\theta_{cc} = \sin^{-1}\left(\frac{y_{cc}}{\sqrt{x_{cc}^2 + y_{cc}^2}}\right) \quad (5.1)$$

sendo  $\theta_{cc}$  o ângulo resultante calculado através das posições corrigidas  $x_{cc}, y_{cc}$  em relação aos pontos centrais da circunferência (deslocada no ponto de rótula), formada pelo sistema de posicionamento.

### 5.2.3 Método de Interpolação

Para o desenvolvimento do método de reescalonamento, o vetor com os valores de ângulo exigiu o redimensionamento para uma matriz quadrada, cuja primeira linha fosse formada por todo o vetor de ângulos medidos (tanto no modelo real, como no modelo virtual animado). Considerando  $n_\theta$  o número de elementos do vetor  $\mathbf{v}_\theta$  (vetor de posição angular que formará a primeira linha da matriz  $\mathbf{M}_\theta$ ), é necessária uma matriz para o reescalonamento. Inicialmente essa matriz formada em sua primeira linha pelo vetor  $\mathbf{v}_\theta$ , tendo as outras  $n_\theta - 1$  linhas preenchidas por elementos nulos e,  $n_\theta$  colunas. A solução

adotada para o reescalonamento do vetor de ângulos  $\mathbf{M}_\theta$  de ordem  $n_\theta$  está apresentada na Eq. 5.2 :

$$R_e = \frac{n_{frames}}{n_\theta} \quad (5.2)$$

sendo  $R_e$  o fator de escala para nova matriz,  $n_{frames}$  o número de quadros extraídos do vídeo.

O método de interpolação utilizado no presente estudo é do tipo bilinear. Em geral esse método é utilizado em aplicações voltadas à reamostragem de imagens, com finalidade de minimização do efeito de serrilhamento ocasionado pelo método de interpolação tipo *Nearest Neighbor* (elementos adjacentes mais próximos) [18]. O método adotado realiza interpolações lineares horizontais e verticais entre os elementos da matriz de reescalonamento  $\mathbf{M}_\theta$ . Considerando uma imagem como uma  $f(x, y)$  (Figura 41), será representada graficamente a interpolação bilinear em uma matriz:

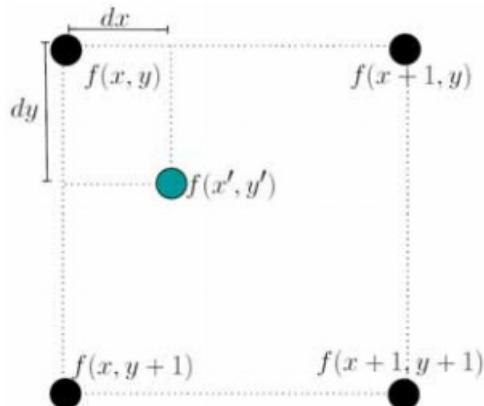


Figura 41 Interpolação bilinear aplicada a uma função  $f(x,y)$ . Fonte:(Dourado.,2014).

Pode-se observar na Figura 41 os pontos em preto na extremidades correspondendo aos elementos adjacentes e o ponto em verde ao centro demonstrando o elemento a ser calculado com base no método de interpolação bilinear.

A equação que demonstra a interpolação bilinear apresentada na Figura 41:

$$f(x', y') = (1-d_x)(1-d_y)f(x, y) + dx(1-d_y)f(x+1, y) + (1-d_x)d_yf(x, y+1) + d_xd_yf(x+1, y+1) \quad (5.3)$$

Pode ser observado na Eq. 5.3 um novo ponto  $(x', y')$  modelado pelos valores das

linhas e colunas adjacentes, ponderado pelas distâncias ao redor do valor desconhecido. O método possui como característica a redução dos valores interpolados, tanto entre os valores das colunas (devido à interpolação com os elementos nulos da matriz  $\mathbf{M}_\theta$ ), como entre os valores das linhas (a interpolação horizontal também é ponderada pelas distâncias verticais).

#### 5.2.4 Background Subtraction e Chroma-Key

O conjunto de imagens e ângulos adquiridos no modelo virtual necessitam ser inseridos em diferentes ambientes ou cenários devido a necessidade de generalização do modelo a ser aprendido pela *CNN*(posicionamento autônomo da sonda em diferentes cenários).

Devido a necessidade da inserção do modelo em diferentes ambientes (10 cenários diferentes para o caso em estudo), realiza-se o “*Background Subtraction*”, ou segmentação do objeto de interesse (caracterizado como “*Foreground*”) do cenário, (caracterizado como “*Background*”).

O *Background Subtraction* é uma técnica de processamento de imagem que utiliza de algoritmos de processamento de imagem e modelos estatísticos [6] [7], desenvolvida com a finalidade de segmentar o objeto de interesse do cenário onde este está inserido. O processo segmentação, do cenário, subtração do cenário primário e inserção do objeto em um cenário ou “*Background*” Secundário, torna-se uma técnica comercialmente conhecida como “*Chroma-Key*” (termo esse já em desuso), modificando o cenário ou ambiente em que um objeto qualquer possa estar inserido, não sendo necessária a presença física do mesmo. O objeto, entretanto, necessita estar primariamente inserido em um cenário limitado à critérios de cor previamente estabelecidos [19]. O sistema de cor das imagens inseridas na ferramenta é o sistema *RGB* (*Red*, *Green*, *Blue* (referência as matrizes ou planos de cor, contidos em cada imagem ou quadro adquirido)).

Para realizar a mudança de cenário, ou mudança de fundo (“*Background*”), foram detectados os limiares de cor em cada plano. No problema em estudo, foi utilizado um cenário cujas intensidades dos planos *RGB* são todos de valores máximos (caracterizando a cor “Branca”). Os limiares para segmentação entre fundo (cenário) e objeto (sonda)

estão caracterizados em parte pelo  $v_{limiar}$ :

$$\mathbf{v}_{limiar} = [limiar_{red}, limiar_{green}, limiar_{blue}] \quad (5.4)$$

sendo  $[limiar_{red}, ,]$ ,  $[, limiar_{green}, ]$  e  $[, , limiar_{blue}]$  os valores de limiar correspondentes aos planos Vermelho (*Red*), Verde (*Green*) e Azul(*Blue*). No desenvolvimento do método o vetor de limiar  $\mathbf{v}_{limiar}$  assume os valores  $[200, 200, 200]$ , tendo como padrão para intensidade de cor o valor máximo de 255 (a cor “branca” ou cor de “intercessão de espectro” em uma imagem é formada por um vetor  $\mathbf{v}$  com intensidades  $[255, 255, 255]$ ). A forma de extração do objeto no fundo branco é mostrada na matriz formada pela seguinte função lógica:

$$\mathbf{M}_{(red/green/blue)} = \begin{cases} 1 & \text{if } (\mathbf{pixel}[\mathbf{R}, \mathbf{G}, \mathbf{B}] - \mathbf{v}_{limiar}) \leq [55, 55, 55] \\ 0 & \text{if } (\mathbf{pixel}[\mathbf{R}, \mathbf{G}, \mathbf{B}] - \mathbf{v}_{limiar}) > [55, 55, 55] \end{cases} \quad (5.5)$$

A equação 5.5 extrai a máscara (matriz cuja posição dos elementos unitários correspondem aos elementos do objeto e os elementos nulos correspondem ao cenário). Onde será adicionado o cenário, sendo  $\mathbf{pixel}[\mathbf{R}, \mathbf{G}, \mathbf{B}]$  o valor do respectivo pixel (considerando um vetor) da imagem original em fundo branco, varrido pela função lógica. Desta forma, para montagem do cenário ao fundo do objeto, recorre-se a equação Eq.5.6:

$$\mathbf{I}_{final(i,j)} = [pixel_{red(i,j)}\mathbf{M}_{red(i,j)}, pixel_{green(i,j)}\mathbf{M}_{green(i,j)}, pixel_{blue(i,j)}\mathbf{M}_{blue(i,j)}] \quad (5.6)$$

sendo  $M_{red(i,j)}$ ,  $M_{green(i,j)}$  e  $M_{blue(i,j)}$  os respectivos elementos da matrizes  $\mathbf{M}_{red}$ ,  $\mathbf{M}_{green}$  e  $\mathbf{M}_{blue}$  resultantes da matrizes geradas pela função lógica  $\mathbf{M}_{(red/green/blue)}$ , sendo  $pixel_{red(i,j)}$ ,  $pixel_{green(i,j)}$  e  $pixel_{blue(i,j)}$  os respectivos pixels do cenário de fundo inseridos em cada plano de cor. Desta forma a imagem ficará conforme apresentado na Figura 42:

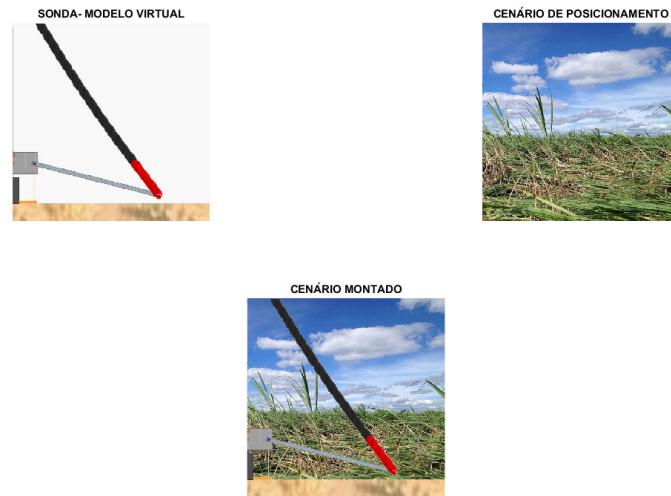


Figura 42 Demonstração do “*Chroma-Key*” aplicado a cada quadro. Fonte: Autor.

Pode ser visto no canto superior esquerdo da Figura 42, o modelo virtual (cenário de fundo em cor ”branca”), o cenário de fundo onde o objeto ou modelo será inserido (canto superior direito) e posteriormente a imagem ou quadro da sonda montada no cenário desejado (centro).

### 5.2.5 Comparação dos Modelos Físico e Simulado

Foram realizadas aquisições finais quanto aos valores de ângulo, em uma operação de içamento, no sistema de posicionamento real, assim como no modelo virtual. O gráfico dos vetores reescalados  $\mathbf{v}$  são apresentados na Figura 43:

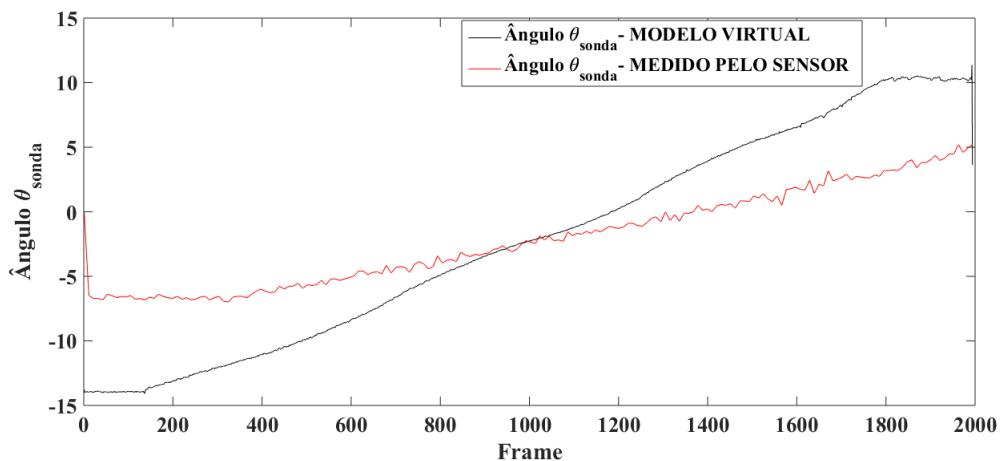


Figura 43 Gráfico de comparação entre o vetor de aprendizado gerado no teste real e o vetor de aprendizado gerado no modelo virtual. Fonte: Autor.

Pode ser observado no vetor de posição angular (adquirido pelo sensor no modelo real), as imperfeições construtivas e interferências na aquisição gerando um sinal ruidoso (linha vermelha no gráfico da Figura 43), em comparação ao sinal adquirido no modelo virtual (linha preta no gráfico). Há também um aumento da faixa de deslocamento angular que, no modelo virtual, consegue alcançar um  $\Delta_\theta$  de aproximadamente  $25^\circ$ , em comparação à um  $\Delta_\theta$  de aproximadamente  $7^\circ$  no modelo real.

### 5.3 Materiais utilizados na Implementação Final

Além dos sistemas de aquisição de imagem (descritos na Tabela 3), existe ainda o *drive* de acionamento é composto por uma placa de desenvolvimento *Arduino®* modelo *UNO*, dispositivo de acionamento do motor elétrico (tensão de operação de 5 Volts) e motor elétrico DC (*Direct Current*) com tensão de operação nominal em 5 Volts e corrente de operação em torno de 800 miliampéres).

### 5.4 Método de classificação pela CNN

Os sistemas de aquisição da imagem (*CÂMERA*), processamento da imagem (*SERVIDOR*) e acionamento do motor elétrico (*DRIVE DE AÇÃOAMENTO*), utilizados no processo de aproximação podem ser observados na Figura 44:

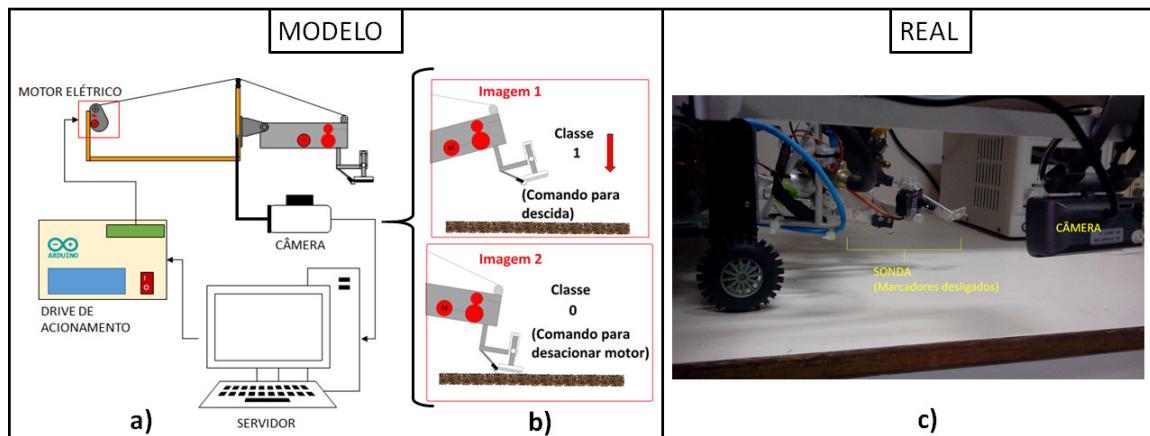


Figura 44 Diagrama de funcionamento de sistema de aproximação. a) Método de aquisição e classificação para reposicionamento do sistema de sondagem ; b) Imagens das classes de inclinação e a descrição dos comandos de reposicionamento; c) Sistema Real de Aproximação.

O processo de aproximação se divide em 3 etapas sendo: aquisição da imagem por

uma câmera posicionada na parte lateral da estrutura mecânica (primeira etapa- Figura 44a) e classificação da imagem por um algoritmo implementado no servidor (segunda etapa). Com base no valor numérico de classificação é gerado um sinal para o acionamento ou parada do motor elétrico (terceira etapa- Figura 44b), reposicionando da sonda.

As posições angulares da sonda (Figura 44b), correspondem a classe 1 (sonda distante do ponto desejado) e classe 0 (sonda no ponto desejado). A classe 1 gera na saída do dispositivo de acionamento um sinal de corrente elétrica para o acionamento do motor elétrico e a classe 0, suspende o fornecimento desse sinal.

O braço mecânico (caracterizado como *Sonda* na páginas 16 e 51) é um mecanismo de 1 Grau de Liberdade (G.D.L), reposicionado angularmente através de um sistema de içamento por cabo (movido pelo motor elétrico), gerando o movimento de subida e descida, movendo a sonda ao ponto desejado.

## 5.5 Método de classificação utilizando a CNN

Os processos de extração de características e posterior classificação pela *CNN* estão descritos na Figura 45:

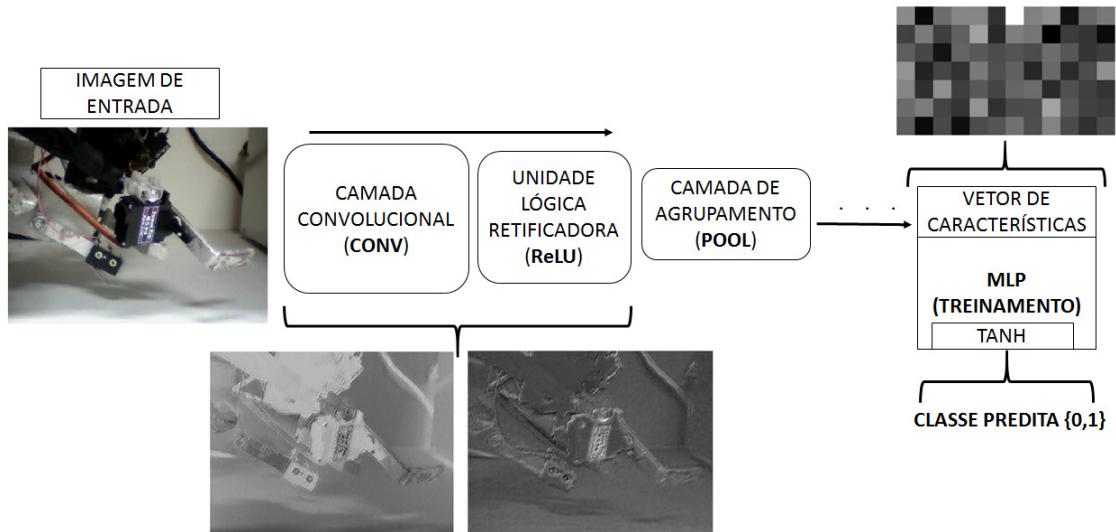


Figura 45 Diagrama de blocos da *CNN*. Imagem do sistema mecânico real sendo convoluída com o filtros na camada convolucional *CONV*, retificada pela camada *ReLU* e agrupada ou sub-amostrada na camada *POOL*, sendo gerado então um conjunto de valores caracterizado como *VETOR DE CARACTERÍSTICAS* inserido na rede *Multi-Layer-Perceptron* (MLP) tipo *Fully-Conected*.

Podem ser observados na Figura 45 os processos de convolução, retificação lógica e agrupamentos de forma a gerar um vetor de características para ser posteriormente

classificado pela estrutura *Fully Conected* (MLP Treinada). É importante ressaltar que todos os processos descritos na Figura 45 fazem parte de uma só rede, descritos em partes para o melhor entendimento. O algoritmo de aprendizado de toda rede tem como base o *Steepest Gradient Descent*.

### 5.5.1 Banco de Dados- Modelo Físico

Para o treinamento final da *CNN* foi realizada uma nova aquisição de imagens (gerado um novo banco de dados), com aquisição de 10.000 imagens ou (*frames*) em 5 diferentes cenários (2.000 imagens por cenário), estruturados de acordo com o diagrama da Figura 46.

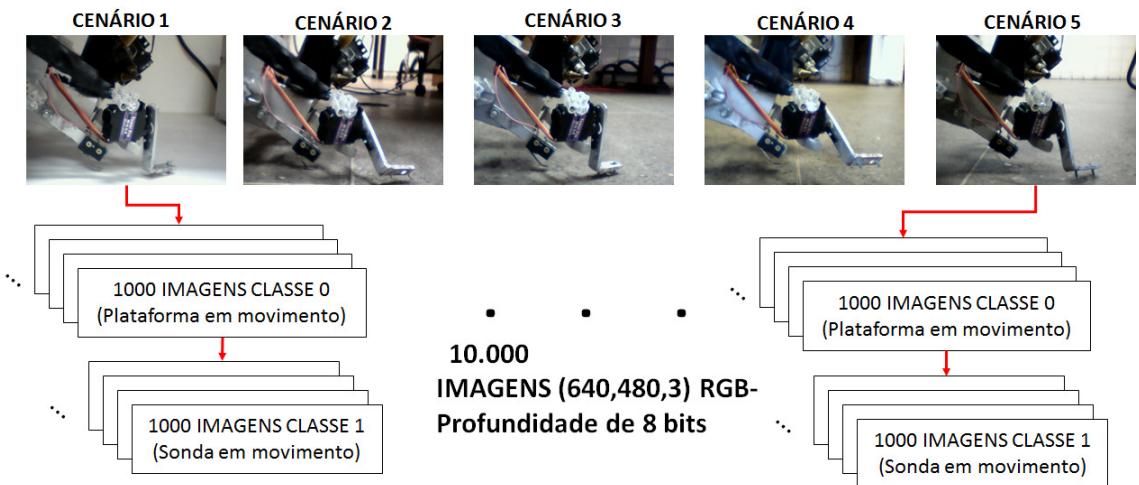


Figura 46 Estrutura do Banco de Dados usado para o treinamento da rede Convolucional.

O banco de imagens foi dividido e rotulado em imagens *classe 0* e imagens *classe 1*. A classe 0 corresponde o sistema mecânico próximo ao solo (ponto desejado), em diferentes cenários. A classe 1 correspondem a toda estrutura mecânica distante do ponto desejado (Figura 46), fixa em um cenário.

A imagens do banco de dados final possuem como dimensão original de 640X480X3 pixels (RGB), reduzida pelas camadas convolucionais à um vetor de características com 256 elementos para entrada na rede tipo *Fully Conected*.

A estrutura da rede implementada é composta por 5 camadas convolucionais (*Conv*) e 5 camadas de agrupamento (*Pool*). A configuração da MLP se resume à 2 camadas escondidas, somadas à uma camada de entrada e uma camada de saída.

## 5.6 Método de classificação utilizando Marcadores

O método de extração de características, baseado em limiar de decisão, utiliza a área formada entre os marcadores de frequência fixa no espectro visível. Esses são implementados na forma de Diodos Emissores de Luz (*LEDs*), descritos na página Figura 47.

O diagrama do sistema mecânico com os marcadores (enumerados como *LED1*, *LED2* e *LED3*-Ponto desejado para o reposicionamento), assim como o método de classificação estão descritos na Figura 47:

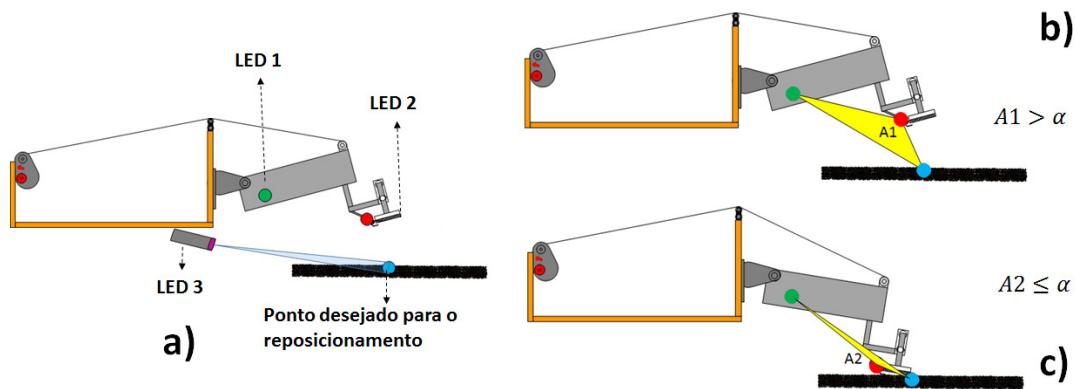


Figura 47 Diagrama dos Marcadores. a) Localização dos Marcadores; b) Área  $A_1$  entre os marcadores que ultrapassa o limiar  $\alpha$  (sonda distante do ponto desejado); c) Área  $A_2$  entre os marcadores que não ultrapassa o limiar  $\alpha$  (sonda estática no ponto desejado).

Pode ser observada as diferentes áreas entre os marcadores (enumerados como *LED1*, *LED2* e *LED3*- Figura 47a) com a sonda próxima e distante do ponto desejado para o reposicionamento (Figura 47 b e c).

O cálculo da área é dado por:

$$A = \frac{1}{2} \left\| \det(\mathbf{M}) \right\|, \quad (5.7)$$

sendo  $A$  a área formada entre os marcadores e a matriz  $\mathbf{M}$  formada pela média das coordenadas (baricentros) dos pixels de máxima intensidade.

A matriz  $\mathbf{M}$  é dada por:

$$\mathbf{M} = \begin{pmatrix} x_{red}^- & y_{red}^- & 1 \\ x_{green}^- & y_{green}^- & 1 \\ x_{blue}^- & y_{blue}^- & 1 \end{pmatrix}$$

sendo  $\bar{x}$ ,  $\bar{y}$  as coordenadas (médias) dos pixels máximo valor, nos respectivos planos *red*, *green*, *blue*.

Baseado na Eq. 5.7, o critério de classificação nas proximidades do ponto desejado é dado por:

$$C = \begin{cases} \text{Classe 1, se } A > \alpha \\ \text{Classe 0, se } A \leq \alpha \end{cases}$$

sendo  $\alpha$  o limiar de decisão. Quando a área entre os marcadores ultrapassar o limiar  $\alpha$ , a sonda deve ser reposicionada para o ponto desejado. Caso contrário, será gerado um comando para finalizar o processo de reposicionamento.

Por último, os valores de área são inseridos em uma função logística para normalização dos valores de área, como pode ser observado na Eq.5.9:

$$z = \frac{L}{1 + e^{(-1)\lambda(A - A_o)}} \quad (5.9)$$

sendo o Valor  $L$  o limite Superior da variável independente  $z$ ,  $\lambda$  a constante de inclinação da curva e  $A_o$  o valor mínimo da área (para o projeto adotado como valor nulo).

O algoritmo de classificação com base no uso dos marcadores coloridos possui o fluxograma descrito na Figura 48.

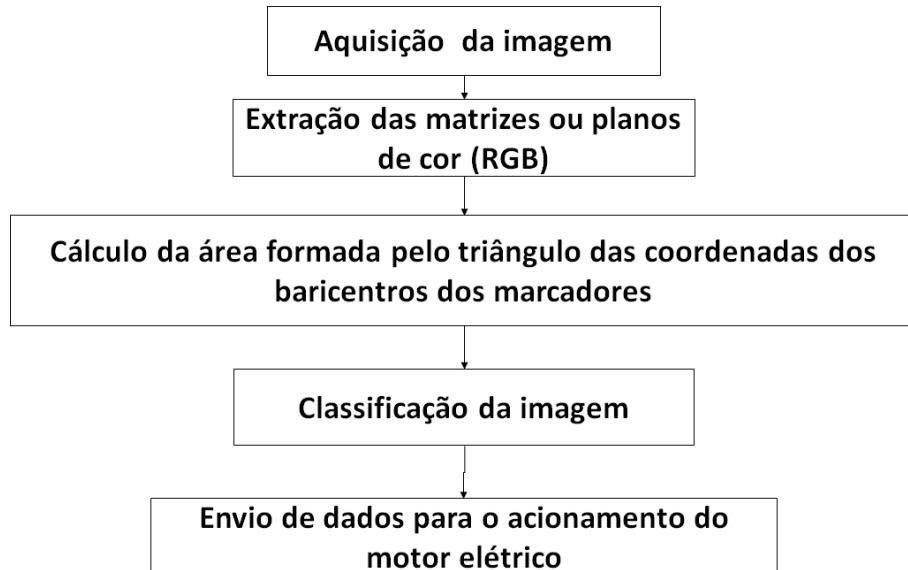


Figura 48 Fluxograma do algoritmo de classificação. Fonte: Autor.

Na Figura 48 pode ser observado que o principal processo anterior à classificação

da imagem é o cálculo da área formada pelos baricentros dos marcadores. Esse tem como base o processamento do histograma de frequência dos píxeis de valor máximo em função de sua posição na imagem. A partir então da coordenada média desses valores máximos em função de sua posição, é calculada analiticamente a área imaginária formada por eles (processo descrito nas equações 5.7 a 5.9).

## 6 RESULTADOS

A aquisição inicial da imagem da sonda pode ser observada na Figura 49 a seguir com os marcadores provisórios:

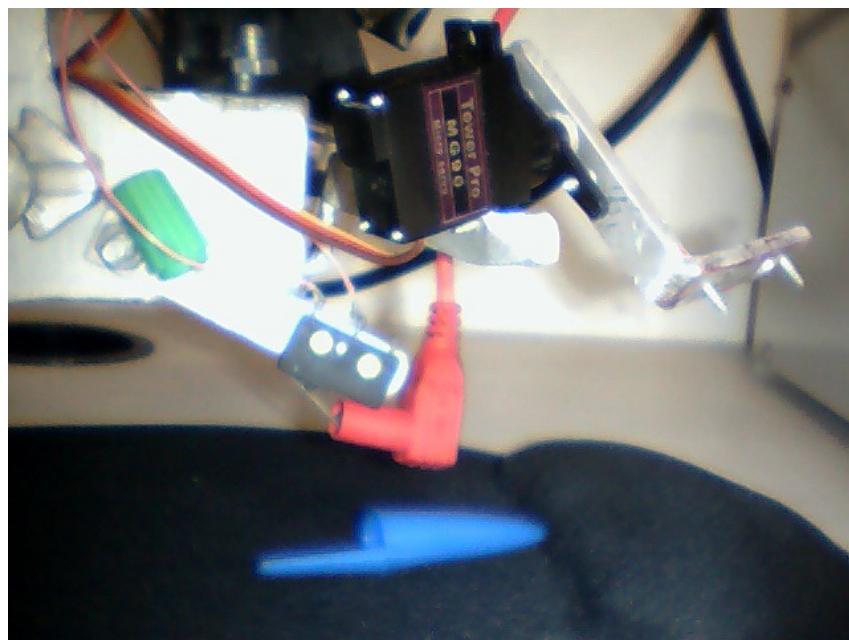


Figura 49 Aquisição da imagem em formato RGB-8 bits. Fonte: Autor.

Pode ser observado na Figura 49 como a imagem é adquirida antes da classificação. Os mapas de agrupamentos dos píxeis de máximo valor são mostradas na Figura 50 e Figura 51:

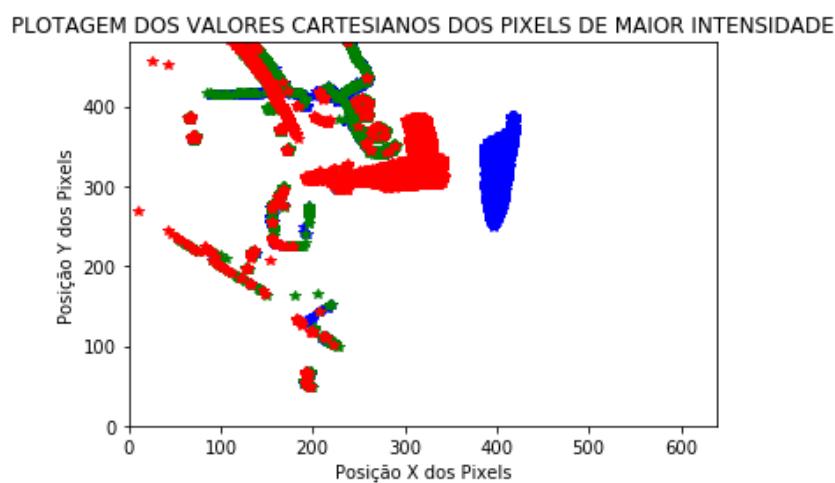


Figura 50 Coordenadas de todos os píxeis de valor máximo de cada plano de cor. Fonte: Autor

Pode-se observar na Figura 50 os píxeis de máximo valor nas suas componentes RGB, correspondendo aos marcadores. É visível a presença de marcadores juntamente com píxeis que não são correspondentes aos mesmos, representando uma deficiência na filtragem prévia da imagem.

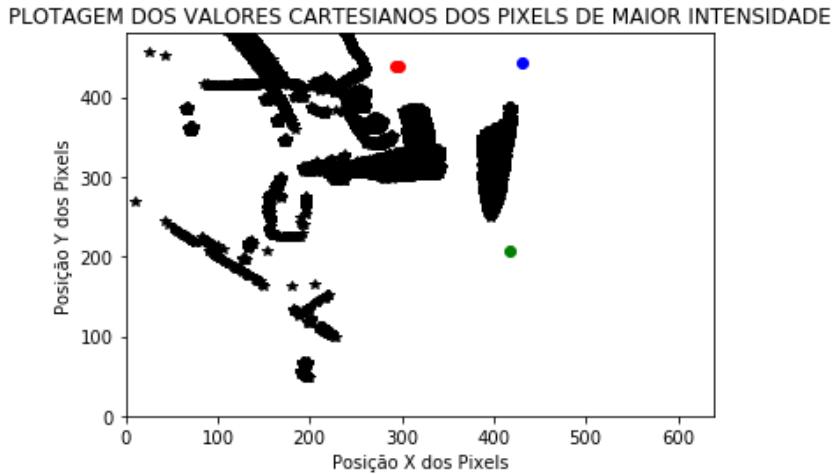


Figura 51 Coordenadas de maior aglomeração dos pixels de intensidade máxima. Fonte: Autor.

As coordenadas dos baricentros dos marcadores podem ser observadas na Figura 51 (pontos vermelho, verde e azul). É observável a mudança de coordenadas que os marcadores sofrem ao longo dos processos de aquisição e processamento da imagem adquirida, não gerando influência sobre o comportamento do algoritmo, haja vista que o modelo se baseia no cálculo da área para classificação.

Na Figura 52 poderá ser observado as respectivas áreas formadas pelos marcadores com a sonda em diferentes posições do objeto:

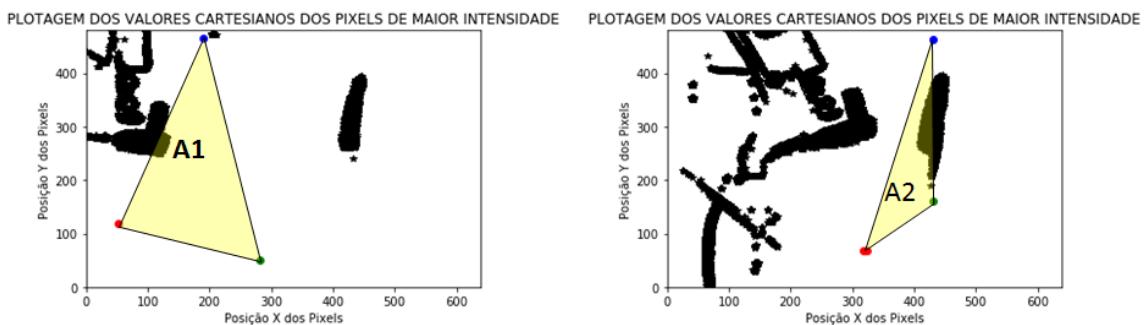


Figura 52 Comparaçao das áreas para classes diferentes de posicionamento (1-esquerda e 0-direita). Fonte: Autor.

Na Figura 52 (lado direito) pode ser observado o aumento da área  $A_1$  em função do distanciamento do objeto ou o inverso observado com a  $A_2$  no lado direito. é visível

que a forma e os lados do triângulo (erros no cálculo da área), porém o comportamento da área permanece o desejado.

Como forma de aumentar a intensidade dos marcadores foram inseridos Diodos Emissores de Luz (LEDs) em frequências específicas para redução do efeito de variação de luminosidade sobre os marcadores. Os marcadores ativos ou LEDs implementados na sonda podem ser observados na Figura 53:

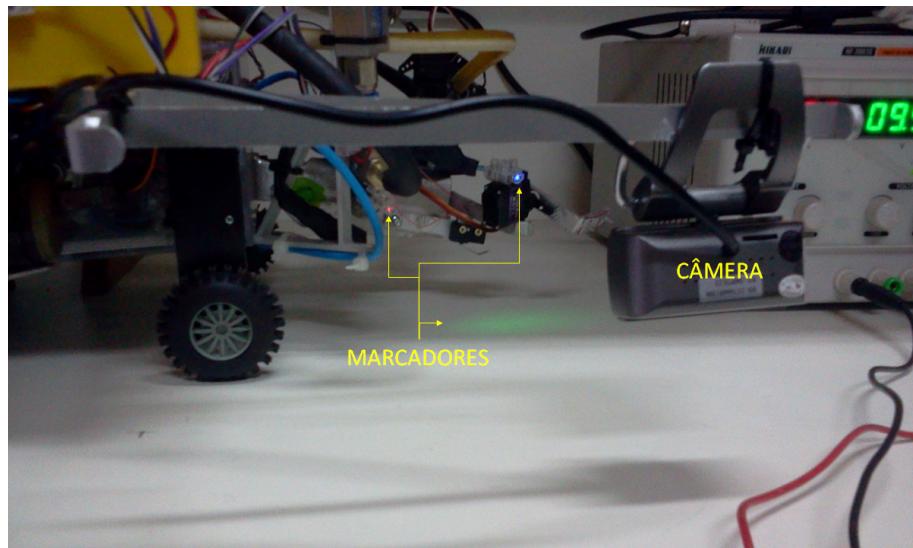


Figura 53 Imagem real da sonda e dos marcadores (ao centro) e a câmera que irá captar a imagem dos marcadores (à direita). Fonte: Autor.

Pode ser observado na Figura 53 que existem variações de luminosidade na imagem que será captada pela câmera. É possível também observar que o terceiro marcador, representado pela aglomeração de raios luminosos na frequência correspondente à cor Verde no espectro visível representa ponto para o qual a sonda deverá ser reposicionada.

No método dos marcadores é assumido que um operador define qual objeto ou ponto para onde a sonda deverá ser aproximada, podendo ser reajustado para cada operação de posicionamento.

Para a verificação de performance dos métodos (*CNN* e Marcadores) foram utilizadas 10.000 imagens sendo 9.000 de treinamento e 1.000 para teste (*CNN*), comparado à 1.000 imagens de teste para o método dos marcadores (performance descrita na Tabela 4). Os resultados apresentados na Tabela 1, avaliam os classificadores em relação as métricas de acurácia (*performance* geral do classificador), precisão (habilidade do classificador em rotular corretamente as amostras) e sensitividade (habilidade do classificador

para encontrar amostras positivas).

Tabela 4 Comparação entre os dois classificadores pelas métricas de Acurácia, Precisão e Sensitividade. Os verdadeiros-positivos correspondem à ( $tp$ ), os verdadeiros-negativos à ( $tn$ ), os falso-positivos à ( $fp$ ) e falso-negativos à ( $fn$ ).

Métricas	Marcadores	CNN
Acurácia ( $\frac{tp+tn}{tp+tn+fp+fn}$ )	51,80%	82%
Precisão (classificação) ( $\frac{tp}{tp+fp}$ )	55,475%	91%
Sensitividade ( $\frac{tp}{tp+fn}$ )	52%	64%

Os valores de área nas diferentes situações para o método dos marcadores estão descritos na 6:

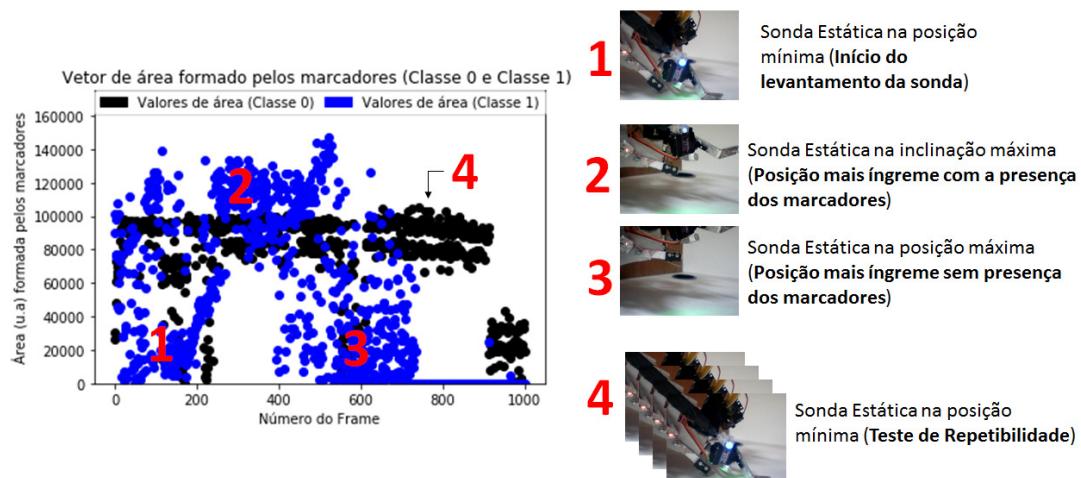


Figura 54 Gráfico dos valores de área agrupados em dois vetores (pontos azuis e pontos pretos), cujos números (em vermelho) indicam: (1) Sonda no início da operação de içamento; (2) Posição cujos marcadores encontram-se nas extremidades da imagem; (3) Sonda fora da área contemplada pela imagem (situação de emergência); (4) Sonda estática na inclinação mínima (verificação da repetibilidade na classificação-pontos pretos).

## CONCLUSÃO

A *CNN* mostrou-se superior em todas as métricas utilizadas (Tabela 4), não necessitando do uso de marcadores, sendo testada em um cenário para o qual essa não foi previamente treinada o que justifica seu uso. O desenvolvimento deste sistema de aproximação (69), permite uma serie de aplicações em sistemas mecânicos, diferindo apenas as características a serem extraídas pela *CNN* (sofrendo novo treinamento e/ou pequenas reconfigurações, caso necessário).

Os valores de inclinação mínima ( $\theta_{inclinacão} \approx -15^\circ$ ) e inclinação máxima ( $\theta_{inclinacão} \approx 30^\circ$ ) correspondem à inclinação da sonda em relação à plataforma (78). Na segunda operação a sonda permanece estática na inclinação mínima (correspondente aos pontos pretos no gráfico).

Pode ser observada instabilidade na extração de características de baixo nível (área entre os marcadores) quando a sonda encontra-se estática na posição mínima (posição de número (4), 6). Os erros de classificação estão em valores próximos ou maiores ao limiar  $\alpha$  fixado em 100000 u.a (pontos pretos na 6), devendo assumir valores menores à 40000 u.a. (erros atribuídos à vibrações e variações de luminosidade durante a aquisição). Refinos na Eq.(5.7) e Eq.(5.9) assim como adição de filtros durante o processamento da imagem (anteriores ao cálculo da área) podem aumentar a acurácia e precisão do classificador.

A posição de número (3) na 6 (caracterizada como situação de emergência), obteve valores de área abaixo do limiar  $\alpha$ , gerando o comando de parada da sonda (comando necessário neste tipo de situação).

Como desvantagens a *CNN* necessita de um vasto banco de dados, sendo necessário, para uso em plataformas agrícolas ou ambientes urbanos, ferramentas que automatizem a aquisição e rotulação das imagens (esforço este já realizado pelo autor no presente trabalho) [20]. Uma outra desvantagem da utilização de Redes Convolucionais é a tendência para o *overfitting* ou sobreajuste para um cenário específico, desvantagem essa que ocorreu em algumas partes do processo de treinamento.

### 6.1 Contribuição Científica da Ferramenta para Geração do Banco de Dados

A ferramenta desenvolvida torna o uso de Redes Neurais Convolucionais mais robusto à problemas de engenharia e visão computacional. Isso ocorre através da geração au-

tomática do banco de dados de treinamento, onde seria necessária uma exaustiva aquisição de dados particular a cada problema. Na Figura 55 a ferramenta está sendo aplicada em um modelo virtual:

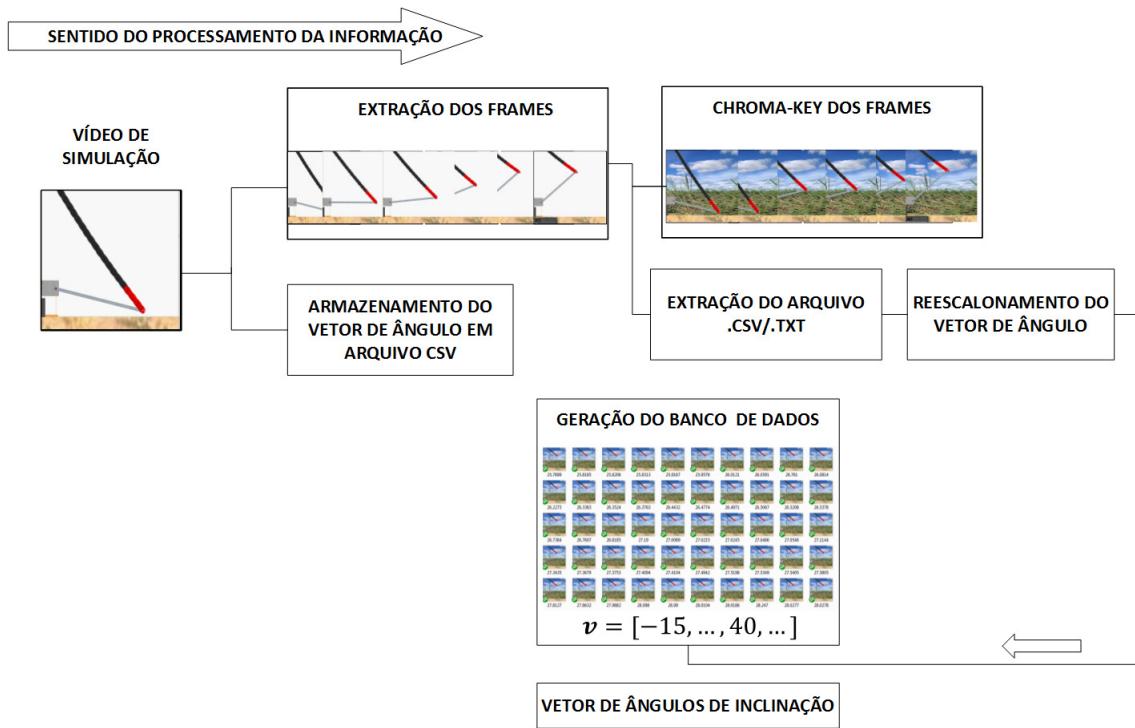


Figura 55 Geração do banco de dados do modelo virtual em cenário desejado. Fonte: Autor.

Pode ser observado na Figura 55 a geração automática do banco de dados no referido modelo computacional, podendo ser replicado em vários outros cenários (minimizando a possibilidade de *overfitting* na rede neuronal), reduzindo assim custos de implementação, ou melhorando o aprendizado de máquina em ambientes totalmente simulados.

## REFERÊNCIAS

- [1] CHONG, E. K. P.; ZAK., S. H. *An Introduction to Optimization, Fourth Edition.* [S.l.]: Jon Wiley Sons, Inc, 2013.
- [2] PEREIRA, V. A. F. V.; SEQUEIRA, J. Low cost object sorting robotic arm using raspberry pi. *IEEE Global Humanitarian Technology Conference - South Asia Satellite (GHTC-SAS)*, p. 1–6, 2014.
- [3] GHAZALI MOHD FAIS MANSOR, M. M. M. K. H.; HUSSAIN, A. Feature extraction technique using discrete wavelet transform for image classification. IEEE, 2007.
- [4] XIAO, B. Principal component analysis for feature extraction of image sequence. IEEE, 2010.
- [5] LIAO, J. J. W. Image feature extraction based on kernel ica. IEEE, 2008.
- [6] OLIVEIRA ANDRÉ LUIZ PINTES, R. C. S. F. E. U. K. M. Jozias Parente de; SILVA, I. S. de Souza e. Fpga architecture for object segmentation in real time. EURASIP, 2006. ISSN 2219-5491.
- [7] OLIVEIRA ANDRÉ LUIZ PINTES, R. C. S. F. E. U. K. M. Jozias Parente de; SILVA, I. S. de Souza e. Fpga architecture for object segmentation in real time. ACM New York, NY, USA ©2006, 2006.
- [8] GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning.* [S.l.]: MIT Press, 2016. <http://www.deeplearningbook.org>.
- [9] GIUSTI JéRÔME GUZZI, D. C. C. F.-L. H. J. P. R. F. F. M. F. C. F. J. S. G. D. C. D. S. L. M. G. A. A machine learning approach to visual perception of forest trails for mobile robots. *IEEE Robotics and Automation Letters*, v. 1, p. 661–667, 2014.
- [10] JALLED, I. V. F. *Object Detection Using Image Processing.* 2016. Disponível em: <<http://www.stat.ucla.edu/ywu/GenerativeConvNet/doc/GConvNet.pdf>>.
- [11] GONZALEZ C., R. W. E. R. *Digital Image Processing. Third Edition.* [S.l.]: Person, 2007.

- [12] TORRALBA, A.; OLIVA, A. Statistics of natural image categories. *Network: Computation in Neural Systems*, v. 14, n. 3, p. 391–412, 2003. ISSN 0954-898X.
- [13] IBARRA, M. V. N. Hempel de; MENZEL, R. Mechanisms, functions and ecology of colour vision in the honeybee. *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioural Physiology*, v. 200, n. 6, p. 411—433, 2014.
- [14] HUBEL, D. H. Evolution of ideas on the primary visual cortex, 1955-1975: A biased historical account. In: *Bioscience Reports 2*. [S.l.]: The Nobel Foundation 1982, 1982. p. 435–469.
- [15] SAXENA, A.
- [16] BRAGA, A. P. L. F. d. C. e. T. B. L. Antônio de P. *Redes Neurais Artificiais-Teoria e Aplicações.-2ed.* [S.l.]: LTC-Livros Técnicos e Científicos Editora S.A.-Rio de Janeiro, 2007.
- [17] GRANDSTRAND, O. Innovation and intellectual property rights. In: FAGERBERG, J.; MOWERY, D. C.; NELSON, R. R. (Ed.). *The Oxford Handbook of Innovation*. [S.l.]: Springer International Publishing Switzerland 2, 2014. p. 818–833. Vol 8689.
- [18] DOURADO, W. B. *AVALIAÇÃO DE TÉCNICAS DE INTERPOLAÇÃO DE IMAGENS DIGITAIS*.
- [19] APRILE M., A. C. A. F. V. M. *SEGMENTATION AND TRACKING OF MULTIPLE OBJECTS IN VIDEO SEQUENCES*. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.140.4597&rep=rep1&type=pdf>>.
- [20] Desenvolvimento de uma ferramenta automática para geração de um banco de dados com base em sistema mecânico simulado para treinamento de uma deep neural network-dnn. In: LNCC (Ed.). *Anais do X Encontro Acadêmico de Modelagem Computacional*. [S.l.: s.n.], 2017. p. 29–39.