



UNIDADE DE ENSINO SUPERIOR DE FEIRA DE SANTANA

YAN NASCIMENTO RAMOS

APLICAÇÃO DE REDE NEURAL CONVOLUCIONAL (DARKNET) EM SISTEMA
INTELIGENTE DE AUXÍLIO À ACESSIBILIDADE DE DEFICIENTES VISUAIS

Feira de Santana

2021

Yan Nascimento Ramos

APLICAÇÃO DE REDE NEURAL CONVOLUCIONAL (DARKNET) EM SISTEMA
INTELIGENTE DE AUXÍLIO À ACESSIBILIDADE DE DEFICIENTES VISUAIS

Trabalho de conclusão de curso da
Unidade de Ensino Superior de Feira de
Santana, como requisito para a obtenção
do título em Bacharel em Engenharia
Mecânica.

Orientadora: Profa. Ma. Ivonete Maciel
Lima Oliveira

Feira de Santana

2021

R175a Ramos, Yan Nascimento.

Aplicação de rede neural convolucional (*darknet*) em sistema inteligente de auxílio à acessibilidade de deficientes visuais. / Yan Nascimento Ramos. Edição do autor. – Feira de Santana, 2021.

75f; il.

Orientadora: Prof^a. Me. Ivonete Maciel Lima Oliveira

Coorientador: Prof^o. Me. Márcio Pereira da Silva Bove

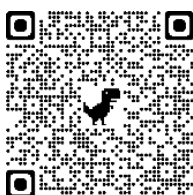
TCC (graduação) Bacharelado em Engenharia mecânica – Unidade de Ensino Superior de Feira de Santana – UNEF – 2021.

Data da defesa: 22/07/2021.

1. Deficientes visuais. 2. Inteligência artificial. 3. Redes Neurais Convolucionais. 4. Tecnologias assistivas. 5. Visão computacional I. Unidade de Ensino Superior de Feira de Santana - UNEF. II. Título.

CDU: 621:004-056.262

Ficha Catalográfica Elaborada por: Deivisson Lopes Pimentel Bibliotecário CRB 5/1562



FOLHA DE APROVAÇÃO

YAN NASCIMENTO RAMOS

APLICAÇÃO DE REDE NEURAL CONVOLUCIONAL EM SISTEMA ELETRÔNICO INTELIGENTE DE AUXÍLIO À ACESSIBILIDADE DE DEFICIENTES VISUAIS

Feira de Santana, ____ / ____ / ____.

Banca Examinadora:

Profa. Ma. Ivonete Maciel Lima Oliveira
Unidade de Ensino Superior de Feira de Santana
Orientadora

Prof. Me. Márcio da Silva Pereira Bove
Unidade de Ensino Superior de Feira de Santana
Coorientador

Prof. Me. Aldi Rui Morais Silva
Unidade de Ensino Superior de Feira de Santana
Examinador

DEDICATÓRIA

Dedico este trabalho a toda comunidade deficiente visual da
República Federativa do Brasil.

AGRADECIMENTOS

Agradeço primeiramente a Deus por me conceder o dom da vida e me conduzir em todo meu trajeto pessoal, acadêmico e profissional com saúde e felicidade.

Ao meu pai, Givaldo da Cruz Ramos, minha mãe, Valdenaide Nascimento Ramos e aos meus irmãos, Iago Nascimento Ramos e Ingrid Nascimento Ramos, por todo apoio, paciência, cumplicidade em todos os momentos de minha vida, principalmente nos mais difíceis. Sou eternamente grato a todos vocês!

A todos os professores (do maternal ao ensino superior) que cruzaram meu caminho, pela paciência e contribuição na construção de meus conhecimentos que culminaram direta ou indiretamente no desenvolvimento deste trabalho.

Em especial, agradeço a professora Ivonete Maciel Lima Oliveira, ao professor Márcio da Silva Pereira Bove e ao professor João Felipe Lacerda da Cruz pela condução desde o planejamento até a finalização deste trabalho.

Por fim, a todos os meus amigos, colegas e familiares que sempre estiveram ao meu lado, mesmo nos momentos mais difíceis desta caminhada.

*“A tarefa não é tanto ver aquilo que ninguém viu, mas pensar o que
ninguém ainda pensou sobre aquilo que todo mundo vê.”*

Arthur Schopenhauer

RESUMO

Ao longo do tempo, os deficientes visuais têm passado por dificuldades de acessibilidade nos mais variados âmbitos de seu cotidiano. Muitas das atividades podem gerar riscos para a integridade física destes indivíduos. Estas problemáticas podem apresentar-se por ocasião de fatores voltados ao fluxo de pessoas ou nos demais perigos intrínsecos na sociedade. Apesar dos meios assistivos, tais como cão-guia e bengala, os deficientes visuais ainda passam por problemas de locomoção devido a baixa eficiência apresentada por tais equipamentos. Com a evolução tecnológica, em especial da Inteligência Artificial, soluções começam a ser pensadas para melhoria de vida deste grupo patológico. Através do *Deep Learning*, da visão computacional e da rede neural convolucional é possível desenvolver um sistema capaz de fornecer melhor qualidade de vida nas atividades diárias destes indivíduos. Desta forma, o presente trabalho objetivou o desenvolvimento de um sistema eletrônico de auxílio aos deficientes visuais através do uso de redes neurais convolucionais. Por meio deste foi possível validar o uso da rede Darknet no reconhecimento de diversos corpos, evidenciando ao deficiente visual aquilo que está à sua volta, propiciando uma melhor interação e inclusão no âmbito social.

Palavras-chave: Deficientes visuais. Tecnologias Assistivas. Inteligência Artificial. Visão Computacional. Redes Neurais Convolucionais.

ABSTRACT

By the pass of time, the blind people have been through accessibility difficulties all over their days. Many activities may generate hazards for this group of person's physical integrity. That problem may be introduced by occasion of factors focused on people's flow or in the others society hazards. Despite of the assistive means, such as guild dog and walking stick, the visually impaired people still passing through locomotion problems, due the low efficiency shown by these previous equipment's. With the technology evolution, specifically the artificial intelligence, solutions were designed to increase the life quality of this group. Using deep learning, computer vision, convolution neural network and electronics, is possible to develop a system capable to provide better life quality in the daily activities as those persons. So, this work seeks to develop an electronic system to help visually impaired people, using convolution neural network. Through it was possible to validate the use of the Darknet network in the recognition of different bodies, showing the visually impaired what is around them, providing a better interaction and inclusion in the social sphere.

Keywords: Visually impaired. Assistive Technologies. Artificial intelligence. Computer vision. Convolutional Neural Networks.

LISTA DE ILUSTRAÇÕES

Figura 1 - Estrutura do globo ocular

Figura 2 - Componentes da túnica externa

Figura 3 - Componentes da túnica média

Figura 4 – Componente da túnica interna

Figura 5 – Ângulos do campo visual

Figura 6 – Método Snellen

Figura 7 – Método Jaeger

Figura 8 – Sinalização de alerta

Figura 9 – Sinalização direcional

Figura 10 – Diferença entre sinal analógico e digital de sensores

Figura 11 – Formação de cores RGB

Figura 12 – Processo de redução de ruído por filtro

Figura 13 – Relação entre Inteligência Artificial, *Machine Learning* e *Deep Learning*

Figura 14 – Diferença de camadas entre *Machine* e *Deep Learning*

Figura 15 - Diferença de extração de atributos entre ML e DL

Figura 16 – Processamento da camada de *pooling*

Figura 17 – Modelo unificado

Figura 18 – Configurações do *Ultrabook* utilizado

Figura 19 – Inclusão das dependências fundamentais

Figura 20 – Variação das cores por classes

Figura 21 – Carregamento das classes pré-treinadas

Figura 22 – Definição da câmera a ser utilizada

Figura 23 – Carregamento da rede e definição dos parâmetros

Figura 24 – Redimensionamento e interpolação do frame capturado

Figura 25 – Cronometragem e detecção do objeto

Figura 26 – Definições para apresentação dos resultados

Figura 27 – Criação e execução da emissão sonora do objeto identificado

Figura 28 – Finalização da detecção e liberação da câmera

Figura 29 – Raspberry® Pi Model 3B+

Figura 30 – *SD Card Formatter*

Figura 31 – Instalando *Raspbian*

Figura 32 – Montagem Modelo Embarcado

Figura 33 – Objetos a serem identificados

Figura 34 – Rede identificando motocicletas distintas e estáticas

Figura 35 – Rede identificando pessoa em ambiente fechado

Figura 36 – Rede identificando aparelho celular em ambiente com baixa luminosidade

LISTA DE TABELAS

Tabela 1 – Deficientes visuais no Brasil por região

Tabela 2 – Classificação das atividades assistivas

Tabela 3 – Componentes básicos do micro controlador

Tabela 4 – Classificação das aplicações da IA

LISTA DE ABREVIATURAS E SIGLAS

ABA - Associação Brasileira de Ações Humanitárias
ABNT - Associação Brasileira de Normas Técnicas
AV - Acuidade Visual
CNI - Confederação Nacional da Indústria
CPU - Unidade de Processamento Central
DL - *Deep Learning*
DV - Deficiente Visual
IA - Inteligência Artificial
IDE - Ambientes de Desenvolvimento
IRIS - Instituto de Responsabilidade e Inclusão Social
GPU - *Graphics Processing Unit*
LBI - Lei Brasileira de Inclusão da Pessoa com Deficiência
LER - Lesão por Esforço Repetitivo
ML – *Machine Learning*
NBR - Norma Brasileira
ONGs - Organizações Não Governamentais
OpenCV - *Open Source Computer Vision*
PDP - Processo de Desenvolvimento de Produtos
PROM - Memória Programável Somente Leitura
RAM - Memória de Acesso Aleatório
RNA - Rede Neural Artificial
RNC - Rede Neural Convolucional
TA - Tecnologias assistivas
VC - Visão Computacional
VSC- *Visual Studio Code*

LISTA DE SÍMBOLOS

m – metro

mm - milímetro

I/O - *Input e Output*

Ohms – Unidade de medida para resistência elétrica

® - Marca registrada

SUMÁRIO

1	INTRODUÇÃO	15
2	OBJETIVOS	17
2.1	Objetivo geral	17
2.2	Objetivos específicos	17
3	REVISÃO DE LITERATURA	18
3.1	Anatomia ocular	18
3.1.1	Túnica externa	19
3.1.2	Túnica média	20
3.1.3	Túnica interna	21
3.2	Fisiologia ocular	22
3.2.1	Campo Visual	22
3.2.2	Acuidade Visual	23
3.3	Patologias oculares	25
3.3.1	Congênitas	26
3.3.2	Adquiridas	26
3.3.3	Baixa visão	26
3.3.4	Cegueira	27
3.4	Deficientes visuais no Brasil	27
3.5	Tecnologias assistivas	29
3.6	Evolução computacional e tecnológica	30
3.6.1	Microcontroladores	32
3.6.2	Sensores	32
3.6.3	Atuadores	33
3.7	Inteligência artificial	34
3.7.1	Visão computacional	35
3.7.2	<i>Deep Learning</i>	37
3.7.3	Rede neural convolucional	40
4	METODOLOGIA	41
4.1	Identificação e definição das necessidades e requisitos	41
4.2	Desenvolvimento do projeto	42
4.3	Ferramentas aplicadas à rede neural	43
4.4	Desenvolvimento do código	44

4.5	Aplicação embarcada no Raspberry® Pi Model 3B+-----	49
5	RESULTADOS E DISCUSSÕES-----	52
6	CONCLUSÃO-----	56
	REFERÊNCIAS-----	57
	APÊNDICE 1 – CÓDIGO COMPLETO-----	67
	APÊNDICE 2 – NOMES DAS CLASSES IDENTIFICÁVEIS-----	71

1 INTRODUÇÃO

Dentro da perspectiva médica e científica, o conceito de deficiência visual em síntese faz referência a impedimentos ligados as patologias oculares, que podem ser congênitas ou adquiridas de acordo com a manifestação do desvio anatômico e/ou fisiológico. Ainda são compreendidos em dois modos de classificação denominados de cegueira e baixa visão, também conhecida como visão subnormal. A base para esta classificação é a acuidade visual, campo visual, sensibilidade ao contraste, percepção de cores e outros aspectos ligados à visão (BATISTA; ENUMO, 2000).

Segundo o Ministério da Saúde (2016), no escopo geral, a cegueira refere-se a total perda da visão de forma que o indivíduo fique incapacitado de realizar leitura através do seu sistema ocular, sendo necessário o uso de Sistema Braille para tal atividade. Já a visão subnormal é caracterizada pela baixa capacidade de funcionamento dos olhos, mesmo que haja tratamento de correção. Neste caso a leitura é possível, no entanto, com auxílio de sistemas artificiais que possibilitam maior acessibilidade aos DVs¹, tais como óculos com lentes especiais, lupas, dentre outros.

Para Guyton e Hall (2011), a visão é um dos sentidos mais importantes para o ser humano e corresponde a um sistema complexo no que diz respeito a sua estrutura fisiológica. O intuito da visão é captar as imagens por intermédio da composição do globo ocular, processar e direcionar sinais elétricos ao nosso cérebro, tornando assim, os olhos análogos a câmeras no que diz respeito ao princípio de funcionamento.

No entanto, nem todos os indivíduos possuem capacidade plena deste sensor devido a questões patológicas e/ou genéticas. Vale salientar que apesar da ineficiência do globo ocular, a capacidade de processamento de informações e de tomada de decisões não se tornam obsoletas. Contudo, acometidos por tais patologias, limitações são impostas na maioria dos casos, o que compromete a independência dos indivíduos (GARCIA, 2014).

Em razão disto, os deficientes visuais encontram em suas rotinas grandes dificuldades de locomoção, atribuídas a diversos fatores intrínsecos ao âmbito

¹ Indivíduos acometido por patologias visuais.

social. Estes fatores, tais como aglomeração de pessoas, obstáculos físicos, trânsito intenso, dentre outros, corroboram para que os indivíduos acometidos por patologias oculares não tenham a acessibilidade capaz de garantir que sua circulação ocorra de forma segura e eficiente. Todavia, existem métodos que auxiliam os DVs nestas atividades (RAHIM, 2017).

Segundo Luiz (2012), uma das formas de gerar maior autonomia e mobilidade aos deficientes visuais é através do cão-guia. No entanto, para que isso seja possível faz-se necessário que o deficiente se enquadre nos requisitos solicitados pelos órgãos brasileiros responsáveis pelo tema em questão. Os requisitos mínimos para esta aquisição é que o deficiente visual tenha mais de 18 anos de idade, resida no território brasileiro e tenha certificado que comprove o curso de orientação e mobilidade.

Outra forma de melhorar a mobilidade do deficiente é com o uso de bengalas específicas para este fim. No entanto, este tipo de acessório é limitado, pois o campo de detecção do obstáculo limitasse apenas aos objetos que se encontram à frente do usuário e numa altura restrita. Deste modo, caso existam obstáculos acima da altura lombar do usuário ou em outros pontos ao seu redor, a segurança do mesmo pode ser comprometida (SOUZA; FERREIRA, 2018).

Neste viés, existe uma problemática que faz relação direta com a necessidade que os deficientes têm de um sistema que seja acessível e sem dependência de cursos, que possibilite o seu uso de forma simples e que tenha eficiência superior às bengalas já comercializadas no mercado. Para tanto, o uso das Redes Neurais Convolucionais (RNC) é de grande valia, pois contempla a real necessidade dos deficientes visuais, podendo realizar o trabalho de visão computacional e processamento de imagens de tal forma a evidenciar ao indivíduo o que está a sua volta, possibilitando que o mesmo tome as decisões adequadas a cada situação no âmbito social cotidiano (EBERMAM; KROHLING, 2018).

Neste contexto, este trabalho objetiva o desenvolvimento de um sistema inteligente de auxílio à acessibilidade de deficientes visuais com o uso da rede neural convolucional, que possibilite a obtenção de imagens e realize o processamento capaz de evidenciar ao usuário, por meio de sinais sonoros, o que o cerca, viabilizando a tomada de decisão por parte do mesmo.

2 OBJETIVOS

2.1 Objetivo geral

Desenvolver um sistema inteligente de auxílio a acessibilidade de deficientes visuais com o uso de rede neural convolucional, em especial com a rede Darknet.

2.2 Objetivos específicos

- Identificar as necessidades dos deficientes visuais perante situações cotidianas;
- Desenvolver código capaz de emitir sinais sonoros referindo-se a diversificados objetos e corpos identificados;
- Executar com eficiência a rede neural convolucional, sem comprometer os custos do projeto;
- Embarcar rede no Raspberry® Pi Model 3B+.

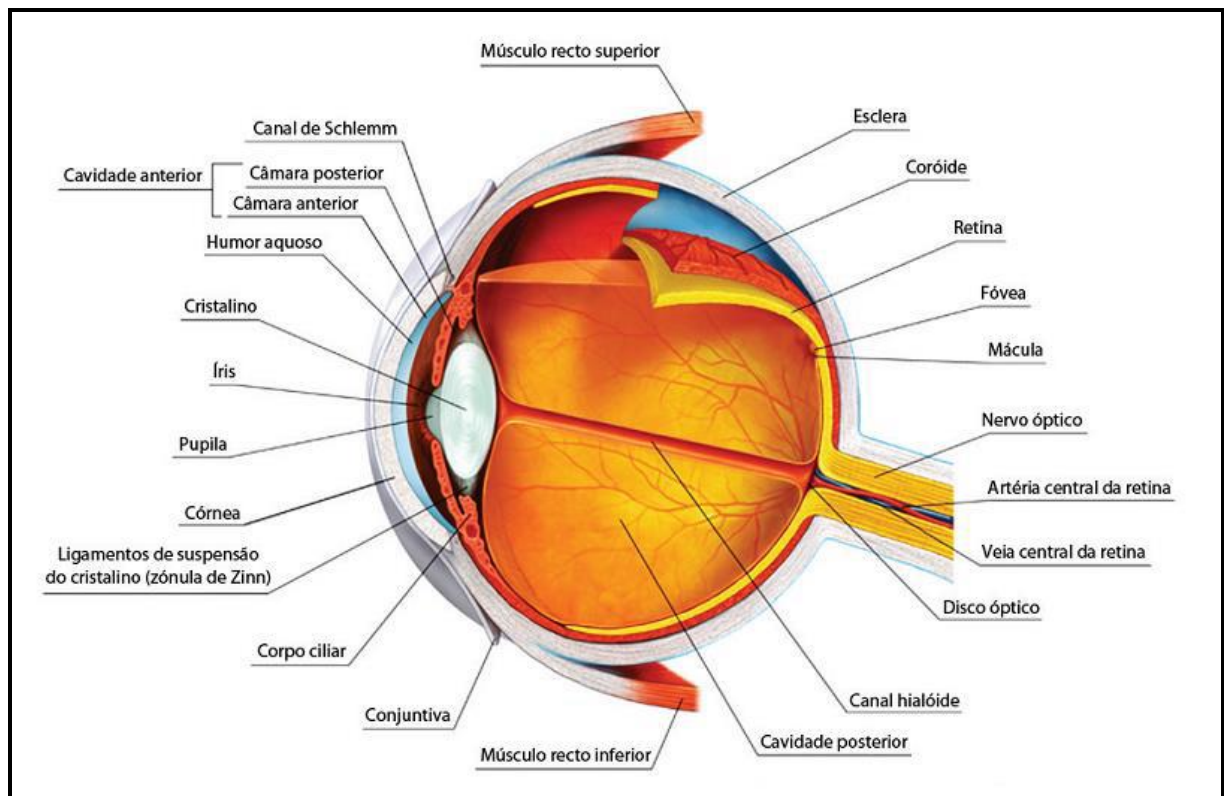
3 REVISÃO DE LITERATURA

3.1 Anatomia ocular

O olho humano é caracterizado por sua complexidade anatômica e fisiológica, sendo compreendido como um sistema sensorial. Este sistema, denominado de globo ocular é responsável por captar energia luminosa e converter em impulsos nervosos que são guiados até o cérebro, onde ocorre o processo de geração de imagens (PASSOS; NETO; LEMAIRE, 2008).

Constituído por diversos componentes que propiciam o seu funcionamento, o globo ocular é compreendido em três camadas, também denominadas de túnica. A primeira é a camada externa, constituída pela esclera e pela córnea. A segunda é denominada como camada média e formada pela coróide, corpo ciliar e íris. Já a terceira túnica corresponde a camada interna, onde se localiza a retina. A estrutura do globo ocular contempla os principais componentes (**Figura 1**) deste sistema (BAIÃO, 2013).

Figura 1 - Estrutura do globo ocular



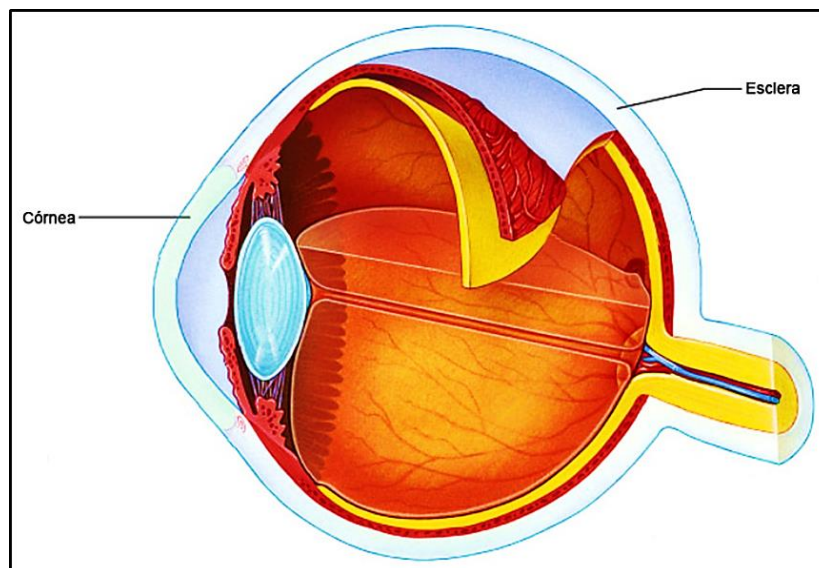
Fonte: BAIÃO (2013)

3.1.1 Túnica externa

A túnica externa é constituída, pela esclera e pela córnea (**Figura 2**) e, segundo Netto (2010), a esclera corresponde a camada fibrosa do bulbo ocular e desempenha algumas funções importantes que garantem o formato e proteção de aproximadamente 85% do bulbo. Além disto, é nela que estão fixados os músculos extraoculares e intraoculares. Desta forma, é um componente caracterizado por sua resistência e opacidade. Ainda, é revestido pela conjuntiva bulbar transparente, o que possibilita a sua visualização como sendo “a parte branca do olho”.

Tratando-se da córnea, a mesma é compreendida como uma calota transparente e convexa que permite a entrada dos raios luminosos no globo ocular através da pupila. Deste modo, é possível a geração das imagens por intermédio dos demais componentes em convergência com o processamento cerebral. Para tanto, a córnea é disposta no segmento anterior do olho e corresponde a cerca de 15% da dimensão do bulbo (NETTO, 2010).

Figura 2 - Componentes da túnica externa



Fonte: GOMES (2019)

Segundo (NISHIDA, 2012), as diferenças entre a esclera e a córnea, além das funções que as mesmas desempenham no sistema de visão, são as diferenças de composição. Dentre estas divergências, explícitas nos limítrofes² entre ambas estão à organização das fibras colágenas constituintes e o grau de hidratação de

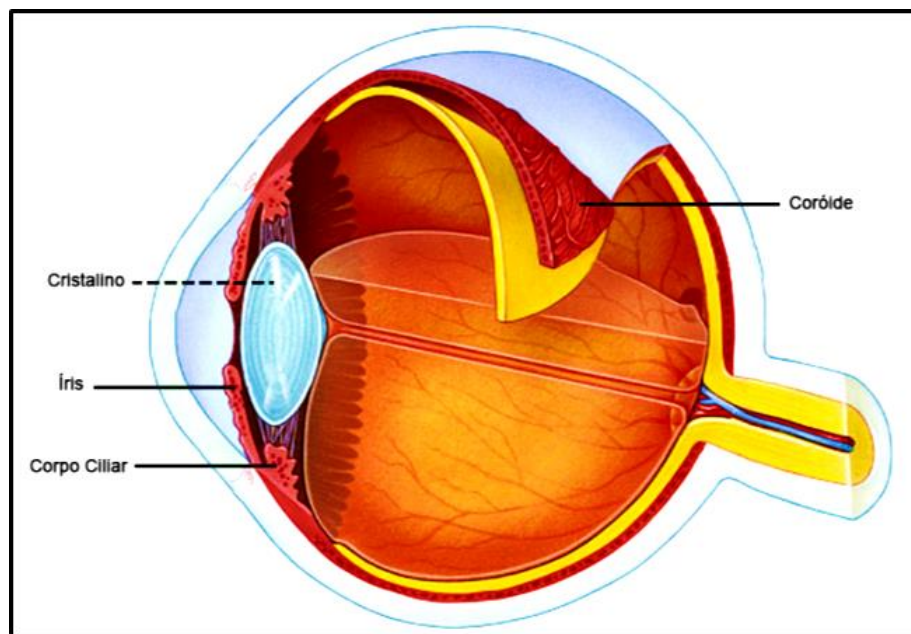
² Que se situa nos limites de uma extensão, que tem limites comuns.

cada uma. Estas diferenças possibilitam que a incidência luminosa tenha apenas um ponto de entrada no globo ocular, garantindo a eficiência na obtenção das imagens.

3.1.2 Túnica média

A camada média é composta, basicamente, pela coróide, corpo ciliar e íris (**Figura 3**). A coróide trata-se de uma camada intermediária do globo ocular demasiadamente vascularizada e sua principal função é absorver excesso luminoso que adentra o sistema sensorial da visão. Com isso, as reflexões sobre os fotorreceptores da retina são minimizadas. No entanto, para que esta função seja cumprida, a coróide é envolvida por uma camada epitelial pigmentada em tom de marrom-avermelhado escuro (NISHIDA, 2012 e NETTO, 2010).

Figura 3 – Componentes da túnica média



Fonte: GOMES (2019)

O corpo ciliar trata-se de um tecido muscular vascularizado e desempenha algumas funções importantes no processo de visão. Estas funções correspondem a produção e manutenção das zônulas (responsáveis por manter o cristalino suspenso), a produção do humor aquoso que preenche a câmara anterior e posterior do globo ocular e a realização da acomodação. Neste processo, ocorre a contração e relaxamento do músculo liso presente no corpo ciliar. Com esta dinâmica, a curvatura do cristalino é modificada, o que possibilita o foco do cristalino em distâncias variáveis (RODRIGUES, 2016).

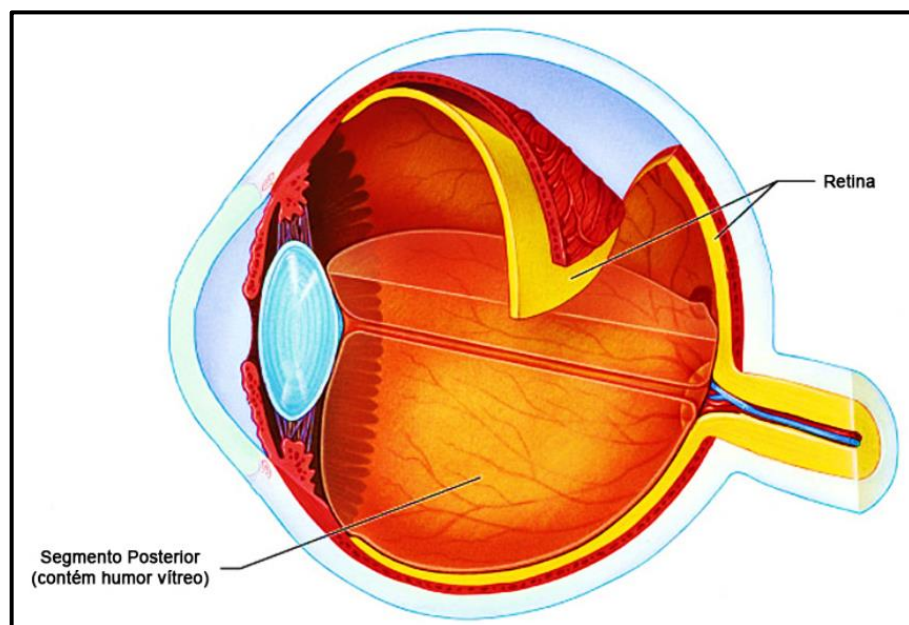
Já a íris, é uma composição fibrosa muscular pigmentada localizada sobre a superfície anterior do cristalino e determina a coloração externa do olho. Ainda, tem a função de regular a incidência luminosa na pupila através da dinâmica de abertura ou fechamento de sua estrutura (RODRIGUES, 2016).

3.1.3 Túnica interna

A túnica interna é basicamente constituída pela retina (**Figura 4**). Segundo Netto (2010), a retina é compreendida como extensão do sistema nervoso central, localizada no globo ocular. A sua composição é compreendida por duas estruturas funcionais situadas em locais distintos do bulbo. Estas partes são diferenciadas pela função visual e denominadas como estrutura óptica e estrutura cega.

Para Rodrigues (2016), a estrutura óptica é sensível aos raios luminosos, enquanto a parte cega não responde a este tipo de estímulo, pois não contém fotorreceptores em sua estrutura. Desta forma, a retina desempenha a função de converter raios luminosos em impulsos nervosos e conduzi-los ao cérebro por intermédio do nervo óptico.

Figura 4 - Componente da túnica interna



Fonte: GOMES (2019)

Ainda visto na Figura 4, o segmento posterior do bulbo ocular é preenchido por um líquido gelatinoso denominado de humor vítreo. Este líquido garante o

formato arredondado do globo e é composto por, aproximadamente, 99% de água e 1% de colágeno e ácido hialurônico (BAIÃO, 2013).

3.2 Fisiologia ocular

Para Farinha (2011), o mecanismo de visão é um dos mais importantes para o ser humano. Isso se dá pela capacidade de transferência de informações do meio para com o indivíduo. Para que isto ocorra, a experiência da visão passa por três fundamentais etapas, a física, a fisiológica e a psicológica. Em convergência, estas etapas tornam a vida humana dinâmica em seus aspectos sociais. Mas, para que isto ocorra, fisiologicamente, a visão necessita de compatibilidade com distintas incidências luminosas.

Portanto, o olho é composto por dois tipos de fotorreceptores, os bastonetes e os cones. Os bastonetes são utilizados na visão noturna, onde a incidência luminosa é baixa e por isso gera imagens em tons de preto e branco. Já os cones, tem responsabilidade de gerar imagens coloridas sob incidência luminosa propícia. Por conta da alta luminosidade, estas imagens apresentam maior nitidez se comparadas com as oriundas dos bastonetes (FARINHA, 2011).

Após o exercício da função dos cones e dos bastonetes, os sinais são propagados em camadas de neurônios ainda na retina. Posteriormente, são transmitidos para as fibras presentes no nervo óptico e para o córtex visual. No córtex ocorrem os processos de percepção de alguns parâmetros de visão, tais como contraste, orientação de linhas e bordas, dentre outros (HALL; GUYTON, 2011).

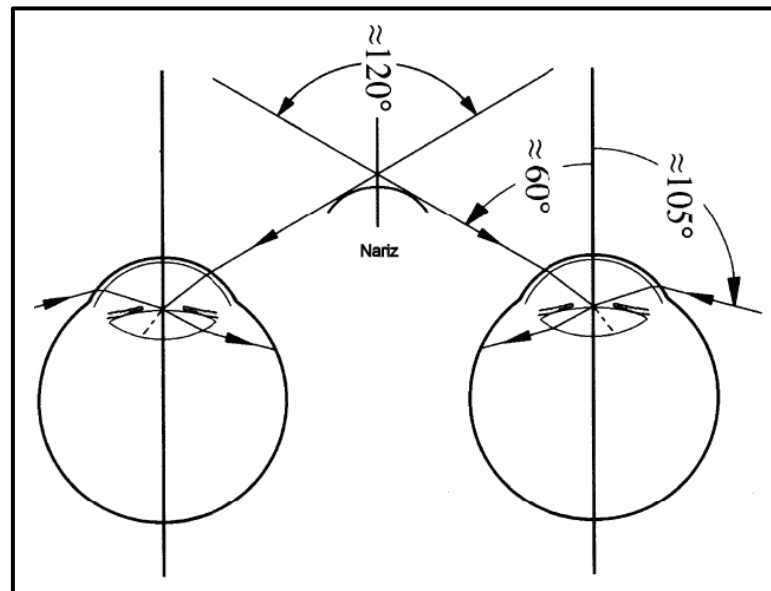
Além dos conceitos fundamentais apresentados anteriormente, é de suma importância a compreensão de outros dois, pois também impactam diretamente no entendimento da interação entre o homem e o meio, por intermédio da visão. Logo, estes serão apresentados nos tópicos 6.2.1 e 6.2.2, descritos como campo visual e acuidade visual, respectivamente.

3.2.1 Campo visual

Para Passos, Neto e Lemaire (2008), a área correspondente à obtenção das imagens no processo de visão é concebida como campo visual. Em síntese, tudo o que é possível ser visualizado em frente, lateral direito e esquerdo, acima ou abaixo do olho humano quando fixado num ponto específico, está inserido no campo visual.

Este campo é compreendido pela visão binocular, que capta as imagens em ângulos distintos e consolida em uma única imagem. Estes ângulos são definidos e seguem um padrão perante o sistema ideal de visão **(Figura 5)**.

Figura 5 – Ângulos do campo visual



Fonte: PASSO et al (2008)

O campo visual pode ser segregado em dois campos, denominados de campo visual nasal e campo visual lateral. Como subentendido pela nomenclatura, o campo visual nasal trata-se de toda a área de captação posicionada frontalmente ao nariz. Enquanto o campo visual lateral corresponde a captação de área excedente do campo visual nasal (HALL; GUYTON, 2011).

Com a determinação dos dois campos, as imagens que são obtidas por cada globo ocular (60°) são sobrepostas na retina obedecendo ao ângulo de 120° no cruzamento. Esta sobreposição é responsável pela geração tridimensional de tudo o que é captado pelo olho humano, possibilitando a determinação de profundidade e relevo das imagens (NISHIDA, 2012).

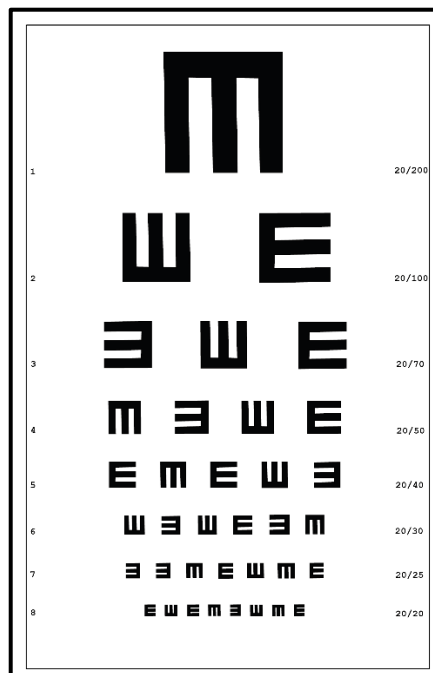
3.2.2 Acuidade visual

Em síntese, a acuidade visual (AV) corresponde a um parâmetro da visão, caracterizado pela capacidade de percepção de contornos e formas por parte da composição ocular. Um dos principais fatores para a boa acuidade visual é determinado pelos detalhes captados pela retina. Em nível fisiológico, a AV tem ligação direta com a mácula, pois é esta parte da retina responsável pelos detalhes

a serem captados nas imagens. Logo, para que os objetos sejam distinguidos com nitidez e definição de bordas, é necessário que o foco de luz incida com perfeição sobre a retina e possibilite o exercício pleno da função da mácula (INSS, 2014).

Ainda segundo o INSS (2014), no Brasil assim como em todo o mundo, o método mais utilizado para medir o nível de AV para longe é através do método Snellen (**Figura 6**). Este método determina a capacidade de reconhecimento oriunda do processo de visão e consiste na identificação, por parte do paciente, do posicionamento do símbolo, geralmente análogo a letra “E”. Esta avaliação é realizada nos olhos separadamente e na visão binocular. Inicialmente, o indivíduo é posicionado em longa distância e define o posicionamento sem correção por meio de lentes e, posteriormente, com a aplicação das lentes corretivas, de forma a obter os valores referentes a sua acuidade visual.

Figura 6 – Método Snellen



Fonte: Adaptado de CÔRREA *et al.* (2016)

Já no caso do método Jaeger trata-se de um procedimento utilizado para determinar a acuidade visual para perto. Neste exame, o paciente é posicionado próximo a tabela (**Figura 7**) e busca identificar os escritos numéricos ou simbólicos da primeira para a última linha. Neste caso, assim como no Snellen, o procedimento é aplicado em visão monocular e binocular, com e sem uso de correção por lentes. Com ambos os métodos, é possível identificar os valores obtidos nos exames e caso

este seja menor que 20/20, é possível que o paciente esteja acometido por alguma ametropia, tais como miopia, astigmatismo e hipermetropia (SALOMÃO, 2007).

Figura 7 – Método Jaeger

0,37m	11111111	11111111	J1
0,50m	11111111	11111111	J2
0,67m	11111111	11111111	J3
0,75m	11111111	11111111	J4
1,00m	11111111	11111111	J5
1,25m	11111111	11111111	J6

Fonte: Adaptado de COSTA et al. (2018)

3.3 Patologias oculares

As patologias oculares se apresentam em decorrência de problemas que afetam a visão dos indivíduos. Fisiologicamente, estes problemas são caracterizados pela degeneração do globo ocular, do nervo óptico ou até mesmo das ligações nervosas entre o olho e o cérebro. Contudo, não são apenas estas as causas evidentes das patologias, pois elas são multifatoriais, podendo apresentar causas não especificadas. Desta forma, as patologias são classificadas em dois modos, a depender de como é manifestada, sendo conferidas como congênitas ou adquiridas (GARCIA, 2014).

Ainda, é prevista a classificação patológica ocular quanto a seu nível de impacto oftalmológico, que pode ser enquadrado como baixa visão ou cegueira. Como forma de medição para definição de tal nível, são utilizados os parâmetros de acuidade e campo visual já explicados nos tópicos 6.2.1 e 6.2.2. Contudo, a avaliação para cada parâmetro difere em suas aplicações metodológicas (CBO, 2019).

3.3.1 Congênitas

Existem algumas divergências temporais na classificação das patologias congênitas. Segundo Garcia (2014), alguns autores convergem para a ideia de que a patologia é congênita quando manifestada durante ou logo após o nascimento. Já outros autores, entendem que este período deve ser considerado até os 12 (doze) meses de idade. Ainda, existem autores que defendem a tese de que se manifestada até os 5 (cinco) anos de idade, tais patologias também devem ser compreendidas como congênitas.

Pelo pressuposto de que é manifestado no nascimento, o INSS (2014), evidencia algumas doenças que podem ser a causa deste tipo de patologia, tais como cataratas, glaucomas, tumores oculares, estrabismos, uveítes, malformação do canal lacrimal, nistagmos e ametropias.

3.3.2 Adquiridas

Como exposto por Santos e Nascimento (2015), as patologias adquiridas correspondem aquelas que são manifestadas ao longo do tempo. Nestes casos, a visão do indivíduo tem pleno funcionamento no início de sua vida e em algum momento ele a perde. Nesta classificação, preexistem as memórias visuais que foram obtidas até o momento da perda da visão. Segundo o INSS (2014), alguns fatores que podem levar a este tipo de patologia estão os traumas, corpos estranhos, infecções, doenças sistêmicas, neurológicas e alérgicas.

Um ponto a ressaltar é que a preexistência das memórias visuais pode ser favorável ou desfavorável em casos de readaptação do indivíduo. Segundo Santos E. e Nascimento (2015), estas lembranças são úteis neste processo. Já para Felicetti *et al.* (2016), nestes casos a readaptação torna-se um processo complexo e pode evidenciar alguns traumas e dificuldades nas relações sociais, sendo necessário o acompanhamento psicológico em paralelo.

3.3.3 Baixa visão

Segundo o Ministério da Educação, (2000), também denominada de visão subnormal, a baixa visão é caracterizada pela diminuição das funções oculares por razão do rebaixamento demasiado da acuidade visual, do campo visual,

sensibilidade ao contraste, dentre outros parâmetros. Em síntese, pode ser conceituada como a incapacidade de visão nítida em locais com boa iluminação.

Pelo Ministério da Saúde (2016), um dos fatores primordiais na diferenciação da visão subnormal em relação à cegueira, é que a baixa visão permite tratamento e correção por meio de métodos usuais, como o uso de lupas, lentes, óculos ou tecnologias assistivas voltadas a este fim. Com isso, os indivíduos podem ler textos impressos, o que é algo com alto grau de dificuldade para os mesmos.

3.3.4 Cegueira

Mediante conceitos apresentados pelo INSS (2014), a amaurose, como também é conhecida a cegueira, está relacionado ao indivíduo que não dispõe da percepção luminosa ou conta com a perda total da visão. Para Gambarato (2012), a cegueira é considerada como uma deficiência severa, mas pode ser minimizada através de formas específicas de aprendizado e com o uso de tecnologias assistivas.

Para os deficientes visuais, é necessário que a acessibilidade seja adequada e propicie as suas atividades cotidianas e sociais, tanto no âmbito físico, quanto no meio digital. Deste modo, as dificuldades oriundas da patologia podem ser minimizadas, trazendo maior conforto e qualidade de vida a essas pessoas. A partir disto, os mesmos são incluídos e se sentem úteis na sociedade em que estão inseridos (GAMBARATO; BATISTA; GIANDONI, 2012).

3.4 Deficientes visuais no Brasil

No cotidiano, os deficientes visuais (DV) encontram diversas dificuldades de locomoção, pois os meios em que os mesmos estão inseridos apresentam diversos obstáculos. Estes são apresentados de diversas formas, podendo ser localizados no solo ou entorno do deficiente visual, em diversos níveis de altura. Muitos destes obstáculos, não são passíveis de identificação pela bengala devido as suas dimensões e localização no meio. Desta forma a acessibilidade dos indivíduos acometidos por tal patologia fica comprometida. (RAHIM, 2017)

No entanto, para que o acesso à acessibilidade seja difundido e regulamentado, o tema é tratado no Brasil mediante redação dada pela Lei nº 13.146, de 2015. Segundo Silva C., Gomes e Júnior (2017), esta lei foi sancionada

em 2015 (dois mil e quinze) e recebeu o nome de Lei Brasileira de Inclusão da Pessoa com Deficiência (LBI). Nela a acessibilidade é conceituada como:

Possibilidade e condição de alcance para utilização, com segurança e autonomia, de espaços, mobiliários, equipamentos urbanos, edificações, transportes, informação e comunicação, inclusive seus sistemas e tecnologias, bem como de outros serviços e instalações abertos ao público, de uso público ou privados de uso coletivo, tanto na zona urbana como na rural, por pessoa com deficiência ou com mobilidade reduzida. (Diário Oficial da República Federativa do Brasil, 2015 p. 20).

Ainda pela Lei nº 13.146, de 2015, divulgada no Diário Oficial da República Federativa do Brasil (2015), em seu Art. 4º, todas as pessoas com deficiência têm direito a igualdade no que tange as oportunidades e não sofrerá nenhuma espécie de discriminação no meio social. Vale ressaltar que a LBI conta com 127 (cento e vinte e sete) artigos. Nestes artigos, é garantido a estes indivíduos o atendimento prioritário, os direitos fundamentais (vida, habilitação e reabilitação, saúde), direito ao trabalho, moradia, cultura, esporte, lazer, assistência social, acessibilidade, mobilidade, informação, comunicação, tecnologia assistiva, justiça, dentre outros.

Para que estas leis sejam aplicadas em todo o território nacional igualmente, o número de deficientes visuais é monitorado pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Nos dados divulgados em 2013 (dois mil e treze), o IBGE aponta o índice destes deficientes que residem no país e a contração destes indivíduos em suas respectivas regiões (**Tabela 1**).

Tabela 1 – Deficientes visuais no Brasil por região

REGIÃO	NÚMERO DE DV
<i>Norte</i>	509.963
<i>Nordeste</i>	1.877.607
<i>Centro-Oeste</i>	491.571
<i>Sudeste</i>	2.705.636
<i>Sul</i>	1.698.061

Fonte: Adaptado de IBGE (2013)

A principal barreira para os deficientes visuais no Brasil trata-se da falta de acessibilidade. Vale ressaltar que as barreiras se referem a tudo aquilo que impossibilita ou limita o livre direito de ir e vir de qualquer indivíduo. Este direito é

garantido pela constituição da República Federativa do Brasil, mas nem sempre seguida em todo o território nacional. A questão a ser evidenciada é que a problemática para os deficientes visuais não é exclusiva das limitações impostas pela patologia, mas também pelas impostas pela sociedade (GUEDES; GIRÃO, 2017).

Ainda segundo Guedes e Girão (2017), na tratativa social, além das implicações atitudinais preconceituosas, o espaço físico e seu planejamento inadequado é outra forma de explicitar as necessidades destas pessoas. Para que isto seja revertido, é necessário que haja a reanálise das condições em que os deficientes visuais estão inseridos no âmbito social Brasileiro. Desta forma, o empenho para propagar a atenção e desenvolvimento de métodos eficientes quanto a acessibilidade deve ocorrer incansavelmente, a fim de propiciar uma sociedade cada vez mais democrática.

3.5 Tecnologias assistivas

As tecnologias assistivas (TA), por análise de Costa R. (2016), englobam todos os métodos ou recursos que favoreçam a acessibilidade dos deficientes, propiciando independência e autonomia perante a sociedade. Estas tecnologias têm o intuito de ampliar a inclusão social dos indivíduos e permitir mobilidade nos mais distintos meios sociais. Vale ressaltar que estas tecnologias podem ser aplicadas por intermédio de sistemas mecânicos, elétricos, computadorizados, dentre outros.

A fim de propiciar a organização dos estudos, pesquisas e desenvolvimentos na área das TAS, a mesma foi dividida em classificações quanto as suas funcionalidades no cotidiano dos deficientes. Esta classificação (**Tabela 2**) pode variar em dependência dos autores, mas serve como direcionamento para os pesquisadores em seus estudos (ZAQUEU, 2012).

Tabela 2 – Classificação das atividades assistivas

CLASSIFICAÇÃO	DESCRIÇÃO
Auxílios para a vida diária	Materiais e produtos para auxílio em tarefas: comer, vestir, etc.
Recursos para computador	Síntese de voz, Braille, auxílios alternativos de acesso, teclados modificados ou alternativos, etc.
Sistemas de controle	Sistemas permitem as pessoas com limitações moto-locomotoras controlar remotamente aparelhos.

Projetos arquitetônicos para acessibilidade	Adaptações estruturais e reformas na casa e/ou ambiente de trabalho, elevadores, dentre outras.
Órteses e próteses	Troca ou ajuste de partes do corpo, faltantes ou de funcionamento comprometido, por membros artificiais.
Auxílios para pessoas cegas ou com visão subnormal	Auxílios para grupos específicos que inclui lupas e lentes, grandes telas de impressão, publicações, etc.
Auxílios para surdos	Aparelhos para surdez, sistemas com alerta tátil-visual, etc..
Adaptações em veículos	Acessórios e adaptações para condução do veículo, elevadores para cadeiras de rodas, etc.

Fonte: Adaptado de ZAQUEU (2012)

3.6 Evolução computacional e tecnológica

A computação tem longevidade histórica em concordância com a evolução humana, pois prende-se ao suprimento das necessidades oriundas de tal evolução. Apesar da computação ter seus métodos definidos por volta de 1930, os princípios são conhecidos a milhares de anos. Tendo em vista que a principal motivação para o surgimento deste tipo de tecnologia foi a necessidade de realização de cálculos matemáticos, o ábaco foi o primeiro sistema mecânico que faz analogia aos computadores atuais. Mas, houveram outros sistemas mecânicos que foram essenciais para os moldes atuais da computação, sendo compreendidos como geração zero da computação (FERNANDEZ; CORTÉS, 2015).

No entanto, por Fernandez e Cortés (2015), a computação aproximasse dos métodos atuais por volta de 1940. Para melhor compreensão dos marcos de evolução, a computação é classificada em gerações. A primeira geração dos computadores é caracterizada pelo início da utilização de circuitos eletrônico, tendo como base os relés e válvulas. Em síntese, os relés correspondem a eletroímãs, cuja função é interromper ou propiciar os contatos elétricos em circuitos. Já as válvulas, são componentes condutores de corrente elétrica em um único sentido previamente-definido.

Em 1948, surge o transistor e, por volta de 1960, este componente passa a ser utilizado na computação, caracterizando a segunda geração. Ainda, foi neste período que houve a substituição dos tubos de raios catódicos por núcleos magnéticos. Com o surgimento dessas tecnologias, os computadores passam a ser

menores em suas dimensões, porém com maior capacidade de processamento em relação aos anteriores (VIEIRA, 2018).

Por Vieira (2018), a terceira geração teve início em 1965, e ficou marcada pela inserção do circuito integrado e da microeletrônica. Com isto, os componentes passam a serem integrados em chips, culminando ainda mais na redução dimensional dos computadores. Sendo assim, o uso destes aparelhos começa a ser disseminado em diversos âmbitos, sendo ainda mais velozes no processamento e nas operações básicas para que os mesmos eram projetados.

A quarta geração tem início em 1971, tendo como marco principal o surgimento do microprocessador, com foco na redução dimensional dos aparelhos. Além desta evolução, começam a surgir às linguagens de programação de alto nível. Estas correspondem às linguagens mais próximas da compreensão humana, minimizando a necessidade por meio da programação binária. Outro fator importante foi o surgimento da teleinformática, ou seja, a transmissão de dados por rede entre os computadores, conhecida como intranet (COUTINHO, 2008).

A partir de 1981, segundo Coutinho (2008), os computadores atingem alto nível de processamento e interatividade entre máquina e usuário. Por meio dos chips de silício, os computadores passam a ter dimensões e custos ainda menores. A partir deste momento, surgem os minicomputadores, a fim de propiciar o conforto e atender as necessidades dos usuários. No mesmo período surge o conceito de inteligência artificial e o mesmo começa a ser estudado e desenvolvido por estudiosos da área. Além da evolução computacional, o comportamento dos usuários perante as máquinas se modifica, possibilitando a exploração mercadológica cada vez mais evidente. É neste momento que a tecnologia computacional passa a ser explorada efetivamente em resposta a demanda que se apresenta no mercado global.

Vale ressaltar que a tecnologia em sua diversidade, propicia o desenvolvimento de distintas soluções voltadas as problemáticas intrínsecas ao cotidiano do ser humano. Tais soluções podem ocorrer por meio de *softwares* e/ou *hardwares*, a depender da aplicabilidade desejada pelo desenvolvedor e da relação com o problema em questão. Dado o advento da evolução tecnológica, o mercado tem buscado gerar soluções de *hardwares* cada vez mais viáveis e tangíveis para a sociedade global. Isto se apresenta através do surgimento dos microcontroladores e seus periféricos de entrada (sensores) e saída de dados (atuadores) (RAHIM, 2017).

3.6.1 Microcontroladores

Os microcontroladores são, em síntese, circuitos integrados que necessitam apenas de uma fonte de alimentação externa para funcionamento, pois o mesmo é provido de todos os componentes, podendo ser equiparados a computadores inseridos em um único e pequeno chip. Por se tratar de um equipamento de baixo custo e de alta versatilidade, o mesmo pode ser empregado nos mais distintos meios. Para seu funcionamento, é necessário que seja realizada uma programação prévia, que coordenará as funções dos componentes internos (KERSCHBAUMER, 2018).

Cada componente desempenha uma função (Tabela 3) na arquitetura proposta. Dentre estes temos a Unidade de Processamento Central (CPU), Memória Programável Somente Leitura (PROM), Memória de Acesso Aleatório (RAM), Linhas de *input* e *output* (I/O) (PENIDO; TRINDADE, 2013).

Tabela 3 – Componentes básicos do microcontrolador

COMPONENTE	FUNÇÃO
CPU	Interpreta as instruções programadas
PROM	Memória onde é gravada as instruções
RAM	Memorizar as variáveis dos programas
Linhas I/O	Recebe informações dos sensores e envia dados para atuadores

Fonte: Adaptado de KERSCHBAUMER (2018)

Segundo Kerschbaumer (2018), os microcontroladores são aplicados em quase todos os equipamentos eletrônicos, tais como, alarmes, computadores ar-condicionado, relógios, smartphones, carros, dentre outros. No entanto, para que sejam executadas as ações programadas, o mesmo necessita de dados de entrada, obtidos pelos sensores.

3.6.2 Sensores

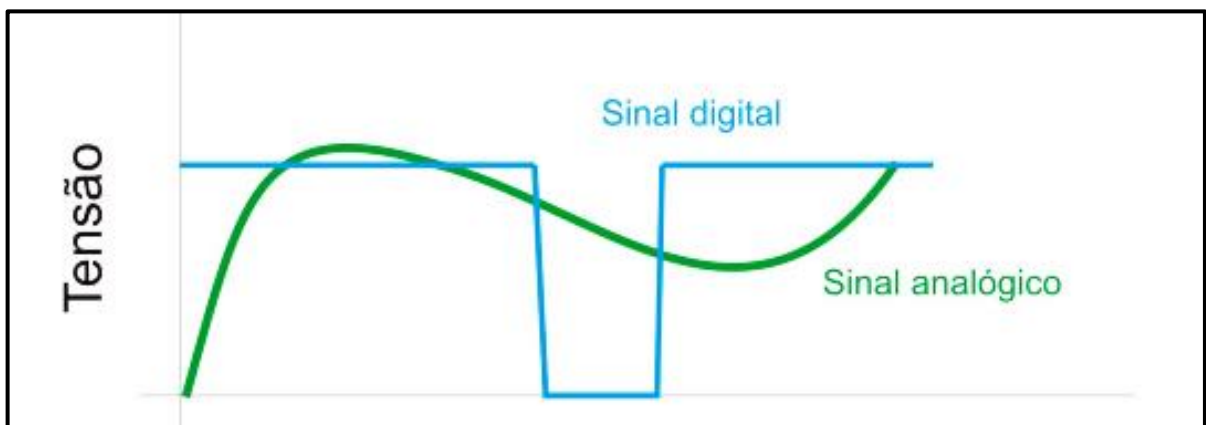
Segundo Patsko (2006), os sensores impactam positivamente no cotidiano do ser humano moderno, pois o mesmo possibilita o conforto, segurança e comodidade em diversas situações. Como exemplo, o carro e suas evoluções tecnológicas,

contemplam o uso destes componentes. O objetivo dos sensores é agir como um transdutor, ou seja, captar um tipo de energia e converter em outra.

Os sensores são dispositivos sensíveis ao meio, com capacidade de captação de energia, podendo ser luminosa, térmica, cinética, etc. Através das variações físicas pertinentes, é possível atribuir ações condicionadas aos valores obtidos no processamento de entrada analógica de determinadas atividades. Para que isto ocorra, os sensores apresentam características elétricas propícias para tais atividades e são classificados como analógicos ou digitais (WENDLING, 2010).

De acordo com Wendling (2010), os sensores analógicos assumem valores dentro de uma faixa específica, podendo variar ao longo do tempo. Neste caso, os componentes atribuídos captam a variação, com isso é possível obter dados definidos conforme as grandezas. Já os sensores digitais, assumem apenas dois valores, tratados na computação como binário (zero ou um). Neste caso, é possível captar o estado de um determinado corpo no espaço, onde é possível verificar a diferença de sinal elétrico gerado entre o sensor analógico e digital (**Figura 10**).

Figura 10 – Diferença entre sinal analógico e digital de sensores



Fonte: PATSKO (2006)

3.6.3 Atuadores

Os atuadores correspondem aos periféricos que recebem sinais elétricos provenientes de um controlador e atua na realização de atividades previamente programadas. Existem diversos tipos de atuadores, por exemplo, válvulas pneumáticas e hidráulicas, pistões, motores elétricos, solenoides, emissores de áudio, dentre outros. Com estes, é possível desenvolver sistemas capazes de

atender as mais variadas necessidades humanas, podendo ter uso específico em máquinas, veículos e sistemas mecânicos (WENDLING, 2010).

Em resumo, os atuadores convertem estímulos elétricos em energia mecânica, térmica, luminosa, calorífica, etc. Para tanto a classificação básica é determinada pela fonte de energia utilizada para que o atuador intervenha no sistema, podendo ser hidráulico, pneumático ou elétrico. O hidráulico utiliza a energia de um líquido sob pressão, enquanto o pneumático faz uso da energia proveniente do ar comprimido proveniente da compressão por meios dinâmicos de compressores. No caso dos elétricos, a energia é proveniente de meios elétricos, como baterias ou fontes de alimentação, gerados por sistemas específicos (STEBEL; SATO, 2017).

3.7 Inteligência artificial

Em tratativas etimológicas, os termos que norteiam o tema podem ser descritos separadamente, a fim de melhorar a compreensão dos conceitos oriundos deste tipo de tecnologia. A palavra inteligência vem da composição latina de *inter* e *legere*. O termo “*inter*” significa, no português, “entre” e “*legere*” condiz com o verbo “escolher”. Desta forma, é compreendido que a palavra inteligência conceitua-se como a capacidade de escolha em meio a possibilidades apresentadas. Já o termo artificial, também derivado do latim, *artificiale*, faz menção àquilo que não é natural, ou seja, que necessitou do esforço humano para ser produzido (NAKABAYASHI, 2009).

Ainda por Nakabayashi (2009), partindo do pressuposto etimológico, pode-se resumir a inteligência artificial como o poder de escolha de uma máquina em simulação da capacidade humana na tomada de decisões, sendo este construído pelo próprio ser humano. Ou seja, o homem criando meios de reproduzir a sua inteligência em máquinas.

Segundo Vicari (2018), o termo Inteligência Artificial (IA), foi concebido por J. McCarthy, um dos estudiosos que conceberam este ramo científico. A concepção do termo é datada em 1956 e seus objetivos eram descritos pela multidisciplinaridade, sendo aplicada em áreas da filosofia, psicologia, matemática, ciência da computação, dentre outros. Sendo assim, as aplicações da IA pode ser compreendida em 3 (três) formas de aplicação (**Tabela 4**).

Tabela 4 – Classificação das aplicações da IA

ATIVIDADES	EXEMPLOS
Corriqueiras	Percepção (fala, visão) Língua Natural (entendimento, geração, tradução) Raciocínio de Senso Comum (Controle de Robôs)
Formais	Jogos (xadrez) Matemática (geometria, lógica, cálculo integral)
Especialistas	Engenharia (projetos, descoberta de falhas, planejamento de manufatura) Análise Científica Diagnóstico Médico Análise Financeira

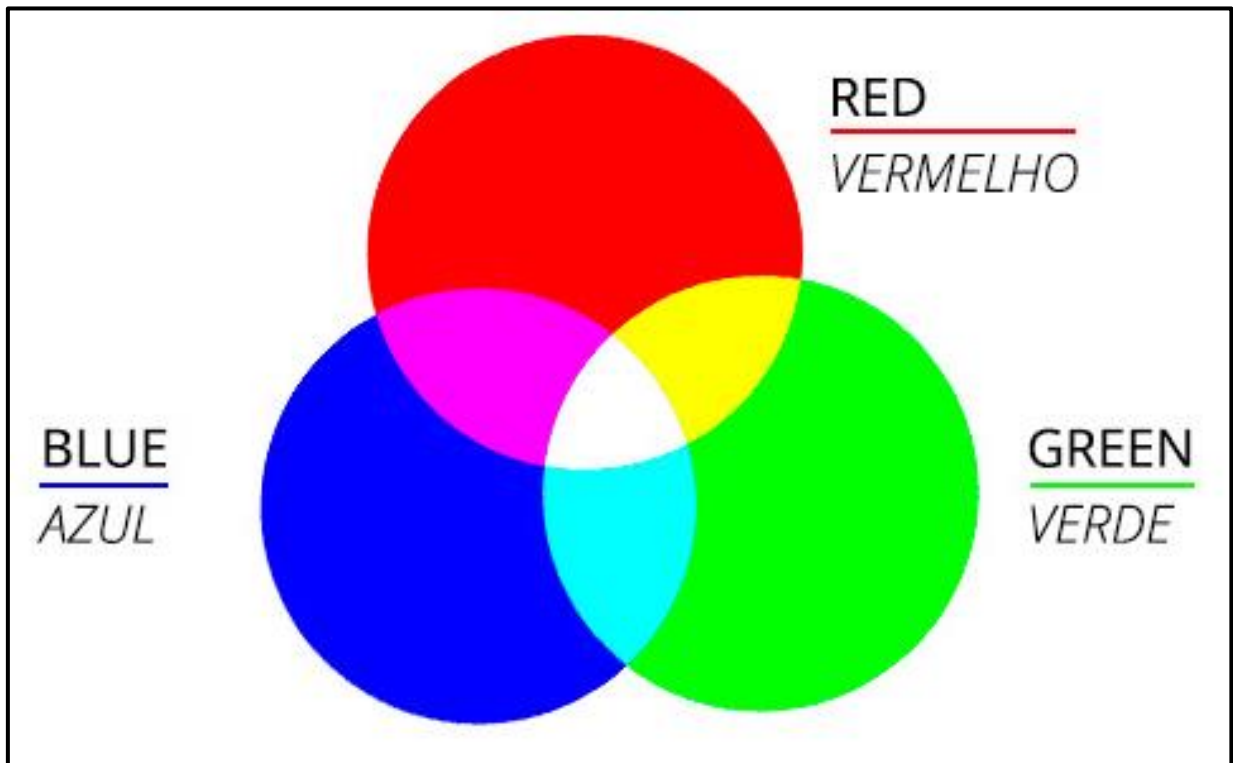
Fonte: Adaptado de ROSA (2011)

3.7.1 Visão computacional

O processo de visão computacional (VC) consiste na obtenção de imagens digitais por meio de câmeras que são direcionadas para a realização de tratamentos, também denominados de pré-processamento de imagens. Logo após, é iniciada a execução da rede neural, onde é realizado o processo de segmentação e extração de atributos que possibilitam a percepção de padrões de cada imagem. Posteriormente, as imagens passam a ser dados de entrada para uma Rede Neural Artificial (RNA), que tem a função de realizar o reconhecimento a partir dos padrões obtidos no treinamento (JUNIOR N., 2008).

Na imagem digital um sistema de cores é utilizado para gerenciamento numérico da coloração dos pixels das imagens de forma individual. Vale ressaltar que o pixel corresponde a menor dimensão de um elemento capaz de ser reproduzido em um dispositivo eletrônico (monitores, televisores e demais tipos de telas). Cada pixel é formado por um conjunto de composições de cores, sendo o mais utilizado, o sistema compostos 3 (três) camadas de cores. Este sistema é denominado de RGB (*Red, Green and Blue*) e compreendem camadas de vermelho, verde e azul definidas em uma escala entre 0 (zero) e 255 (duzentos e cinquenta e cinco). A variação destes valores permite a formação de qualquer cor (**Figura 11**), onde, com a união destes pixels é possível à geração das imagens (HERRMANN, 2019).

Figura 11 – Formação de cores RGB



Fonte: <http://www.afixgraf.com.br/o-que-significa-rgb/>

O processo de aquisição das imagens corresponde ao início da visão computacional, pois se trata da entrada de dados no processo de visão computacional. Em síntese, é composto por uma câmera que capta as imagens e converte as mesmas em digitais para serem enviadas ao pré-processamento. Uma das principais variáveis para esta etapa é a iluminação da cena a ser capturada, pois a luminosidade impacta diretamente da reflexão de luz oriunda dos objetos em direção as lentes da câmera (GALLON, 2013).

No processo de aquisição das imagens, quando a luminosidade não está adequada ou existem outros fatores externos que impactam negativamente na qualidade da obtenção dos dados (imagens), é necessário que o mesmo passe por um pré-processamento, a fim de minimizar tais impactos negativos. Nesta etapa são aplicados pela arquitetura computacional alguns filtros que reduzem tal interferência (**Figura 12**), muitas vezes apresentadas em forma de ruído que comprometem o processamento e eficiência da rede neural durante o processo de execução, de forma a mitigar tal problemática e garantir que o reconhecimento das mesmas ocorra de forma assertiva (MARENGONI; STRINGHINI, 2009).

Figura 12 – Processo de redução de ruído por filtro



Fonte: MARENGONI E STRINGHINI (2009)

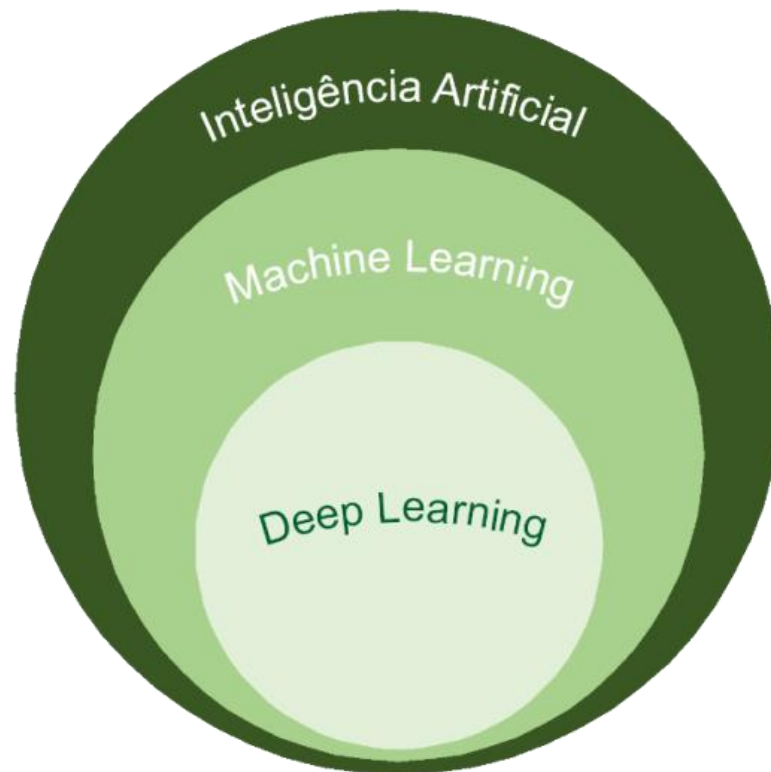
Além deste, existem outros filtros, com distintas funções podem ser aplicados, a fim de contribuir para o reconhecimento da imagem pela Rede Neural Convolucional (RNC). Estes filtros variam conforme arquitetura da rede neural a ser utilizada, podendo ter aplicação na alteração de contraste, saturação, inversão de cores, passa-alta, redução de tamanho da imagem, dentre outros. Geralmente, as redes utilizam métodos que convertem as imagens em formatos e tamanhos específicos, também definidos pelo desenvolvedor da rede. O objetivo destas aplicações é melhorar o desempenho e processamento das imagens durante o reconhecimento (ORLANDINI, 2012).

3.7.2 Deep Learning

O *Deep Learning* (DP), que traduzido para o português significa aprendizagem profunda é um método intrínseco ao *Machine Learning* (ML) caracterizado pela aprendizagem de máquinas que, por sua vez, está inserido no ramo de estudo da IA (**Figura 13**), onde é possível verificar a relação existente entre os três temas citados nos tópicos anteriores.

Vale ressaltar que o surgimento da aprendizagem profunda de máquinas trouxe à tona novamente a IA, pois, por volta de 1970, a mesma passava por um período de redução nos financiamentos dos estudos e mitigação dos avanços em decorrência da limitação tecnológica naquele período (CHAGAS, 2019).

Figura 13 – Relação entre Inteligência Artificial, *Machine Learning* e *Deep Learning*

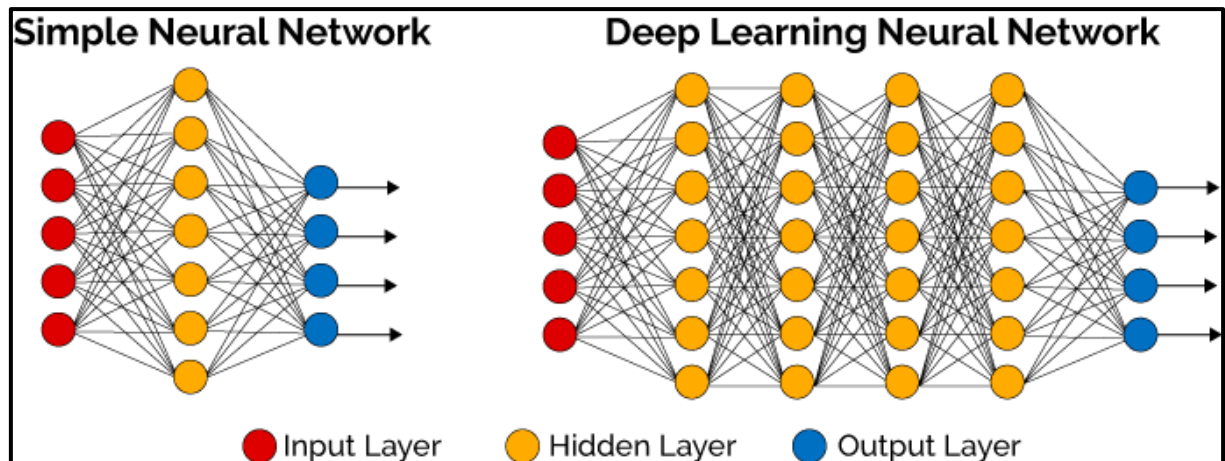


Fonte: Adaptado de CHAGAS (2019)

O modelo de neurônio computacional não se difere em grandes proporções do neurônio humano. No sistema nervoso humano, cada neurônio recebe, processa e emite sinais por intermédio de uma estrutura denominada axônio. Da mesma forma ocorre no neurônio artificial, onde o mesmo é munido de uma entrada de dados (*input layers*), filtra os dados importantes (*hidden layers*) e emite sinais de saída (*output layers*). Quando tem mais de uma camada entre a entrada e saída, conhecidas como camadas ocultas, caracteriza-se o *Deep Learning* (GUIMARÃES, 2018).

Desta forma, pela existência de multicamadas, os dados que saem de uma camada entram na próxima sucessivamente até que seja alcançada a camada de saída. Com isso, é possível extrair pontos essenciais de cada imagem captadas pela visão computacional, melhorando a eficiência dos padrões obtidos (PONTI; COSTA, 2017). No entanto, vale salientar que existem diferenças fundamentais a serem compreendidas na comparação entre o ML e DL, principalmente, em relação ao número de camadas ocultas, que outrora ocorria em monocamada e com o surgimento do *Deep Learning*, em multicamadas (**Figura 14**).

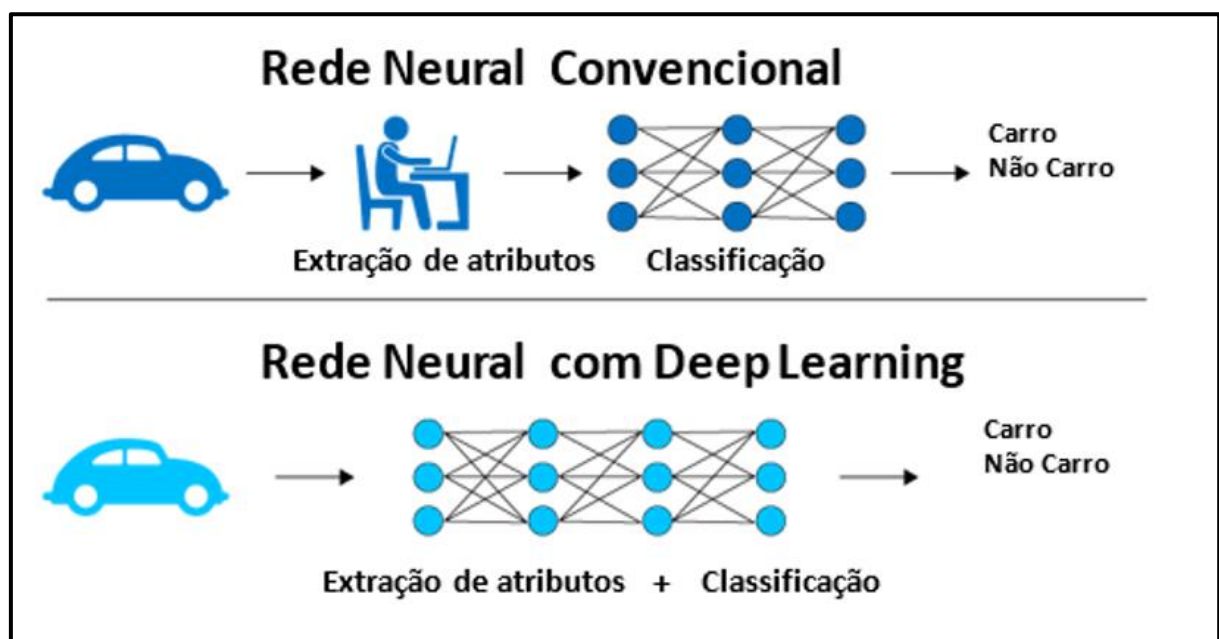
Figura 14 – Diferença de camadas entre *Machine* e *Deep Learning*



Fonte: CHAGAS (2019)

De acordo com Chagas (2019), um ponto que revolucionou a IA, foi a automatização da extração de atributos (Figura 15). Os atributos correspondem as características retiradas das imagens para realização do treinamento e validação da rede. Deste modo, no DL não é necessária a intervenção humana nesta etapa do processo, garantindo maior autonomia e eficiência da tecnologia. Isto também é uma das principais diferenças entre ML e DL.

Figura 15 - Diferença de extração de atributos entre ML e DL



Fonte: <https://www.3dimensoes.com.br/post/deep-learning-conceitos-e-aplicações>. Acesso em 07/06/2020

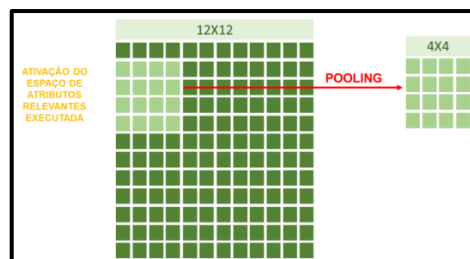
3.7.3 Rede Neural Convolucional

A rede neural convolucional trabalha diretamente com reconhecimento de imagens, utilizando métodos matemáticos de convolução. Existem diversos algoritmos que tornam o processo mais eficiente, sendo umas das principais a técnica de segmentação de Imagens e camada de *Pooling* (PONTI; COSTA, 2017).

Para que o processo seja eficiente, as camadas necessitam ser ativadas em casos onde exista uma característica relevante para o reconhecimento das imagens. No entanto, estas são constituídas de grandes quantidades de pixels. É neste momento que entra em ação a técnica de segmentação das imagens, que obtém as regiões de interesse para a rede neural. Logo após a segmentação, ocorre o processo de ativação dos neurônios que emite sinais para a camada de *pooling* (GALLON, 2013).

A camada de *pooling* tem o objetivo de reduzir o espaço da imagem que anteriormente foi ativada pelos neurônios da rede neural. Com isso é possível reduzir o processamento e a necessidade de memória do *hardware* utilizado no processo, evitando assim, sobrecargas no local aonde a IA está funcionando. Deste modo (**Figura 16**), são processados somente os atributos que serão importantes na contribuição do reconhecimento da imagem (HERRMANN, 2019).

Figura 16 – Processamento da camada de *pooling*



Fonte: Próprio Autor (2020)

Para a aplicação da RNC, existem alguns métodos, dentre eles o *Transfer Line*, onde uma rede pré-treinada é utilizada com base de detecção. Uma das principais redes para este método é a rede Darknet, pois a mesma já é treinada e de livre uso, além de utilizar o método YOLO como base de rastreamento dos objetos a serem identificados. Neste caso, a mesma conta com 80 classes possíveis de serem identificadas (Apêndice 1), mas possibilita o treinamento em casos mais específicos (REDMON, 2016).

4 METODOLOGIA

Para organização metodológica, este trabalho foi dividido em três etapas distintas, seguindo o Processo de Desenvolvimento de Produtos (**Figura 17**). Para Rozenfeld *et al.* (2006), o Processo de Desenvolvimento de Produtos (PDP), trata da estruturação lógica das atividades no processo de desenvolvimento de determinado produto. Este processo engloba a percepção de necessidades mercadológicas, viabilidade econômica e tecnológica, além das estratégias a serem empregadas no projeto.

Figura 17 – Modelo unificado



Fonte: Adaptado de ROZENFELD et al. (2006)

4.1 Identificação e definição das necessidades e requisitos

A primeira etapa tratou da identificação e definição das necessidades e requisitos dos deficientes visuais e dos métodos a serem utilizados no desenvolvimento do projeto. Sendo assim, tendo em vista a existência da real necessidade de incorporar o DV ao meio, foi decidido que a identificação deveria contemplar o maior número de objetos possíveis, de forma a mitigar a problemática,

além de comprovar a eficiência da RNC e demonstrar a sua aplicabilidade no âmbito social.

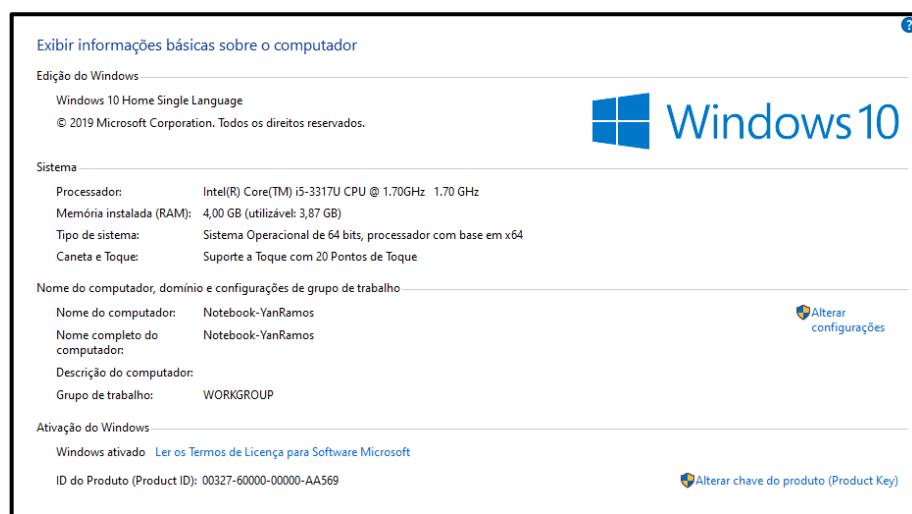
No entanto, foi compreendido que o sistema deveria atender alguns requisitos, onde o sistema fosse capaz de identificar os mais variados tipos de objetos e corpos (versatilidade de detecção), mas sem comprometer a qualidade do processamento (*hardware*), nem o custo do projeto, a fim de possibilitar a aplicabilidade econômica e tecnológica dentro do contexto apresentado neste trabalho.

Sendo assim, pela viabilidade apresentada, verificou-se que a aplicação mais adequada seria através de uma rede neural convolucional pré-treinada, onde a redução da carga de processamento seria menor, tendo em vista que a arquitetura computacional de desenvolvimento do código já é aprimorada para uso no âmbito experimental. Desta forma, foi empregada a rede *Darknet*, pois a mesma utiliza o método Yolo em seu estado de funcionamento.

4.2 Desenvolvimento do projeto

Para a fase de desenvolvimento do projeto, foi necessário inicialmente segregar a construção em duas etapas, sendo a etapa inicial desenvolvida e testada alocada em um computador, neste caso o *Ultrabook* Lenovo® munido das configurações mínimas necessárias para execução da rede (**Figura 18**).

Figura 18 – Configurações do *Ultrabook* utilizado



Fonte: Próprio Autor (2021)

Vale ressaltar que a configuração apresentada, trata-se da básica e minimamente necessária para o funcionamento do código a ser desenvolvido em conjunto com os *softwares* a serem utilizados, além do processamento necessário para execução da RNC.

Já munido do *hardware* inicialmente necessário, foi levado em consideração às necessidades do sistema em contrapartida com a linguagem de programação a ser utilizada no projeto. Devido à vasta aplicação e facilidade de compreensão em razão do nível da linguagem e da gama de produções teóricas intrínsecas ao meio acadêmico e científico, foi decidido que o código deveria ser escrito em linguagem de programação Python.

4.3 Ferramentas aplicadas à rede neural

Mediante aplicabilidade, a linguagem Python que se caracteriza pela simplicidade de escrita e pelo alto nível de programação, foi à opção adotada para desenvolvimento do *software* a ser empregado no dispositivo. Além disto, pela rede *Darknet* já ser desenvolvida e treinada foi considerado que o Python seria a melhor opção para desenvolvimento do trabalho em questão. Um ponto fundamental a ser avaliado, é que esta rede consegue identificar até 80 classes distintas, o que a torna eficiente para experimentos acadêmicos. Mas, vale ressaltar que existem outras linguagens e ambientes de desenvolvimentos capazes de realizar a mesma atividade de rede neural artificial a qual foi utilizada.

Além da definição da linguagem de programação, para o desenvolvimento do código, foi necessário à instalação de alguns módulos computacionais (*softwares*) que possibilitaram o desenvolvimento e execução da rede neural convolucional. O primeiro deles é o ambiente de desenvolvimento (IDE), *Visual Studio Code* (VSC), onde o código é escrito e compilado, mas também foi testado na IDE Python 3.9, tendo o mesmo funcionamento entre ambos ambientes.

Além deste, também foi necessário a instalação do OpenCV (*Open Source Computer Vision*), sendo esta uma biblioteca de código aberto voltada a aplicação nos métodos de visão computacional, aprendizagem de máquina e aceleração de GPU (*Graphics Processing Unit*) simultâneo ao funcionamento da rede neural, que originalmente foi desenvolvida pela Intel. Esta biblioteca foi escolhida, pois, além de ser livre para uso acadêmico, é compatível com diversas linguagens de

programação e sistemas operacionais, dentre eles o Python, Windows e Linux que foram utilizados no desenvolvimento do trabalho em questão.

De mesmo modo, o uso do método YOLO, foi de fundamental importância, pois ao inter-relacionar o sistema desenvolvido com a problemática apresentada, o mesmo permite que o deficiente visual se desloque e seja realizado o rastreamento dos objetos identificados pela câmera em tempo real, o que é tratado como um fator preponderante para avaliação e aplicação da RNA dentro do contexto do DV.

4.4 Desenvolvimento do código

A partir da obtenção e instalação das tecnologias citadas anteriormente, foi possível o início da escrita do código (Apêndice 1). Primitivamente, como na maioria das linguagens de programação, é imprescindível a inclusão das bibliotecas, também denominadas de dependências, necessárias para o uso das funções pertinentes aos códigos. No caso deste trabalho, foram necessárias seis, sendo elas, *cv2*, *time*, *numpy*, *speech_recognition*, *gTTS* e *playsound* (**Figura 19**)

Figura 19 – Inclusão das dependências fundamentais

```
#Importa as dependências necessárias
import cv2
from cv2 import cv2
import time
import numpy as np
import speech_recognition as sr
from gtts import gTTS
from playsound import playsound
```

Fonte: Próprio Autor (2021)

Como forma de melhoramento da identificação e rastreamento dos corpos de forma individual e simultânea, foi percebido que a utilização de cores distintas seria de suma importância neste projeto, pois possibilitaria durante a execução uma melhor visualização por parte dos testes de bancada a serem realizados. Para tanto, foi incluído uma linha de código que trataria das limitações da escala de cor a serem apresentadas, fazendo uso da função *colors*, onde a variação para cada classe é aplicada (**Figura 20**).

Figura 20 – Variação das cores por classes

```
#Estabelece variação das cores do retângulo de detecção
COLORS = [(0,255,255), (255,255,0), (0,255,0), (255,0,0)]
```

Fonte: Próprio Autor (2021)

Posteriormente, foi necessário fazer uso de um arquivo de texto que contivesse o nome das classes a serem identificadas pela rede YOLO. Neste caso, foi feito uso das funções que carregassem o arquivo “classes.name” dentro da rede e fizesse a ligação das linhas deste arquivo com os objetos identificados pela RNC (Figura 21).

Figura 21 – Carregamento das classes pré-treinadas

```
#Carrega as classes de identificação
class_names = []
with open("classes.names", "r") as f:
    class_names = [cname.strip() for cname in f.readlines()]
```

Fonte: Próprio Autor (2021)

O próximo passo foi de suma importância, pois neste definimos a câmera a ser utilizada para captação das imagens, a serem usadas na execução da RNA, mediante rastreamento da rede YOLO (Figura 22). O valor definido dentro da função *VideoCapture* é responsável por tal definição, onde o valor 0, neste momento, refere-se a câmera padrão do computador de execução da rede, podendo variar conforme o número de câmeras pré-instaladas no equipamento em questão.

Figura 22 – Definição da câmera a ser utilizada

```
#Define a câmera utilizada para detecção
cap = cv2.VideoCapture(0)
```

Fonte: Próprio Autor (2021)

Como citado anteriormente, o processamento da rede deve ocorrer de forma fluída, onde seja possível identificar o objeto sem que comprometa o funcionamento do sistema. Para tanto, a rede *Darknet* faz disponível dois modos de rede, sendo

que o modelo *tiny* apresenta menor uso de processamento de memória, mas em contrapartida tem um *delay* maior na detecção, além de apresentar menor acurácia na identificação de alguns corpos. Como o objetivo deste trabalho foi definido para aplicação embarcada, foi decidido o uso do modelo simplificado, acompanhado da definição de parâmetros básicos, a fim de possibilitar uso de equipamentos de baixo custo, viabilizando o projeto (**Figura 23**).

Figura 23 – Carregamento da rede e definição dos parâmetros

```
#Carrega a rede, podendo ser leve ou pesada
#net = cv2.dnn.readNet("yolov4.weights", "yolov4.cfg")
net = cv2.dnn.readNet("yolov4-tiny.weights", "yolov4-tiny.cfg")

#Define os parâmetros da rede
model = cv2.dnn_DetectionModel(net)
model.setInputParams(size=(416,416),scale=1/255)
num = 0
```

Fonte: Próprio Autor (2021)

Com todas as definições e configurações já aplicadas, foi estabelecido um *loop* infinito (até que seja finalizada a execução do sistema) para que a detecção se inicie. Intrínseco a este, são obtidos os *frames* do vídeo capturado, de forma a segrega-los para que todos os parâmetros de detecção sejam aplicados em cada um deles. Dentre estes parâmetros estão o *label* do objeto identificado, cor da classe, retângulo de delimitação do objeto rastreado e o FPS (quantidade de fotos por segundos).

Inicialmente, dentro do *loop* o frame é capturado e passa por redimensionamento e interpolação. Estas ações são aplicadas a fim de padronizar as imagens e possibilitar que a RNC faça a detecção do objeto. Para isso, utilizam-se algumas funções definidas pela dependência cv2 (**Figura 24**).

Figura 24 – Redimensionamento e interpolação do frame capturado

```
#Segregação dos frames
ret, frame = cap.read()
try:
    frame = cv2.resize(frame, (416, 416), interpolation=cv2.INTER_AREA)
except:
    break
```

Fonte: Próprio Autor (2021)

Como forma de controle da rede, foi necessário que fosse calculado o tempo de execução da função principal do sistema, que é a detecção propriamente dita do objeto alocado no campo de visão da câmera utilizada. Para isto, foi contado o tempo em milissegundos logo anterior e posterior a execução da função “*model.detect*” (**Figura 25**). Vale ressaltar que esta função é a responsável por estabelecer a conexão entre o objeto capturado pela câmera e as classes pré-treinadas da rede, além de medir a acurácia de detecção.

Figura 25 – Cronometragem e detecção do objeto

```
#Início da contagem do tempo levado para detecção em milissegundos
start = time.time()

#Define o que foi identificado
classes, scores, boxes = model.detect(frame, 0.1, 0.2)

#Fim da contagem do tempo levado para detecção em milissegundos
end = time.time()
```

Fonte: Próprio Autor (2021)

Com o objeto devidamente identificado, foi necessária a apresentação do resultado, onde o nome da classe foi apresentado em um objeto de texto, denominado de *label*, com uso das cores estabelecidas, além do retângulo de rastreo (**Figura 26**).

Figura 26 – Definições para apresentação dos resultados

```
#Percorre todas as detecções
for(classid, score, box) in zip(classes, scores, boxes):

    #Gera uma cor para a classe definida
    color = COLORS[int(classid)%len(COLORS)]

    #Busca o nome da classe pelo ID dentro do arquivo classes.name, calcula
    label = f"{class_names[classid[0]]}: {score*100}%"

    #Desenha o retângulo de rastreo de detecção
    cv2.rectangle(frame, box, color, 2)

    #Escreve o nome da classe em cima do retângulo do objeto
    cv2.putText(frame, label, (box[0], box[1]-10), cv2.FONT_HERSHEY_SIMPLEX,
```

Fonte: Próprio Autor (2021)

No entanto, é também neste momento que a atuação principal do objetivo deste trabalho deve acontecer, onde a detecção realizada deve ser informada por intermédio de sinais sonoros ao deficiente visual o que encontra a sua volta, possibilitando a definição da ação por parte do indivíduo. Para isso, são utilizadas funções das bibliotecas *gTTS*, *speech_recognition* e *playsound* já declaradas no início do código.

Deste modo, é definida inicialmente uma *string* que contém em seu escopo a frase a ser ouvida pelo DV. Para isto, a frase é convertida em um arquivo de áudio através da função *gTTS* e salva no diretório estabelecido como arquivo de extensão *mp3*. Com a frase já convertida, é chamada a função *playsound* que inicia a execução do áudio, explicitando deste modo o objeto ao portador da necessidade especial (**Figura 27**).

Figura 27 – Criação e execução da emissão sonora do objeto identificado

```
#Emissão de áudio do que foi detectado
audio = 'Estou vendo: ' + class_names[classid[0]]
tts = gTTS(audio, lang='pt-br')
num = num + 1
arquivo = 'audios/arquivo'+ str(num) + ".mp3"
tts.save(arquivo)
playsound(arquivo)
```

Fonte: Próprio Autor (2021)

Para finalizar a execução do *loop*, o FPS é calculado e exibido conjuntamente com a imagem já totalmente munida dos atributos citados nas funções anteriores (classe, cor, FPS, acurácia, etc.), finalizando com a definição do estado do processamento por parte do *cv2*, até que a janela do sistema seja fechada e destruída da memória *flash*, liberando a câmera para outras aplicações do equipamento em que a rede estiver alocada e em execução.

Nestas funções (**Figura 28**), é possível perceber a identificação como forma de organização do código, além de ser um dos fatores de extrema importância na linguagem Python, pois a lógica de programação é totalmente modificada caso haja ou não o espaçamento entre a lateral esquerda do IDE e a linha do código escrito.

Figura 28 – Finalização da detecção e liberação da câmera

```
#Calcula o FPS
fps_label = f"FPS: {round((1.0/(end-start)),2)}"

#Escreve o FPS na imagem
cv2.putText(frame, fps_label, (0,25), cv2.FONT_HERSHEY_SIMPLEX, 1, (0,0,0),
cv2.putText(frame, fps_label, (0,25), cv2.FONT_HERSHEY_SIMPLEX, 1, (0,255,0)

#Mostra a imagem detectada com as escritas do objeto detectado, do FPS,
#da acurácia, da detecção e o retângulo de rastreo
cv2.imshow("detections", frame)

#Aguarda resposta de processamento
if cv2.waitKey(1) & 0xFF == ord('q'):
    break

#Libera a câmera e fecha todas as janelas
cap.release()
cv2.destroyAllWindows()
```

Fonte: Próprio Autor (2021)

4.5 Aplicação embarcada no Raspberry® Pi Model 3B+

Análogo ao modelo já apresentado na metodologia deste trabalho aplicado ao sistema alocado em computador convencional, o modelo embarcado obedeceu aos mesmos parâmetros e passos descritos anteriormente. No entanto, para que isto seja possível, foi necessário que o Raspberry® Pi Model 3B+ (**Figura 29**) ou semelhantes, estivesse munido do sistema operacional Raspbian para que o mesmo funcionasse da mesma forma que um computador tradicional e possibilitasse a execução do código sem anomalias.

Figura 29 – Raspberry® Pi Model 3B+



Fonte: Próprio Autor (2021)

Para tanto, foi necessário o uso de um SD Card de no mínimo 16Gb de memória, onde o sistema operacional foi alocado. Inicialmente, o cartão foi formatado mediante uso do *software SD Card Formatter* (**Figura 30**), após ser baixado e instalado de forma simplificada pelo navegador padrão do computador.

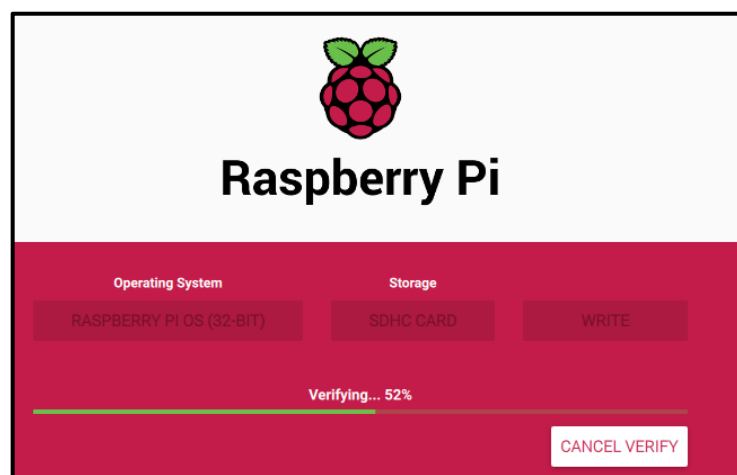
Figura 30 – SD Card Formatter



Fonte: Próprio Autor (2021)

Para que a realização da formatação bastou iniciar o *software*, selecionar o dispositivo conectado e iniciar o processo. Feito isto, o mesmo esteve pronto para receber o Raspbian. Deste modo, bastou realizar o *download* do mesmo (**Figura 31**), instalar dentro do SD Card e inserir no Raspberry® para que fosse reconhecida a instalação e o processo de embarcamento prosseguisse conforme o modo já apresentado anteriormente. Existem diversas versões que podem ser utilizadas, sendo que neste trabalho foi instalada a versão oficial que a Raspberry® disponibilizado no site oficial da empresa, em específico para instalação mediante uso do Windows 10.

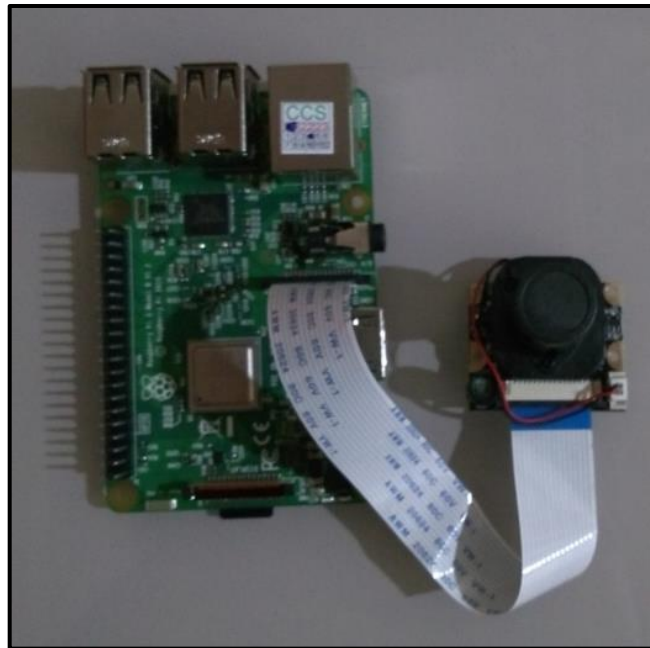
Figura 31 – Instalando Raspbian



Fonte: Próprio Autor (2021)

Vale salientar que para o trabalho em questão, foi utilizada uma câmera de uso específico ao Raspberry® (neste caso, um modelo genérico com infravermelho disponível para melhor captação noturna das imagens), onde através de um cabo flat é estabelecida tal conexão (**Figura 32**).

Figura 32 – Montagem Modelo Embarcado



Fonte: Próprio Autor (2021)

Mediante disponibilidade de conexão P2 na placa Raspberry®, o uso do fone de ouvidos é de suma importância por parte do DV, pois desta forma, é minimizada a influência dos sons intrínsecos ao meio social, de forma que torne o uso do equipamento mais assertivo e confiável perante situações práticas do cotidiano dos indivíduos acometidos pela patologia visual.

Vale ressaltar que apenas o embarcamento foi responsável pelos custos deste trabalho, tendo em vista que o *Ultrabook* já estava disponível e não foi contabilizado nos cálculos. Desta forma, o Raspberry® representou R\$350,00, a câmera específica R\$80,00 e o *SD Card* R\$40,00, totalizando um custo de R\$470,00 em todo o projeto.

5 RESULTADOS E DISCUSSÕES

Como forma de teste, foi iniciado o funcionamento da rede neural convolucional de forma a detectar alguns objetos de possível reconhecimento por parte da Darknet, como forma de verificação de acurácia e viabilidade do sistema desenvolvido. Para tanto, verificou-se que, primordialmente, devia ocorrer em ambiente sem implicações de controles, onde as influências externas são evidenciadas, favorecendo assim, melhor avaliação do trabalho para aplicação prática em auxílio aos deficientes visuais. Logo, a influência do ambiente ocorreu de forma livre e sem nenhum tipo de controle, onde se buscou um ambiente mais próximo do meio ao quais os DVs estão inseridos em suas rotinas.

Com esta definição, o sistema foi posto em estado de funcionamento, onde foi possível verificar os resultados apresentados posteriormente neste tópico. No entanto, foi realizado o teste, primitivamente, no ambiente de desenvolvimento alocado no Ultrabook, a fim de avaliar as variações impostas pelo dispositivo a qual a rede estava alocada, além de controlar o funcionamento de forma visual por meio do monitor. Neste teste, foi utilizado o aplicativo DroidCam para enviar imagens em tempo real via Wi-fi para o computador responsável por rodar a rede neural. Para tanto, o teste foi realizado no interior de uma residência de forma livre. Como tratado anteriormente, os objetos reconhecidos (**Figura 33**) correspondem a algumas das 80 classes que se encontram listadas no Apêndice 2.

Figura 33 – Alguns dos objetos a identificados



Fonte: Próprio Autor (2021)

Este teste foi realizado, inicialmente, com duas motocicletas distintas e de forma parcial, a fim de verificar que a rede identificava de forma adequada ambos os veículos, sem levar em consideração apenas um dos parâmetros, evidenciando assim o pleno funcionamento da rede neural convolucional, onde é possível verificar a aplicação em execução para este caso em específico (**Figura 34**).

Figura 34 – Rede identificando motocicletas distintas e estáticas



Fonte: Próprio Autor (2021)

Como forma de avaliação em meios distintos, foi realizado o teste também de uma pessoa, mas agora em um ambiente fechado e com luminosidade comprometida. Neste, foi possível que apesar das circunstâncias implicadas por tais influências, a rede teve bom comportamento no reconhecimento e classificação do indivíduo (**Figura 35**).

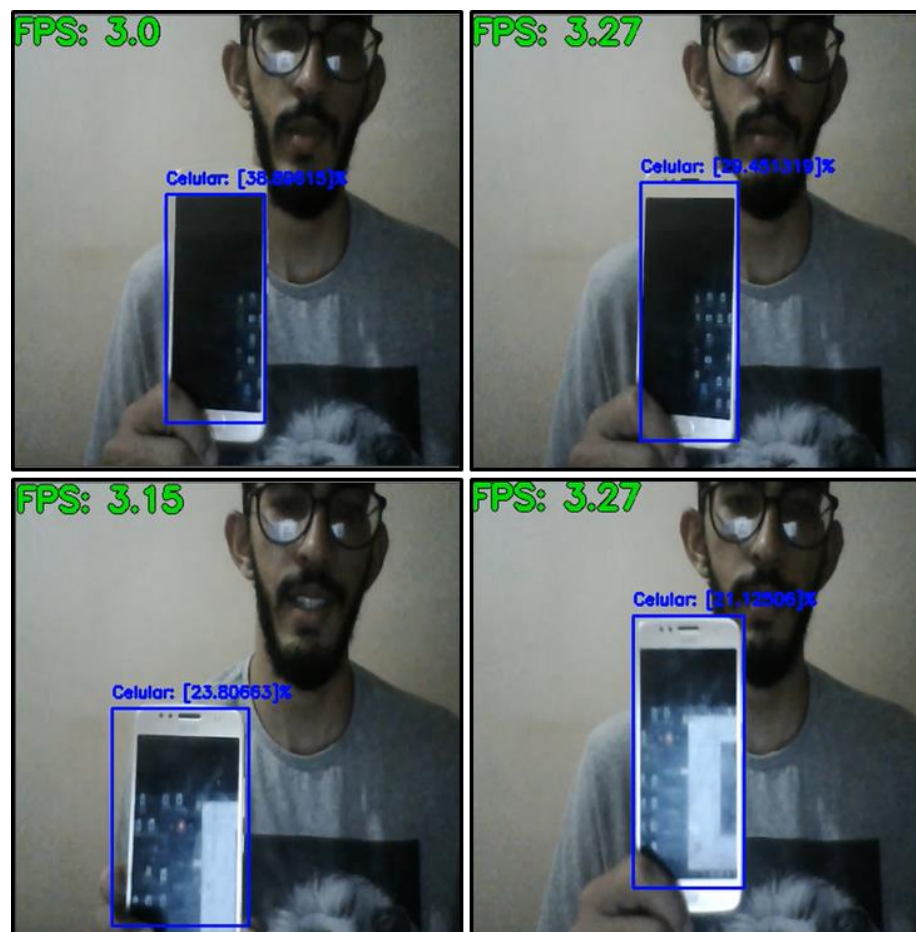
Figura 35 – Rede identificando pessoa em ambiente fechado



Fonte: Próprio Autor (2021)

Ainda, foram realizados testes no reconhecimento de aparelhos celulares, tendo como principal objetivo, avaliar se a rede teria seu funcionamento comprometido neste tipo de situação, de forma a simular condições de aplicabilidade reais do cotidiano. Com forma avaliativa, foi identificado um aparelho celular em ambiente com baixa luminosidade (**Figura 36**) e de forma adequada informada por meio sonoro pelo sistema.

Figura 36 – Rede identificando aparelho celular em ambiente com baixa luminosidade



Fonte: Próprio Autor (2021)

Posteriormente, foi estabelecido o teste principal, onde foi possível realizar um teste em tempo real, simulando as necessidades dos deficientes visuais no interior de suas residências. Para isso, o teste foi alocado no computador, da mesma forma apresentada anteriormente, pois desta forma poderiam ser aplicadas diversas variações de acurácia, versão da rede, dentre estabelecimento de demais parâmetro, além do uso gráfico de algumas funções aplicadas no código, a fim de

evidenciar o pleno funcionamento do sistema. Dentre todas as variações, as que apresentaram melhores resultados, foram às redes definidas para emissão de áudio à partir de 60% da acurácia, tanto na versão compacta (uso de vídeo como apresentação de resultado), quanto com o uso da versão normal da Darknet, porém com uso apenas de imagens como validação dos resultados. Ambos os resultados estão descritos em forma de vídeo no link: <https://youtu.be/CPCtsK4JkfM>.

Para validação do sistema móvel, os mesmos testes foram realizados dentro da Raspberry®, mas devido à limitação imposta pelo *hardware*, não foi realizado a validação visual (foto ou vídeo) durante o funcionamento da rede, restringindo-se ao módulo de voz. Ainda, houve limitação imposta pelos custos do projeto, onde a identificação ocorreu em um espaço limitado, tento em vista que a fonte de energia tratou-se de uma fonte com fio aplicada na tomada residencial. Além disso, o áudio foi gravado através de um *notebook*, captando o volume máximo do fone de ouvidos conectado a placa.

Mas, vale salientar, que da mesma forma que apresentada quando alocada no computador, a versão compacta da rede teve os mesmos resultados apresentados na versão alocada em computador e pode ser visto no mesmo link disponibilizado nos demais testes.

Com tais testes, foi possível verificar os pontos positivos no funcionamento da rede neural e validar que o uso desta se dá de forma adequada tendo em vista as influências a qual foi submetida. Todavia, ainda foi possível verificar as necessidades de melhorias a serem aplicadas, objetivando maior autonomia do deficiente visual em seus trajetos rotineiros.

6 CONCLUSÃO

Como foi possível verificar nos testes descritos anteriormente, o objetivo proposto neste trabalho foi atingido, onde se prendeu a necessidade de identificação diversificada de objetos e corpos intrínsecos ao meio, onde o código desenvolvido da rede neural convolucional (Darknet) teve seu funcionamento mediante processo de visão computacional por intermédio da rede neural artificial e emissão sonora para captação por parte do deficiente visual.

Além disto, pela alta capacidade de aplicação em ambientes distintos e com influências sociais divergentes, verificou-se que o uso da mesma é de grande valia na inclusão social por parte dos indivíduos patológicos, principalmente com o uso embarcado, que torna o sistema móvel.

Durante os testes, verificou-se que apesar da alta capacidade da rede utilizada, a alternativa mais adequada (funcionalmente falando) é o treinamento da rede específica para aplicação neste dispositivo, tendo em vista que em alguns momentos de maior exigência da rede o objeto pode ser identificado de forma incoerente, limitando o uso como dispositivo secundário. No entanto, a aplicabilidade econômica para esta realização, incorreria, pois a necessidade de um processamento maior acarretaria no hardware a ser utilizado.

Analogamente, ainda ficaram perceptíveis alguns pontos de melhorias a serem aplicados em versões posteriores, onde a aplicação da impressão 3D na confecção de métodos de fixação da câmera ao usuário e o uso de um hardware específico para atuadores, trariam resultados cada vez mais satisfatórios para aplicação real do dispositivo.

Contudo, pela limitação de hardware aplicada, a rede ficou avaliada qualitativamente como adequada para uso experimental, apresentando resultados satisfatórios que podem ser utilizados como precedentes para desenvolvimento de dispositivos cada vez mais tecnológicos e eficientes no que diz respeito ao auxílio e inclusão dos deficientes visuais no âmbito social com segurança.

REFERÊNCIAS

ASSIS, Diva. **O caminhar da pessoa cega: Análise da exploração de elementos do espaço urbano por meio da bengala longa**. 2018. Trabalho de conclusão de curso (Pós-graduação em arquitetura e urbanismo) - Faculdade de Arquitetura e Urbanismo, Universidade Federal de Alagoas, Maceió, 2018. Disponível em: <http://www.repositorio.ufal.br/handle/riufal/2467>. Acesso em: 8 maio 2020.

BAIÃO, André. **Simulador óptico dinâmico do olho humano**. 2013. Dissertação (Mestrado em Engenharia Biomédica) - Faculdade de Ciência e Tecnologia, Universidade Nova de Lisboa, Lisboa, 2013. Disponível em: https://run.unl.pt/bitstream/10362/9751/1/Baiao_2013.pdf. Acesso em: 13 maio 2020.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS - ABNT. NBR 9050, 2004. **Acessibilidade a edificações mobiliárias, espaço e equipamentos urbanos**, [S. l.], 31 maio 2004. Disponível em: http://www.turismo.gov.br/sites/default/turismo/o_ministerio/publicacoes/downloads_publicacoes/NBR9050.pdf>. Acesso em: 13 maio 2020.

BATISTA, Cecília; ENUMO, Sonia. **Desenvolvimento humano e impedimentos de origem orgânica: o caso da deficiência visual**. 2000. Trabalho de conclusão de curso (Pós-graduação em Psicologia) - Universidade Federal do Espírito Santo, Vitória, 2000. Disponível em: <http://www.lite.fe.unicamp.br/cursos/nt/ta2.4.htm>. Acesso em: 13 jun. 2020.

BRASIL. **Lei nº 13.146, de 6 de julho de 2015**. Diário Oficial da República Federativa do Brasil. Brasília, 6 jul. 2015. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm. Acesso em: 17 maio 2020.

BRASIL. Ministério da Educação. Secretaria de Educação a Distância. **Cadernos da TV Escola**. Brasil, 2000. 6 p.

BRASIL. Ministério da Saúde. Secretaria de Atenção à Saúde. **Cadernos temáticos do PSE – Saúde Ocular**. Brasil, 2016. 28 p.

CBO – Conselho Brasileiro de Oftalmologia. **As condições de Saúde Ocular no Brasil.** 1 ed. 2019. 10 p. Disponível em: <http://www.cbo.com.br/novo/publicacoes/condicoes_saude_ocular_brasil2019.pdf>. Acesso em: 18 de maio de 2020.

CHAGAS, Edgar. Revista Científica Multidisciplinar Núcleo do Conhecimento. **Deep Learning e suas aplicações na atualidade**, [s. l.], v. 4, ed. 5, 2019. Disponível em: <https://www.nucleodoconhecimento.com.br/it/2019/05>. Acesso em: 15 maio 2020.

CÔRREA, Edílson; BOTEON, Joel; MOLINARI, Luiz; MOREIRA, Angela; OLIVEIRA, Bruno. Nescon. **Avaliação ocular de crianças e adolescentes na atenção básica à saúde**, [s. l.], 2016. Disponível em: https://www.nescon.medicina.ufmg.br/biblioteca/imagem/avalia%C3%A7ao_ocular_crian%C3%A7a.pdf. Acesso em: 17 maio 2020.

COSTA, Ana; VAGNER, SANTOS. Revista Brasileira de Oftalmologia. **Da visão à cidadania: tipos de tabelas de avaliação funcional da leitura na educação especial**, São Paulo, 2018. Disponível em: https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-72802018000500296. Acesso em: 20 maio 2020.

COSTA, Rayssa; SANTOS, Diogo; SANTOS, Cleiton; SOUZA, Simone; LIMA, Fernando. **Desenvolvimento de uma bengala automatizada para deficientes visuais utilizando arduino**. 2016. Anais inovação, tecnologia, gestão e sustentabilidade, 2016.

INOVAÇÃO, TECNOLOGIA, GESTÃO E SUSTENTABILIDADE, 2018, Tangará da Serra. **Anais** [...]. [S. l.: s. n.], 2018. v. 4. Disponível em: <https://even3.blob.core.windows.net/anais/102473.pdf>. Acesso em: 21 maio 2020.

COUTINHO, Raul. Histórico da informática. **Introdução à Informática**: História da Informática - Um olhar pelo sistema, [s. l.], 2008. Disponível em: <http://kessia.blogs.unipar.br/files/2009/02/ahistriadocomputador.pdf>. Acesso em: 22 maio 2020.

DUMMER, Karine. **Bengala Eletrônica para Deficientes Visuais: contribuição à locomoção nos centros urbanos**. 2018. Trabalho de conclusão de curso

(Bacharelado em Engenharia de Controle e Automação) - Centro de Engenharias, Universidade Federal de Pelotas, Pelotas, 2018. Disponível em: https://wp.ufpel.edu.br/labserg/files/2018/03/o_tcc_dummer_2018.pdf. Acesso em: 20 maio 2020.

EBERMAM, Elivelto; KROHLING, Renato. Revista de Sistemas de Informação da FSMA. **Uma Introdução Compreensiva às Redes Neurais Convolucionais: Um Estudo de Caso para Reconhecimento de Caracteres Alfabéticos**, Vitória, 2018. Disponível em: http://www.fsma.edu.br/si/edicao21/FSMA_SI_2018_1_Principal_08.pdf. Acesso em: 21 maio 2020.

FARINHA, Paula. **A física da visão no ensino**. 2011. Dissertação (Mestrado em ensino da física e da química) - Universidade da Beira Interior, Covilhã, 2011. Disponível em: <https://ubibliorum.ubi.pt/bitstream/10400.6/2494/1/Dissertac...pdf>. Acesso em: 17 maio 2020.

FELICETTI, Suelen; BERNARDINE, Angelita; CARTELI, Zulméia; SANTOS, Sandro. Revista Eletrônica Interdisciplinar. **Acessibilidade, orientação e mobilidade: um estudo de caso considerando o ponto de vista de pessoas cegas ou com baixa visão**. Matinhos, v. 9, 2016. Disponível em: <https://revistas.ufpr.br/diver/article/view/50082>. Acesso em: 18 maio 2020.

FERNANDEZ, Marcial; CORTÉS, Mariela. **Introdução à computação**. 3. ed. Fortaleza: UECE, 2015.

GALLON, Lucas. **Sistema de visão computacional para classificação de pedras naturais através de vídeo em tempo real**. 2013. Trabalho de conclusão de curso (Graduação em Sistema de Informação) - Centro Universitário Univates, Lajeados, 2013. Disponível em: <https://www.univates.br/bdu/bitstream/10737/501/1/2013LucasGallon.pdf>. Acesso em: 24 maio 2020.

GAMBARATO, Vivian; BATISTA, Ana; GIANDONI, Larissa. Revista Tekhne e Logos. **Uso de tecnologias assistivas na educação superior tecnológica**, [s. l.], v. 3, ed.

1, 2012. Disponível em: <http://revista.fatecbt.edu.br/index.php/tl/article/view/126/0>. Acesso em: 18 maio 2020.

GARCIA, Mário. **Cegueira congênita e adquirida: Implicações na saúde mental e resiliência**. 2014. Dissertação (Mestrado em Psicologia, Aconselhamento e Psicoterapias) - Escola de Psicologia e Ciência da Vida - Universidade Lusófona de Humanidade e Tecnologias, Lisboa, 2014. Disponível em: <http://recil.grupolusofona.pt/bitstream/handle/10437/6424/DISSERTA%C3%87%C3%83O%20-%20M%C3%81RIO%20RUI%20SANCHES%20DA%20COSTA%20GARCIA.pdf?sequence=1>. Acesso em: 18 maio 2020.

GEHLEN, Rubens; NONOHAY, Roberto; AFFONSO, Ligia. **Desenvolvimento de produtos**. 2018. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788595022904/>. Acesso em: 06 maio 2020.

GUEDES, Denyse; GIRÃO, Mardônio. **Meio ambiente e inclusão social: Acessibilidade para os deficientes visuais**. 2017. Tese (Doutorado sanduíche) - Universidade Lusíada do Porto, Porto, Portugal, 2017. Disponível em: <https://www.unaerp.br/documentos/2174-meio-ambiente-e-inclusao-social-acessibilidade-para-os-deficientes-visuais/file>. Acesso em: 25 maio 2020.

GUIMARÃES, Heitor. **Recuperação de informações musicais**: Recuperação de informações musicais: uma abordagem utilizando deep learning. 2018. Trabalho de conclusão de curso (Graduação em Computação e Informação) - Escola Politécnica - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2018. Disponível em: <http://monografias.poli.ufrj.br/monografias/monopoli10025687.pdf>. Acesso em: 24 maio 2020.

HALL, John; GUYTON, Arthur. **Tratado de Fisiologia Médica**. 12. ed. Rio de Janeiro: Elsevier, 2011.

HERRMANN, Gustavo. **Técnicas de visão computacional para classificação de cobras a partir de imagens**. 2019. Trabalho de conclusão de curso (Bacharelado em Ciência da Computação) - Universidade Estadual de Londrina, Londrina, 2019.

Disponível em: <http://www.uel.br/cce/dc/wp-content/uploads/PRELIMINAR-GUSTAVO-CASARIM-HERRMANN.pdf>. Acesso em: 27 maio 2020.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). BRASIL. **Pessoas com deficiência visual**. [S. l.], 2013. Disponível em: <https://sidra.ibge.gov.br/tabela/5753#resultado>. Acesso em: 17 maio 2020.

INSTITUTO NACIONAL DO SEGURO SOCIAL (BRASIL). INSS. **Manual Técnico de Procedimentos de Avaliação Médica Pericial das Funções da Visão**. [S. l.], 2014. Disponível em: http://www.consultaesic.cgu.gov.br/busca/dados/Lists/Pedido/Attachments/510599/R ESPOSTA_PEDIDO_rs396PRESINSSanexo.pdf. Acesso em: 17 maio 2020.

JUNIOR, Archimedes. **Acessibilidade e mobilidade na estimativa de um índice de potencial de viagens utilizando redes neurais artificiais e sistemas de informações geográficas**. 2000. Tese (Doutorado em Engenharia Civil - Transportes) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2001. Disponível em: <https://www.teses.usp.br/teses/disponiveis/18/18137/tde-10112001-160812/publico/RaiaJrT.PDF>. Acesso em: 20 maio 2020.

JUNIOR, Nelson. **Visão Computacional e Interface Homem-Máquina na Animação de Grupos de Humanos Virtuais**. 2008. Dissertação (Mestrado em Computação Aplicada) - Universidade do Vale do Rio dos Sinos, São Leopoldo, 2008. Disponível em: <http://livros01.livrosgratis.com.br/cp057727.pdf>. Acesso em: 27 maio 2020.

KERSCHBAUMER, Ricardo. **Engenharia de controle e Automação: Microcontroladores**. Luzerna: Instituto Federal De Educação, Ciência E Tecnologia Catarinense, 2018. Disponível em: <http://professor.luzerna.ifc.edu.br/ricardo-kerschbaumer/wp-content/uploads/sites/43/2018/02/Apostila-Microcontroladores.pdf>. Acesso em: 19 maio 2020.

LUIZ, Lilian; ROSA, Fabricia; PFITSCHER, Elisete. Encontro de Administração Pública e Governo. **Projeto Cão Guia: Custos para Implementação de um Centro de Treinamento e Formação de Treinadores e Instrutores**, Salvador, 2012.

Disponível em: http://www.anpad.org.br/admin/pdf/2012_EnAPG437.pdf. Acesso em: 29 maio 2020.

MARENGONI, Maurício; STRINGHINI, Denise. Research Gate. **Tutorial: Introdução à Visão Computacional usando OpenCV**, [s. l.], 2009. Disponível em: https://www.researchgate.net/publication/41799461_Tutorial_Introducao_a_Visao_Computacional_usando_OpenCV. Acesso em: 28 maio 2020.

MELEZINSKI, Helen; WIGGERS, Larissa. **Igualdade faz a diferença**: Guia prático para a Inclusão social de pessoas com deficiência. 2015. Trabalho de conclusão de curso (Graduação em Tecnologia em Design Gráfico) - Departamento Acadêmico de Desenho Industrial - Tecnologia em Design Gráfico, Universidade Tecnológica Federal do Paraná, Curitiba, 2015. Disponível em: http://repositorio.roca.utfpr.edu.br/jspui/bitstream/1/4005/1/CT_CODEG_2015_1_01.pdf. Acesso em: 20 maio 2020.

NAKABAYASHI, Luciana. **A contribuição da Inteligência Artificial (IA) na Filosofia da Mente**. 2009. Dissertação (Mestrado em tecnologias da inteligência e design digital) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2009. Disponível em: <https://www.inbot.com.br/artigos/educacional/A-contribuicao-da-Inteligencia-Artificial-na-filosofia-da-mente-Luciana-Akemi-Nakabayashi.pdf>. Acesso em: 28 maio 2020.

OFTALMOLOGIA. In: NETTO, Arlindo. **Semiologia**: Anatomia do olho e exame físico. [S. l.]: MED Resumos, 2010. Disponível em: <https://pt.slideshare.net/JuciVasconcelos/semiologia-08-oftalmologia-anatomia-do-olho-e-exame-fsico-pdf>. Acesso em: 14 maio 2020.

NISHIDA, Silvia. **Sentido da visão**. Botucatu: Apostila do curso de fisiologia - Departamento de Fisiologia, 2012. 85 p. Disponível em: https://www.biologia.bio.br/curso/1%C2%BA%20per%C3%ADodo%20Facioplac/08.sntido_visao.pdf. Acesso em: 16 maio 2020.

OLIVEIRA, Carlos; NETO, Manoel. Instituto Federal de Educação, Ciência e Tecnologia da Bahia. **Bengala Digital e Boné Digital à Serviço da Pessoa com Deficiência Visual**, Salvador, 2015.

ORLANDINI, Guilherme. **Desenvolvimento de Aplicativos Baseados em Técnicas de Visão Computacional para Robô Móvel Autônomo**. 2012. Dissertação (Mestrado em Ciência da Computação) - Faculdade de Ciências Exatas e da Natureza, Universidade Metodista de Piracicaba, Piracicaba, 2012. Disponível em: http://www.xbot.com.br/wp-content/uploads/2012/10/Disserta%C3%A7%C3%A3o_Guilherme_V4_revisado_pos_defesa-2_corrigido_final.pdf. Acesso em: 28 maio 2020.

PASSOS, Erinaldo; NETO, Antônio; LEMAIRE, Thierry. Caderno de física da UEFS. **Comportamento ótico do olho humano e suas Ametropias**, Feira de Santana, n. 6, 2008. Disponível em: <http://dfisweb.uefs.br/caderno/vol6n12/Olhohumano.pdf>. Acesso em: 21 maio 2020.

PATSKO, Luís. Maxwell Bohr. **Tutorial: Aplicações, Funcionamento e Utilização de Sensores**, [s. l.], 2006. Disponível em: https://www.maxwellbohr.com.br/downloads/robotica/mec1000_kdr5000/tutorial_eletronica_-_aplicacoes_e_funcionamento_de_sensores.pdf. Acesso em: 1 jun. 2020.

PENIDO, Édilus; TRINDADE, Ronaldo. **Microcontroladores**. Ouro Preto: Instituto Federal de Educação, Ciência e Tecnologia, 2013. Disponível em: http://estudio01.proj.ufsm.br/cadernos/ifmg/tecnico_automacao_industrial/microcontroladores.pdf. Acesso em: 30 maio 2020.

PONTI, Moacir; COSTA, Gabriel. **Como funciona o Deep Learning**. 1. ed. [S. l.]: SBC, 2017. Disponível em: https://sites.icmc.usp.br/moacir/papers/Ponti_Costa_Como-funciona-o-Deep-Learning_2017.pdf. Acesso em: 2 jun. 2020.

RAHIM, Taliha. **Desenvolvimento de um protótipo para auxílio no deslocamento de deficientes visuais**. 2017. Trabalho de conclusão de curso (Graduação em Tecnologias da informação e comunicação) - Universidade Federal de Santa Catarina, Araranguá, 2017. Disponível em: <https://repositorio.ufsc.br/bitstream/handle/123456789/181868/TCC%20Taliha%20Rahim.pdf?sequence=1>. Acesso em: 4 jun. 2020.

REDMON, Joseph. **Darknet: Redes Neurais de Código Aberto em C**. 2016. Disponível em: <http://pjreddie.com/darknet/>

RODRIGUES, Ana. **Imagem de fundo ocular: Análise comparativa de técnicas e de novas tecnologias no diagnóstico oftalmológico da retina**: Aplicação Clínica. 2016. Dissertação (Mestrado em Engenharia Biomédica) - Faculdade de Ciência e Tecnologia, Universidade de Coimbra, Coimbra, Portugal, 2016. Disponível em: https://estudogeral.sib.uc.pt/bitstream/10316/33007/1/TESE_MULTICOLOR_FFINAL_Filipa%20Rodrigues.pdf. Acesso em: 26 maio 2020.

ROSA, João. **Fundamentos da Inteligência Artificial**. Rio de Janeiro: Livros Técnicos e Científicos - LTC, 2011. Disponível em: <http://walderson.com/2011-2/IA/FIA.pdf>. Acesso em: 31 maio 2020.

ROZENFELD, Henrique *et al.* **Gestão de Desenvolvimento de Produtos**: Uma Referência para a Melhoria do Processo. 1. ed. São Paulo: Saraiva, 2006.

SALOMÃO, Solange. Universidade Federal de São Paulo. **Desenvolvimento da acuidade visual de grades**, São Paulo, 2007. Disponível em: <https://www.scielo.br/pdf/pusp/v18n2/v18n2a05.pdf>. Acesso em: 8 jun. 2020.

SANTOS, Aline. **Tecnologia assistiva para pessoas com deficiência visual**: Avaliação da eficiência de dispositivos para mobilidade pessoal. 2019. Dissertação (Mestrado em Design) - Faculdade de Arquitetura, Artes e Comunicação, Universidade Estadual Paulista, Bauru, 2019. Disponível em: https://repositorio.unesp.br/bitstream/handle/11449/181124/santos_adp_me_bauru.pdf?sequence=3&isAllowed=y. Acesso em: 5 jun. 2020.

SANTOS, Elidiane; NASCIMENTO, Flávia. Ciência e Tecnologia da Bahia: Instituto Federal de Educação. **Uma Solução integrada para Auxílio na Mobilidade de Deficientes Visuais em Ambientes Internos**, [s. l.], 2015. Disponível em: <http://www.prpgi.ifba.edu.br/wp-content/uploads/Elidiane-Pereira-dos-Santos.pdf>. Acesso em: 29 maio 2020.

SILVA, Clara; GOMES, Werley; JÚNIOR, Agenor. Revista Interdisciplinar da PUC Minas no Barreiro. **O direito à acessibilidade do deficiente visual à luz da lei**

brasileira de inclusão da pessoa com deficiência, Belo Horizonte, v. 7, ed. 14, 2017.

SILVA, Kátia. Universidade Federal de Minas Gerais. **Acessibilidade para deficientes visuais: um estudo nos museus de belo horizonte**, Minas Gerais, 2013. Disponível em: https://pdfs.semanticscholar.org/6933/d6f90d327835074051d42d46ec6e69ae9ab6.pdf?_ga=2.81902870.1037593241.1591773716-696053835.1591223538. Acesso em: 27 maio 2020.

SOUZA, Magali; FERREIRA, Luiz. Revista Observatório. **A tecnologia assistiva cões-guia no Brasil**: Uma ação do programa viver sem limite, [s. l.], v. 4, n. 3, 2018. Disponível em: <https://doi.org/10.20873/uft.2447-4266.2018v4n3p307>. Acesso em: 20 maio 2020.

STEBEL, Sérgio; SATO, Gilson. EL 71D: Universidade Tecnológica do Paraná. **Introdução à Engenharia Mecatrônica**, Paraná, 2017. Disponível em: http://paginapessoal.utfpr.edu.br/sato/material-didatico/engenharia-mecatronica/material-da-disciplina-de-introducao-a-engenharia-mecatronica/20170409_microprocessadores.pdf. Acesso em: 7 jun. 2020.

VICARI, Rosa. **Tendências em inteligência artificial na educação no período de 2017 a 2030**. Brasília: Confederação Nacional da Indústria. Sumário executivo, 2018. Disponível em: <http://www2.fiescnet.com.br/web/uploads/recursos/d1dbf03635c1ad8ad3607190f17c9a19.pdf>. Acesso em: 3 jun. 2020.

VIEIRA, Marcelo. **Histórico dos computadores**, São Carlos, 2018. Disponível em: <http://iris.sel.eesc.usp.br/sel415m/Aula%201%20-%20Historico%20dos%20Computadores.pdf>. Acesso em: 1 jun. 2020.

WENDLING, Marcelo. **Sensores**. Guaratinguetá: UNESP - Universidade Estadual Paulista, 2010. Disponível em: <https://docplayer.com.br/3707502-Sensores-unesp-prof-marcelo-wendling-2010-versao-2-0-universidade-estadual-paulista.html>. Acesso em: 2 jun. 2020.

ZAQUEU, Livia. **Política inclusiva educacional I**. São Luís: Universidade Federal do Maranhão, 2012. Disponível em: https://books.google.com.br/books?id=9dTuaQAAQBAJ&pg=PA95&hl=pt-BR&source=gbs_selected_pages&cad=3#v=onepage&q&f=false. Acesso em: 8 jun. 2020.

GOMES, José. Departamento de Oftalmologia - Universidade Federal de São Paulo. **Anatomia e fisiologia ocular**, São Paulo, 2019. Disponível em: https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0004-27491999000500004&lng=pt&nrm=iso&tlng=pt. Acesso em: 19 maio 2020.

APÊNDICE 1 – CÓDIGO COMPLETO

```
#Importa as dependências necessárias
import cv2
from cv2 import cv2
import time
import numpy as np
import speech_recognition as sr
from gtts import gTTS
from playsound import playsound

#Estabelece variação das cores do retângulo de detecção
COLORS = [(0,255,255),(255,255,0),(0,255,0),(255,0,0)]

#Carrega as classes de identificação
class_names = []
with open("classes.names", "r") as f:
    class_names = [cname.strip() for cname in f.readlines()]

#Define a câmera utilizada para detecção
cap = cv2.VideoCapture(0)

#Carrega a rede, podendo ser leve ou pesada
net = cv2.dnn.readNet("yolov4.weights", "yolov4.cfg")
#net = cv2.dnn.readNet("yolov4-tiny.weights", "yolov4-tiny.cfg")

#Define os parâmetros da rede
model = cv2.dnn_DetectionModel(net)
model.setInputParams(size=(416,416),scale=1/255)
num = 0

#Emissão de áudio INICIAL
```

```
audio = 'SEJA BEM-VINDO AO SISTEMA INTELIGENTE DE AUXÍLIO À
DEFICIENTES VISUAIS. ESTE É UM PROJETO EM PRIMEIRA VERSÃO E
CONTINUARÁ EM DESENVOLVIMENTO A FIM DE TORNAR A AUTONOMIA
DESTE GRUPO PATOLÓGICO CADA VEZ MAIS EFICIENTE DENTRO DO
ÂMBITO SOCIAL. VAMOS LÁ?: '
```

```
tts = gTTS(audio,lang='pt-br')
```

```
num = num + 1
```

```
arquivo = 'audios/arquivo'+ 'INICIO' + ".mp3"
```

```
tts.save(arquivo)
```

```
playsound(arquivo)
```

```
#Colhendo os frames do vídeo para detecção com base no padrão da rede
```

```
while True:
```

```
    #Segregação dos frames
```

```
    ret, frame = cap.read()
```

```
    try:
```

```
        frame = cv2.resize(frame, (416, 416), interpolation=cv2.INTER_AREA)
```

```
    except:
```

```
        break
```

```
#Início da contagem do tempo levado para detecção em milisegundos
```

```
start = time.time()
```

```
#Define o que foi identificado
```

```
classes, scores, boxes = model.detect(frame, 0.1, 0.2)
```

```
#Fim da contagem do tempo levado para detecção em milisegundos
```

```
end = time.time()
```

```
#Percorre todas as detecções
```

```
for(classid, score, box) in zip(classes, scores, boxes):
```

```

if score*100>55:
    #Gera uma cor para a classe definida
    color = COLORS[int(classid)%len(COLORS)]

    #Busca o nome da classe pelo ID dentro do arquivo classes.name, calcula
    acurácia e apresenta em label
    label = f"{class_names[classid[0]]: {score*100}%"

    #Desenha o retângulo de rastreo de detecção
    cv2.rectangle(frame, box, color, 2)

    #Escreve o nome da classe em cima do retângulo do objeto
    cv2.putText(frame, label, (box[0], box[1]-10),
cv2.FONT_HERSHEY_SIMPLEX,0.5, color, 2)

    #Emissão de áudio do que foi detectado
    audio = class_names[classid[0]]
    tts = gTTS(audio,lang='pt-br')
    num = num + 1
    arquivo = 'audios/arquivo'+ str(num) + ".mp3"
    tts.save(arquivo)
    playsound(arquivo)

    #Calcula o FPS
    fps_label = f"FPS: {round((1.0/(end-start)),2)}"

    #Escreve o FPS na imagem
    cv2.putText(frame, fps_label, (0,25), cv2.FONT_HERSHEY_SIMPLEX, 1,
(0,0,0), 5)
    cv2.putText(frame, fps_label, (0,25), cv2.FONT_HERSHEY_SIMPLEX, 1,
(0,255,0), 3)

```

```
#Mostra a imagem detectada com as escritas do objeto detectado, do FPS,  
#da acurácia, da detecção e o retângulo de rastreo  
cv2.imshow("Detecção", frame)  
  
#Aguarda resposta de processamento  
if cv2.waitKey(1) & 0xFF == ord('q'):  
    break  
  
#Libera a câmera e fecha todas as janelas  
cap.release()  
cv2.destroyAllWindows()
```


APÊNDICE 2 – NOMES DAS CLASSES IDENTIFICÁVEIS

Pessoa

Bicicleta

Carro

Moto

Aviao

Onibus

Trem

Caminhao

Barco

Semaforo

Hidrante

Sinal de pare

Parquimetro

Banco

Passaro

Gato

Cachorro

Cavalo

Ovelha

Vaca

Elefante

Suporte

Zebra

Girafa

Mochila

Guarda-chuva

Bolsa

Gravata

Mala

Frisbee

Esquis

Snowboard

Bola de esportes

Pipa

Taco de beisebol

Luva de baseball

Skate

Prancha de surfe

Raquete de tenis

Garrafa

Copo de vinho

xicara

Garfo

Faca

Colher

Tigela

Banana

Maca

Sanduiche

Laranja

Brocolis

Cenoura

Cachorro quente

Pizza

Rosquinha

Bolo

Cadeira

Sofa

Vaso de planta

Cama

Mesa de jantar

Banheiro

Tv/monitor

Computador portátil

Mouse

Controle remoto

Teclado

Celular

Microondas

Forno

Torradeira

Afundar

Frigorifico

Livro

Relógio

Vaso

Tesoura

Urso Teddy

Secador de cabelo

Escova de dente