

UNIVERSIDADE DE RIBEIRÃO PRETO
CENTRO DE CIÊNCIAS EXATAS, NATURAIS E TECNOLÓGICAS
ENGENHARIA DE COMPUTAÇÃO

VITOR ANDRADE THOMAZINI

**IDENTIFICAÇÃO BIOMÉTRICA PARA ANÁLISE DA
MARCHA HUMANA**

Ribeirão Preto
2017

VITOR ANDRADE THOMAZINI

IDENTIFICAÇÃO BIOMÉTRICA PARA ANÁLISE DA MARÇA HUMANA

Monografia apresentada ao Centro de Ciências Exatas, Naturais e Tecnológicas da Universidade de Ribeirão Preto para obtenção do grau de bacharel em Engenharia de Computação.

Orientador: Prof. Dr. Rodrigo de Oliveira Plotze

Ribeirão Preto
2017

Ficha catalográfica preparada pelo Centro de Processamento Técnico da
Biblioteca Central da UNAERP

- Universidade de Ribeirão Preto -

T465i Thomazini, Vitor Andrade, 1994-
Identificação biométrica para análise da Marcha Humana /
Vitor Andrade Thomazini. - - Ribeirão Preto, 2017.
43 f.: il. color.

Orientador: Prof. Dr. Rodrigo de Oliveira Plotze.

Monografia (graduação) - Universidade de Ribeirão Preto,
UNAERP, Engenharia da computação. Ribeirão Preto, 2017.

1. Automação Residencial. 2. Biometria. 3. Sistema
computacional. I. Título.

CDD 005.1

FOLHA DE APROVAÇÃO

815486 – Vitor Andrade Thomazini
Identificação Biométrica para Análise da Marcha Humana

Monografia apresentada ao Centro de Ciências Exatas, Naturais e Tecnológicas da Universidade de Ribeirão Preto para obtenção do grau de bacharel em Engenharia de Computação.

Orientador: Prof. Dr. Rodrigo de Oliveira Plotze

Aprovada em: 05/12/2017

Banca Examinadora

Prof. Dr. Rodrigo de Oliveira Plotze
Instituição: Universidade de Ribeirão Preto – UNAERP

Prof. Dr. Milton Faria Junior
Instituição: Universidade de Ribeirão Preto – UNAERP

Prof. Dr. Carlos Eduardo Formigoni
Instituição: Universidade de Ribeirão Preto – UNAERP

*Dedico este trabalho a minha família
por sempre me apoiar independente
das decisões que tive até hoje.*

AGRADECIMENTOS

Agradeço a todos que tiveram paciência e compreensão de meus atos e principalmente conseguiram conviver com minha personalidade. Além disso, sou grato aqueles que me apoiaram e ajudaram quando precisava de ajuda.

RESUMO

THOMAZINI, V. A. **Identificação Biométrica para Análise da Marcha Humana**. 2017. 43 p. Monografia (Graduação em Engenharia de Computação) – Centro de Ciências Exatas, Naturais e Tecnológicas, Universidade de Ribeirão Preto. Ribeirão Preto - SP, 2017.

Este trabalho tem como objetivo o desenvolvimento de um sistema computacional com a finalidade de analisar a caminhada de uma pessoa, podendo identifica-la por meio de imagens binarizadas, este procedimento é possível, porque cada pessoa tem formas diferentes de equilíbrio e controle de queda ao deslocar o corpo. Com isso, o desenvolvimento teve como base a teoria da *Gait Energy Image* que analisa o acumulo de energia durante o período de caminhada, com o diferencial nos procedimentos que antecedem a conservação da energia realizando a variação do deslocamento entre um *frame* e seu estado posterior, em seguida efetuando a somatória dos resultados. Além disso, a etapa que consiste em identificar as características, fez uso da rede neural convolucional. O processo de validação foi realizado por meio da matriz de confusão utilizando como parâmetros os resultados obtidos na rede neural, por meio destes, foi obtido um ótimo resultado no processo de identificação dos indivíduos.

Palavras chave: processamento de imagem, biometria, reconhecimento da marcha humana, sistema computacional.

ABSTRACT

THOMAZINI, V. A. **Biometric Identification for Analysis of the Human Gait**. 2017. 43 p. Monography (Undergraduate Degree in Computer Engineering) – Center of Exact Sciences, Natural and Technological, Universidade de Ribeirão Preto. Ribeirão Preto - SP, 2017.

This work has the objective of developing a computational system with the purpose of analyzing the walk of a person, being able to identify it by means of binarized images, this procedure is possible, because each person has different ways of balance and tumbling control when move the body. With this, the development was based on the theory of Gait Energy Image which analyzes the energy accumulation during the walking period, with the differential in procedures that precede the conservation of energy realizing the variation of the displacement between a frame and its later state, after this effecting the sum of the results. Besides that, the step that consists of identifying the characteristics, used the convolutional neural network. The validation process was performed through the confusion matrix using as parameters the results obtained in the neural network, by means of these, obtained great result in the process of human identification.

Keywords: image processing, biometrics, human gait recognition, computational system.

LISTA DE FIGURAS

Figura 1: Operações e um sistema biométrico, tendo como exemplo a impressão digital.....	12
Figura 2: Classificação das impressões digitais	13
Figura 3: Características da face.....	13
Figura 4: Procedimentos para detecção de íris	14
Figura 5: Representação das fases do ciclo de marcha.....	16
Figura 6: O gráfico apresenta um perfil de usuário comparado com outros perfis	16
Figura 7: Representação das subfases do apoio.....	17
Figura 8: Representação das subfases do balanço	18
Figura 9: Os números representam o desvio do corpo, em que eles não ocorrem simultaneamente	19
Figura 10: Banda de espectro eletromagnético.....	21
Figura 11: A - Imagem original, B - Imagem aplicado uma equalização completa e C Imagem aplicado equalização parcial	22
Figura 12: Aplicação de A – algoritmo de esqueletização, B – centroide e C – assinatura tendo como referência a centroide.....	25
Figura 13: Exemplo de duas pessoas caminhando, e o resultado do GEI no quadro mais a direita.	25
Figura 14: Representação visual do classificador por distância mínima, em que a centroide é o ponto em destaque	27
Figura 15: Representação de uma rede neural com uma camada oculta	28
Figura 16: Deslocamento do filtro no mapa de entrada.....	29
Figura 17: Estrutura de uma rede neural convolucional.....	29
Figura 18: A – Visão frontal 0°; B – Visão de 36°; C – Visão de 90°; D-Visão de 126°.....	30
Figura 19: Distribuição do conjunto de dados.....	31
Figura 20: Estrutura das etapas do projeto	32
Figura 21: Amostra de uma imagem com a função de Centralizar aplicada.....	33
Figura 22: Lógica realizada no processamento do conjunto de imagens	34
Figura 23: Resultado obtido com o processamento de imagem	34
Figura 24: Comparação dos resultados obtidos após o processamento de imagem.	35
Figura 25: Representação do modelo da rede neural convolucional	36
Figura 26: Resultados dos acertos obtidos na matriz de confusão	37

LISTA DE EQUAÇÕES

Equação 1: Distância euclidiana.....	24
Equação 2: Distância city-block.....	24
Equação 3: Distância chessboard	24
Equação 4: <i>Gait Energy Image</i>	25
Equação 5: Modelo matemático para cálculo do espectro de movimento	32

LISTA DE SIGLAS

ADC	<i>Analog Digital Converter</i>
CCD	<i>Charge-Coupled Device</i>
CMOS	<i>Complementary Metal-Oxide Semiconductor</i>
EER	<i>Equal Error Rate</i>
GEI	<i>Gait Energy Image</i>
RGB	<i>Red Green Blue</i>

SUMÁRIO

1	INTRODUÇÃO.....	9
1.1	OBJETIVO GERAL	9
1.2	OBJETIVOS ESPECÍFICOS	10
1.3	ESTRUTURA DA MONOGRAFIA	10
2	REVISÃO DE LITERATURA	11
2.1	BIOMETRIA	11
2.1.1	Impressão Digital.....	12
2.1.2	Característica da Face	13
2.1.3	Íris	14
2.1.4	Voz	14
2.1.5	Marcha Humana	15
2.2	RECONHECIMENTO DE MARCHA HUMANA	15
2.2.1	Fases da Marcha	16
2.2.2	Desvio da Pelve	18
2.3	VISÃO COMPUTACIONAL.....	20
2.3.1	Aquisição da Imagem	20
2.3.2	Processamento da Imagem	21
2.3.3	Análise da Imagem	24
2.4	RECONHECIMENTO DE PADRÕES	26
2.4.1	Classificador por Distância Mínima	26
2.4.2	Rede Neural – Multilayer Perceptron (MLP)	27
2.4.3	Rede Neural Convolucional ou ConvNet	28
3	MATERIAIS E MÉTODOS.....	30
3.1	Conjunto de dados	30
3.2	Requisitos Não-Funcionais	31
3.3	Arquitetura	31
4	RESULTADOS	33
5	CONCLUSÃO.....	38
	REFERÊNCIAS	39
	APÊNDICE A – CÓDIGO DO MODELO DA REDE NEURAL CONVOLUCIONAL CRIADO UTILIZANDO O MÓDULO TENSORFLOW	43

1 INTRODUÇÃO

Com o avanço tecnológico em todas as áreas da ciência e, a grande quantidade de informações que são geradas, atualmente, por qualquer dispositivo, surgiu a necessidade de implementar processos para proteger as informações. Com base nesse contexto a tecnologia biométrica tem sido empregada na proteção de informações e ambientes, e tem como finalidade identificar pessoas baseadas em características físicas e comportamentais. No entanto, a biometria traz consigo diversas questões, como por exemplo, a invasão de privacidade de um indivíduo e, além disso, há aqueles que são mal informados ou não conhecem a tecnologia citada (BORGES, et. al, 2011).

Entre as tecnologias biométricas é possível encontrar: sistema de reconhecimento de íris, sistema de reconhecimento de digitais e sistemas de reconhecimento por face. Além disso, existe uma área de estudo denominada biometria comportamental, que contempla o estudo da marcha humana (WANG; GENG, 2010).

A marcha humana tem como vantagem a possibilidade de analisar e/ou identificar indivíduos a uma longa distância se comparado com métodos de reconhecimento de face, íris, digitais e/ou técnicas de biometria comportamental. Tal vantagem é possível por analisar a silhueta do indivíduo, com isso, suas dimensões em uma imagem são enfatizadas até que a distância entre a câmera e o indivíduo distorça a imagem a um nível que a distribuição dos *pixels* da silhueta não seja possível extrair características suficientes para identificar o indivíduo.

Segundo Boulgouris, Hatzinakos e Plataniotis (2005), com a análise da marcha é possível caracterizar o movimento para reconhecer e identificar indivíduos, além de reconhecer e detectar atividades suspeitas. Outra possível aplicação está relacionada a área da fisioterapia, em que pela análise da marcha pode-se identificar distúrbios no modo de andar, podendo se originar de patologia, de lesões ou de enfermidades e, também, pode ser empregada no acompanhamento pós-cirúrgico (MAGEE, 2010).

1.1 OBJETIVO GERAL

Desenvolver um sistema computacional de identificação biométrica de um indivíduo por meio da análise de sua caminhada.

1.2 OBJETIVOS ESPECÍFICOS

Os objetivos específicos do estudo são:

- Propor um algoritmo de análise do perfil da marcha humana;
- Analisar marcha humana pela composição de *frame*;
- Treinar e balancear uma rede neural artificial supervisionada.

1.3 ESTRUTURA DA MONOGRAFIA

Essa monografia está dividida em cinco capítulos, sendo que no primeiro capítulo apresenta-se a Introdução, os Objetivos Geral e Específicos e a Estrutura da Monografia.

No segundo capítulo apresenta-se a Revisão de Literatura que apresenta os conceitos de biometria, reconhecimento de marcha humana, visão computacional e reconhecimento de padrões.

No terceiro capítulo há a descrição dos Materiais e Métodos, em que se detalhar quais os materiais e métodos foram utilizados para o desenvolvimento da parte experimental da monografia.

No quarto capítulo encontra-se a descrição dos Resultados durante o processo de desenvolvimento. Já o quinto capítulo é apresentado a Conclusão do projeto.

E, por último, são apresentadas as referências bibliográficas utilizadas para o desenvolvimento da monografia.

2 REVISÃO DE LITERATURA

Neste capítulo é apresentada a revisão de literatura que contempla a contextualização sobre biometria, o reconhecimento da marcha humana e a visão computacional, com os processos de aquisição, processamento e análise da imagem. Além disso, também abordará do conceito de reconhecimento de padrões.

2.1 BIOMETRIA

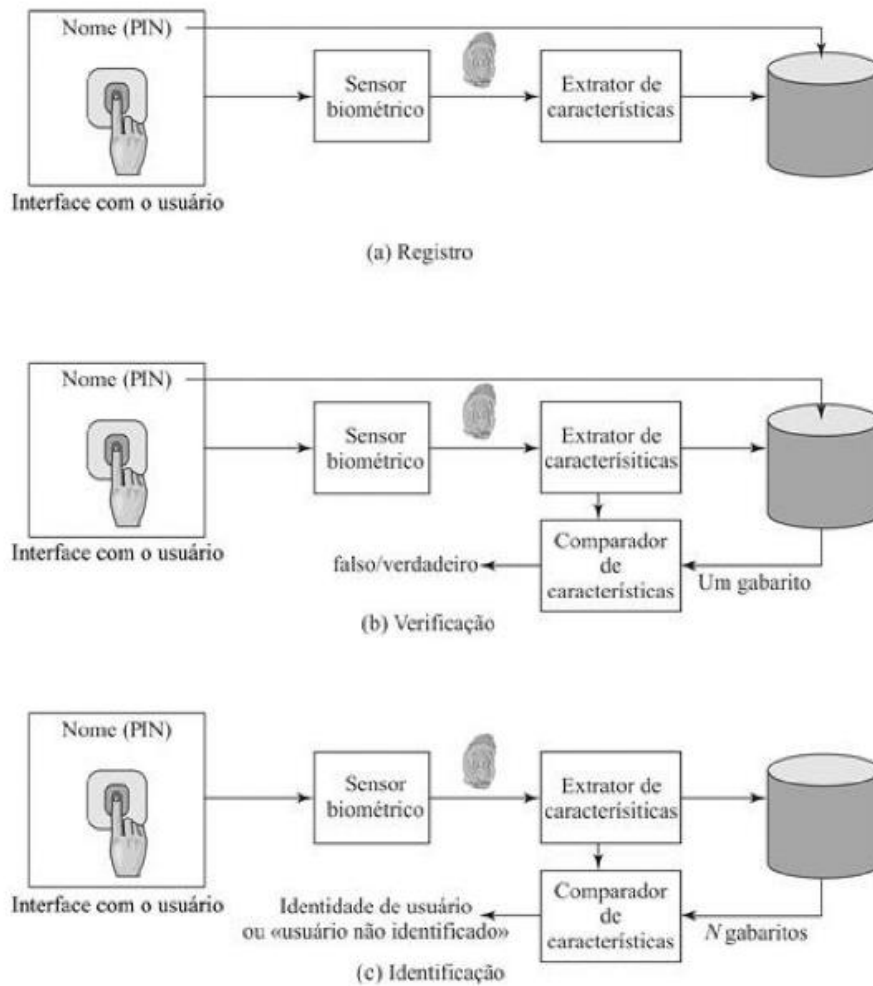
A biometria é uma ciência que consiste em reconhecer e autenticar um indivíduo por meio de suas características físicas, tais como a impressão digital, a estrutura da mão, as características da face, o padrão de íris e a retina (SANTOS, 2007). As características físicas são relativamente estáveis, basicamente imutáveis ou com pouca variação com o envelhecimento (BRANQUINHO et al., 2014).

No contexto da biometria também é realizado um estudo dos padrões de comportamento que incorpora como fator de análise a dimensão temporal, além disso, diferente da biometria convencional pode ser coletada de uma forma não invasiva ou até mesmo sem conhecimento dos usuários (WANG; GENG, 2010). Outro fator que o padrão de comportamento incorpora é a possibilidade de explorar elementos relacionados ao psicológico podendo, assim, extrair novas características, no entanto, isto também afetará a identificação do usuário, pois se o mesmo estiver alterado psicologicamente seu comportamento será afetado (BRANQUINHO et al., 2014).

Alguns exemplos deste tipo de biometria são: a verificação de assinatura, a dinâmica de teclas, o reconhecimento por voz (WANG; GENG, 2010) e também reconhecimento da marcha humana (REVETT, 2008).

Segundo Stallings e Brown (2014), um sistema biométrico é dividido em três operações: registro, verificação e identificação, tais como ilustrados na Figura 1. Nas operações o registro consiste em armazenar as características biométricas obtidas do usuário e um identificador, uma senha, para que, posteriormente, o sistema possa realizar uma verificação do indivíduo e/ou identifica-lo com as características obtidas por meio da biometria.

Figura 1: Operações e um sistema biométrico, tendo como exemplo a impressão digital

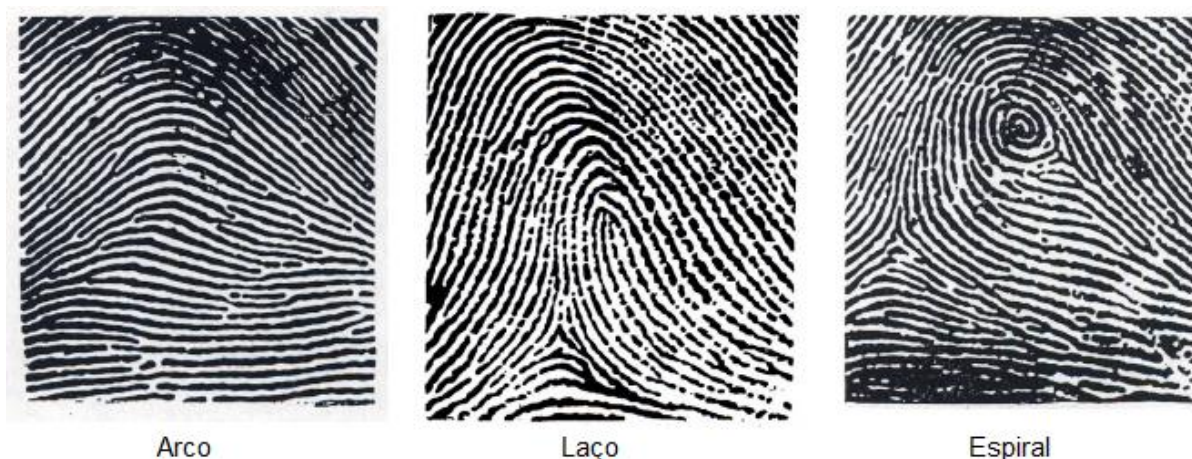


Fonte: STALLINGS; BROWN (2014, p. s/d)

2.1.1 Impressão Digital

Este tipo de reconhecimento pode ser feito por meio de três dispositivos: ótico, em que a luz é refletida sobre o dedo; capacitivo, realiza a medida do calor que sai da digital; e ultrassônico, enviando sinais sonoros e analisando seu retorno (SANTOS, 2007). As impressões digitais obtidas podem ser classificadas em arco, laço ou espiral, conforme demonstrado na Figura 2, para que na sequência sejam analisadas e, assim, identificar o indivíduo (US DEPT. OF JUSTICE, 1984).

Figura 2: Classificação das impressões digitais



Fonte: Adaptação de US DEPT. OF JUSTICE (1984, p. 6).

Este tipo de reconhecimento pode ser utilizado em atividades de forense, auxiliando na identificação de criminosos ou pessoas vinculadas ao mesmo, além disso, tem um vínculo comercial podendo identificar funcionários de uma organização (WAYMAN et al., 2005).

2.1.2 Característica da Face

O reconhecimento facial é um sistema não intrusivo (SANTOS,2007), em que a análise é obtida por meio de imagens frontais do rosto e com iluminação uniforme, suas características são abstraídas da relação das distâncias entre olhos, nariz e boca, demonstrado na Figura 3, podendo ser aplicado em interface humano-computador, verificação de identidade, controle parentais e monitoramento do paciente (SZELISKI, 2010).

Figura 3: Características da face



Fonte: SZELISKI (2010, p. 238).

2.1.3 Íris

A leitura da íris é um dos sistemas mais precisos do mundo, pois ela possui um padrão único para cada pessoa e que nunca é alterado, tal padrão é definido pela coloração do olho, no entanto, é uma das opções de biometria com maior custo (SANTOS, 2007). Embora, no ano de 2017 tal sistema já está sendo integrado com *smarthphones*, Galaxy Note 7 para desbloquear o dispositivo (SAMSUNG, 2017). Na Figura 4 são representadas as etapas do sistema de íris do *smarthphone*.

Figura 4: Procedimentos para detecção de íris



Fonte: SAMSUNG (2017).

2.1.4 Voz

A verificação por voz tem como base a estrutura do espectro de amplitude do indivíduo, identificando seus picos, no qual a extração de características é aplicada com a amplitude do sinal normalizada que ainda é decomposta em vários canais de frequência, sendo um logaritmo do sinal de entrada por meio da Transformada de Fourier a característica frequentemente extraída (WANG; GENG, 2010). Além disso, a verificação pode ser classificada de duas formas, textos dependentes - a frase ou o texto deve estar armazenado para ser utilizado como parâmetro de comparação e textos independentes - são identificados

padrões em um conjunto de falas gravadas para serem comparados com a que está sendo identificada (REVETT, 2008).

2.1.5 Marcha Humana

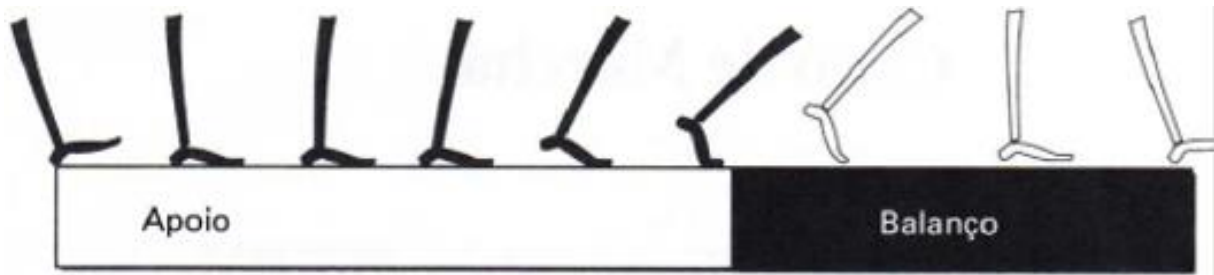
Na análise da marcha humana existe pequenas diferenças decorrentes de cada indivíduo, tais diferenças correspondem a 5% ou menos da *Equal Error Rate* (EER) (REVETT, 2008). O EER é uma medida de verificação para desempenho global, no qual, seu valor deve ser zero, mas na biometria este erro deve ser o menor valor possível tendendo a zero devido à variedade de dados encontrado (CHIN; PONG, 1998). A análise da marcha, pode ser realizada por meio de sensores e também por visão computacional, em que a visão computacional é normalmente utilizada em locais que há um grande volume de pessoa, já os sensores, por exemplo acelerômetros, são adequados para autenticação de usuários, por adquirir dados dinâmicos, medindo a aceleração em três direções ortogonais (REVETT, 2008).

No entanto, a marcha está sujeita as significativas variações com base nas mudanças do peso corporal da pessoa, na ondulação durante a gravidez, lesões nas extremidades ou no cérebro ou devido a intoxicação (WANG; GENG, 2010).

2.2 RECONHECIMENTO DE MARCHA HUMANA

A marcha humana pode ser definida como sequência de movimentos repetitivos dos membros inferiores associados com a oscilação dos membros superiores, tronco e cabeça, fazendo com que o corpo mova para frente mantendo uma postura estável (OTTOBONI; FONTES; FUKUJIMA, 2002). Nesta sequência de movimentos um membro inferior é utilizado como apoio, enquanto o outro é responsável por avançar para uma nova posição, após isso o membro inverte suas funções até que o primeiro membro inferior que foi utilizado como apoio volte a tocar o chão, tal processo é denominado ciclo de marcha (MAGEE, 2010). Este ciclo contém a fase de apoio, representa 60 a 65% do ciclo, e a fase de balanço, representando 35 a 45% do ciclo como demonstrado na Figura 5, contendo também dois períodos de apoio duplo e um período de apoio de um único dos membros inferiores (MAGEE, 2010).

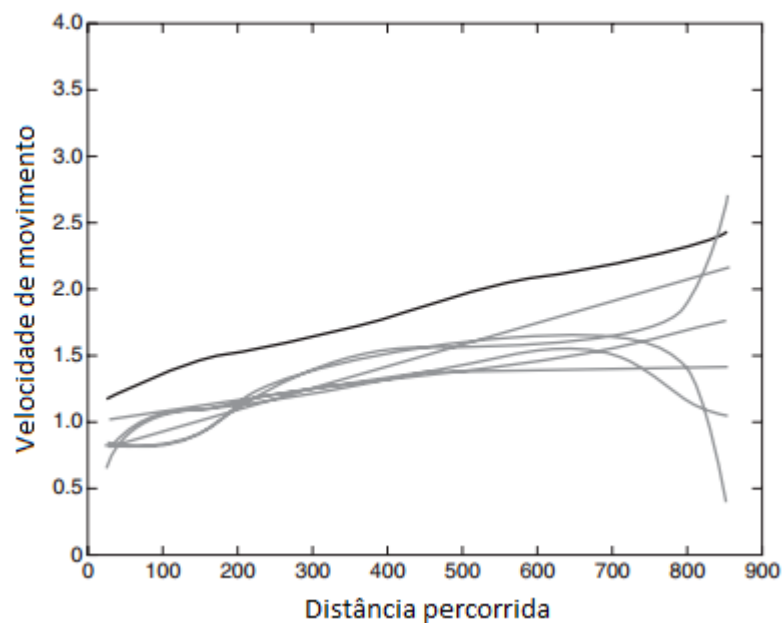
Figura 5: Representação das fases do ciclo de marcha



Fonte: ARANTES (2010, p. 26)

Andriacchi et al. (1977) apud Arantes (2010) relatara que a duração das fases de apoio e balanço variam de acordo com a velocidade da marcha do indivíduo. Com uma velocidade de 80m/min, velocidade normal, os períodos de apoio e balanço representam 62% e 38% do ciclo de marcha (ARANTES, 2010). Na Figura 6 é demonstrada a relação entre velocidade e distância percorrida de alguns perfis, tal que a velocidade foi obtida por meio de sensores.

Figura 6: O gráfico apresenta um perfil de usuário comparado com outros perfis



Fonte: REVETT (2008, p. 7).

2.2.1 Fases da Marcha

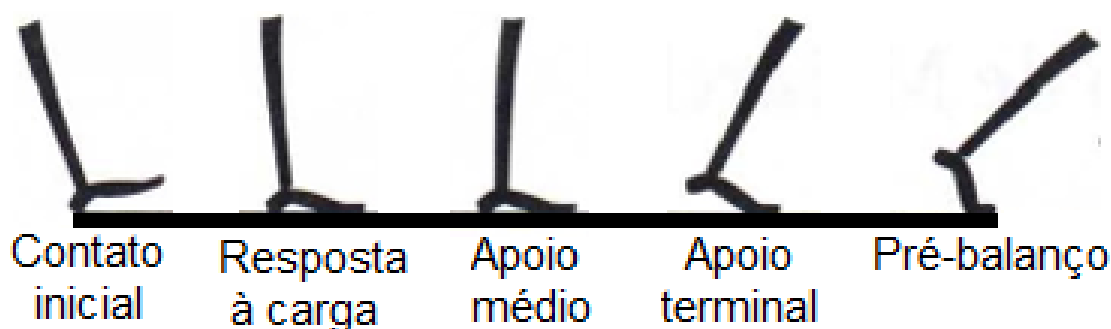
A fase de apoio consiste na sustentação do peso do corpo por um único membro inferior, sendo subdividida em contato inicial, reposta de carga, apoio médio, apoio terminal e pré-balanço (Figura 7). O contato inicial sucede no instante que o pé se posiciona no solo,

para suportar o peso do corpo, enquanto o outro começa a perder contato com a superfície, entretanto, ambos os pés ainda estão apoiados ao solo caracterizando um apoio duplo (MAGEE, 2010).

Ainda conforme Magee (2010), a resposta a carga e o apoio médio acontece após o contato inicial, quando um único membro sustenta peso do corpo e mantém o mesmo em equilíbrio, além disso deve existir estabilidade lateral para que o outro membro avance. Por fim as fases de apoio terminal e pré-balanço ocorrem quando o membro que estava apoiando transfere o peso no corpo no instante em que o membro que estava em balanço posiciona-se em contato com o solo, deixando a posição dos membros invertidos se comparado com o estado inicial, isto também caracteriza um apoio duplo.

Esta também pode ser descrita utilizando a primeira lei de Newton, um objeto que detém velocidade constante e aceleração nula estará em inércia até que uma força o desestabilize (YOUNG; FREEDMAN, 2008), mas é preciso desenvolver uma força para superar a inércia, portanto, quanto maior o peso do corpo, maior a força para que o corpo adquira aceleração suficiente, movimentando o corpo (ASENCIO et al., 2001).

Figura 7: Representação das subfases do apoio.



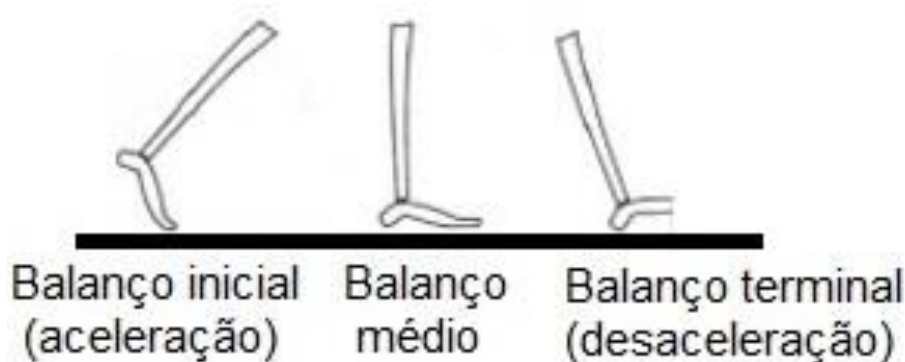
Fonte: Adaptação de ARANTES (2010, p. 26); MAGEE (2010, p. 943).

Já a fase de balanço é caracterizado por um membro não sustentar o peso do corpo e, além disso, mover-se para frente, esta também é subdividida em balanço inicial, balanço médio, balanço terminal, representado na Figura 8. O balanço inicial é caracterizado no momento que o pé não está sustentando o peso do corpo e o joelho é flexionado para que o membro na fase de balanço acelere para frente, esta subfase também é denominada aceleração. Após a aceleração ocorre o balanço médio, quando o membro na fase de balanço fica adjacente com o membro oposto na subfase de apoio médio, para que posteriormente

sucedá-se o balanço terminal ou desaceleração em que o membro na fase de balanço prepara-se, desacelerando a velocidade do corpo, para entrar em contato com a superfície (MAGEE, 2010).

Outra forma de representar este processo pode ser dada com a segunda lei de Newton, uma força resultante é aplicada em um corpo, este acelera e/ou desacelera de acordo com a direção e o sentido da aceleração e/ou força (YOUNG; FREEDMAN, 2008), o corpo continuará acelerando até que exista uma força oposta, suficiente, para desacelerá-lo (ASENCIO et al., 2001).

Figura 8: Representação das subfases do balanço



Fonte: Adaptação de ARRANTES (2010, p. 26); MAGEE (2010, p. 943).

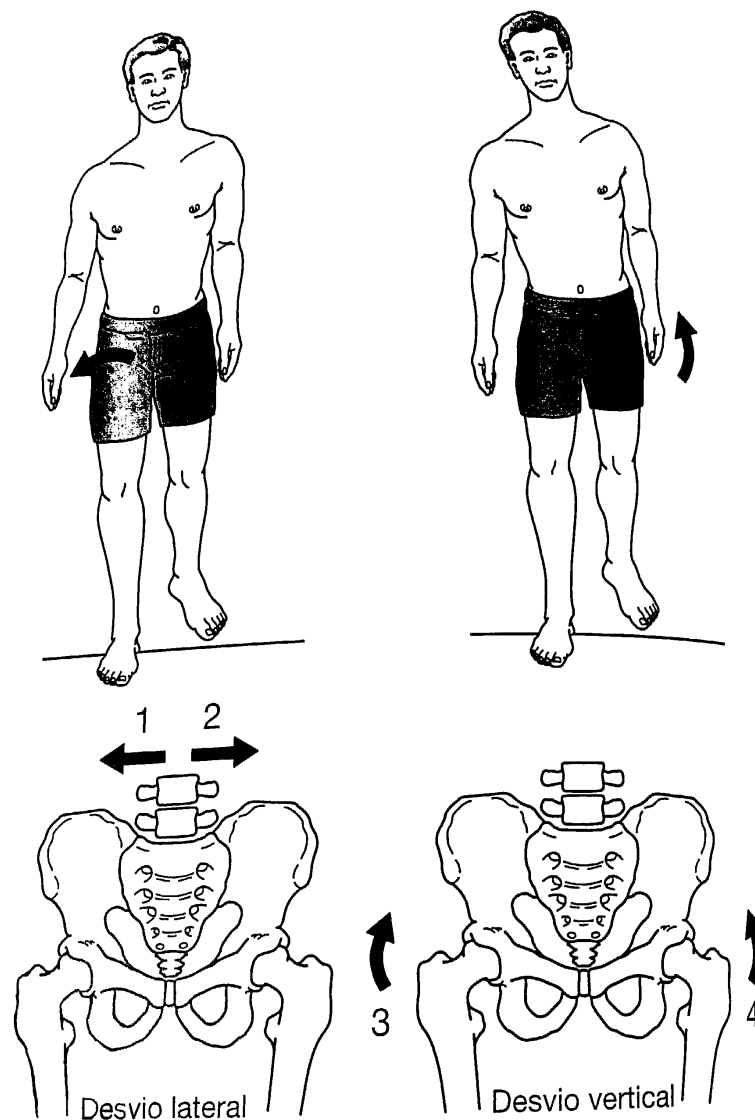
Ainda assim, para que o corpo se mantenha equilibrado, em ambas as fases, a terceira lei de Newton é aplicada, um corpo A exerce uma força/ ação sobre um corpo B, logo, o corpo B exerce uma força/reação sobre corpo A com sentido oposto a força de A (YOUNG; FREEDMAN, 2008), isto é, a fricção do apoio sobre a superfície causa uma reação que auxilia o equilíbrio e propulsão (ASENCIO et al., 2001).

2.2.2 Desvio da Pelve

No instante que a pessoa tem a intenção de levantar o pé direito, por exemplo, nele é repousado o peso do corpo para que posteriormente seja transferido ao pé esquerdo e assim o pé direito esteja em condições para deixar o solo, como consequência da repetição deste movimento para ambos os membros o corpo inclina-se para frente e assim o membro que oscilará irá acelerar (ASENCIO et al., 2001). Neste processo, o desvio lateral da pelve tem como papel manter o centro de massa do corpo sobre o membro inferior de apoio, além disso

a variação do desvio é relativa a distância entre os pés, ou seja, caso esteja afastado o desvio será maior. Já o desvio vertical da pelve consiste em limitar o posicionamento do centro de gravidade do corpo tanto para cima ou quanto para baixo, mantendo a altura do corpo sempre inferior quando o indivíduo está realizando a marcha, se comparado com a posição em que está apoiado com ambos os pés (MAGEE, 2010). A Figura 9 mostra o efeito do desvio da pelve sobre o corpo de uma pessoa.

Figura 9: Os números representam o desvio do corpo, em que eles não ocorrem simultaneamente



Fonte: MAGEE (2010, p. 946).

2.3 VISÃO COMPUTACIONAL

A visão computacional pode ser definida como soluções que proporcionam a máquina um sistema visual similar ao do ser humano, projetando a imagem e interpretando seu significado, embora tal nível ainda não pode ser alcançado por englobar fatores (movimentação, intensidade luminosa, reflexão, entre outros), que assimilados criam a concepção da visão, gerando assim, dificuldade para o desenvolvimento de métodos matemáticos com o intuito de aproximar a visão humana com a visão computacional (SZELISKI, 2010).

Neste contexto, uma imagem é uma representação bidimensional de uma projeção real formada por pontos, denominados elementos pictóricos ou *pixels* em que neles são contidas informações referentes ao posicionamento, coordenadas x e y , e a intensidade ou nível de cinza (GONZALEZ; WOODS, 2010). Além disso, um *pixel* contempla com informações das colorações como por exemplo o sistema de cores *Red Green Blue* – RGB (SZELISKI, 2010).

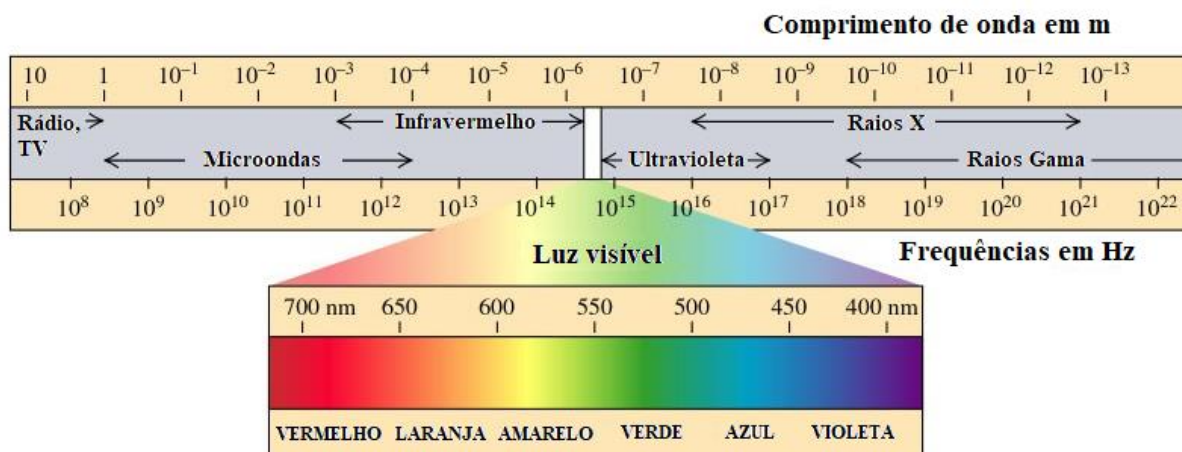
2.3.1 Aquisição da Imagem

Segundo Gonzalez e Woods (2010), uma imagem digital é formada por meio de um conjunto de sensores formando uma estrutura bidimensional que capta a energia de iluminação para que posteriormente seja transformada em tensão mediante a combinação da energia elétrica de entrada e do material do sensor, em que o sensor seja sensível a uma determinada banda de espectro eletromagnético, este representado na Figura 10, logo, para a visão humana identificar e/ou reconhecer a coloração de uma imagem, a banda deve variar de 0,43 μ m(violeta) até 0,79 μ m(vermelho), tal faixa denomina-se luz cromática.

No mercado existe dois tipos de sensores para câmeras convencionais, os *Charge-Coupled Device* (CCD) e os *Complementary Metal-Oxide Semiconductor* (CMOS), ambos têm a finalidade de acumular a carga para produzir uma imagem, no entanto, os CCD geram dados analógicos que representam a coloração, por isso precisam de um conversor *Analog Digital Converter* (ADC), em comparação o sensor CMOS já fornece os dados no formato digital (SZELISKI, 2010).

Na visão computacional existe duas classes de imagens, as estáticas e as dinâmicas, em que a dinâmica é composta por uma sequência de imagens adquiridas por um curto intervalo de tempo criando um aspecto visual uniforme, já estática é uma única imagem (COSTA; CERSAR JUNIOR, 2001).

Figura 10: Banda de espectro eletromagnético



Fonte: GOOGLE (2017).

2.3.2 Processamento da Imagem

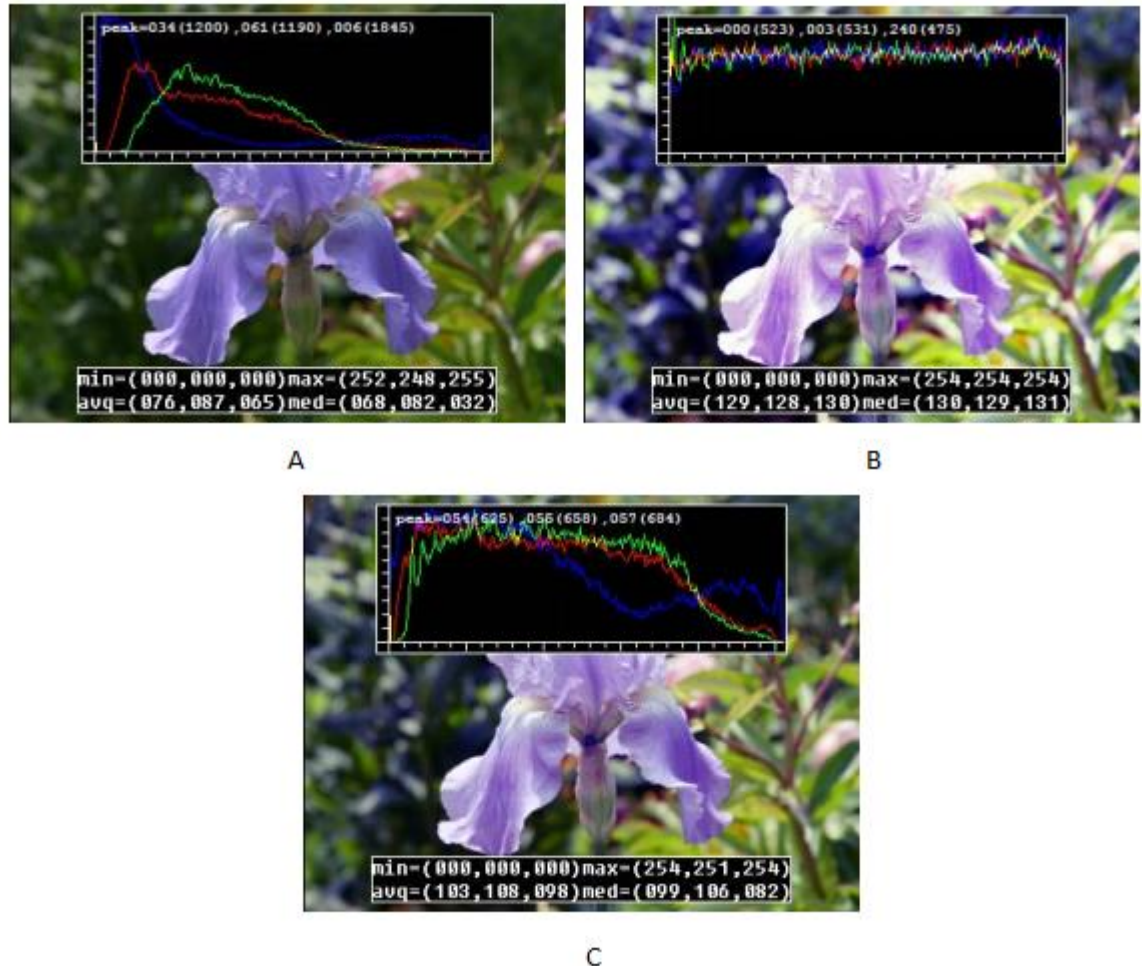
O processamento de uma imagem engloba um conjunto de algoritmos com a finalidade de reestruturar a imagem para que, posteriormente, seja analisada e retirada suas características, além disso, ele pode ser dividido em duas categorias, o pré-processamento e a segmentação, sendo o realce e a restauração relacionadas ao pré-processamento e a segmentação tem como objetivo destacar os objetos de interesse na imagem, normalmente esta deve estar binarizadas, isto é, em preto ou branco, note que, caso a segmentação não esteja adequada ao problema que está sendo analisado, o mesmo não poderá ser solucionado ou será ineficiente em alguns aspectos (GONZALEZ; WOODS, 2010). Neste contexto, existe algumas técnicas de processamento, tais como: histograma, filtros, detecção de bordas e transformada de distâncias.

2.3.2.1 Histograma

Um histograma é o mapeamento de uma imagem por meio da contagem de cores equivalentes dos canais (RGB) e/ou pela variação de tons de cinza de cada pixel representado por um gráfico, demonstrado na Figura 11, embora, normalmente seja utilizado tons de cinzas, neste esta transformação é realizado uma média entre os valores referentes aos canais de cada *pixel* (COSTA; CERSAR JUNIOR, 2001). Por meio de um histograma, é possível identificar níveis concentrados de tons de cinza que prejudica a visibilidade da imagem, podendo assim, aumentar seu contraste por meio da equalização do histograma, isto é, a

coloração dos pixels é recalculada para que os tons de cinza sejam redistribuídos sem muitas variações (SZELISKI, 2010).

Figura 11: A - Imagem original, B - Imagem aplicado uma equalização completa e C Imagem aplicado equalização parcial



Fonte: Adaptação de SZELISKI (2010, p. 102).

2.3.2.2 Filtros

Os filtros são normalmente matrizes aplicadas por meio da convolução de matriz que percorrem cada pixel da imagem independentemente da resolução, embora, os *pixels* da borda da imagem não devem ser considerados nesta convolução, por não conter todos os pixels de sua vizinhança (COSTA; CERSAR JUNIOR, 2001). Neste contexto, vizinhança são os *pixels* que estão em torno de um *pixel*, seja horizontal e vertical, denominada vizinhança-4 ou considerando também suas diagonais, vizinhança-8, embora, a aplicação de filtro utiliza vizinhança-8 (GONZALEZ; WOODS, 2010).

Os filtros passa-baixa, são utilizados para realizar a suavização e reduzir ruídos da imagem, tal processo pode ser efetuado por meio do filtro da média, neste cada valor da coloração do *pixel* (tons de cinza, por exemplo) é substituído pela média de sua vizinhança, além deste, existe o filtro Gaussiano, tendo o mesmo intuito que o filtro da média, no entanto, o *pixel* central do filtro detém o maior valor e os valores dos vizinhos vão reduzido de forma gradativa (COSTA; CERSAR JUNIOR, 2001).

2.3.2.3 Detecção de Bordas

Uma borda pode ser detectada por meio da mudança de coloração da imagem, tais mudanças são detectadas através de derivadas de primeira ordem e/ ou segunda ordem, ou seja, na derivada de primeira ordem quando seus resultados forem zero, implica que são áreas de intensidade constante e quando seus resultados forem diferentes de zero, denota o início de um degrau, de uma rampa e até ao longo de uma rampa, já as derivadas de segunda ordem seus resultados são diferente de zero somente no início e no fim de uma rampa, no entanto, a transição de claro e escuro, ou vice e versa, são definidas por meio do sinal de seu resultado, de claro para escuro o sinal é positivo e o inverso é negativo. Outra diferença entre as derivadas dá-se pela espessura da borda gerada, em que as derivadas de primeira ordem produzem bordas mais largas e as de segunda ordem bordas mais finas, destacando pequenos ruídos isolados na imagem. (GONZALEZ; WOODS, 2010). Alguns exemplos de algoritmos de bordas que utilizam derivadas de primeira ordem são de Sobel, Prewitt e Kirsch (KEKRE; GHARGE, 2010). Já o de segunda ordem, pode ser utilizado o algoritmo de Laplace (COSTA; CERSAR JUNIOR, 2001).

2.3.2.4 Transformada de Distâncias

A transformada de distância é uma das operações mais básicas de processamento de imagem, pois ela pode ser utilizada para cálculo da distância mínima entre dois pontos, esqueletização e ainda em operações morfológicas, dilatação e erosão por exemplo. Neste contexto, algumas equações podem ser utilizadas para calcular a distância, tais como distância euclidiana (1), distância *city-block* (2) e a distância *chessboard* (3), dentre elas a distância Euclidiana possui maior precisão, por não representar a distância em *pixel*, no entanto ela exige maior processamento que as demais (ZAMPIROLL, 2003).

$$d_E(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad (1)$$

$$d_4(x,y) = |x_1 - y_1| + |x_2 - y_2| \quad (2)$$

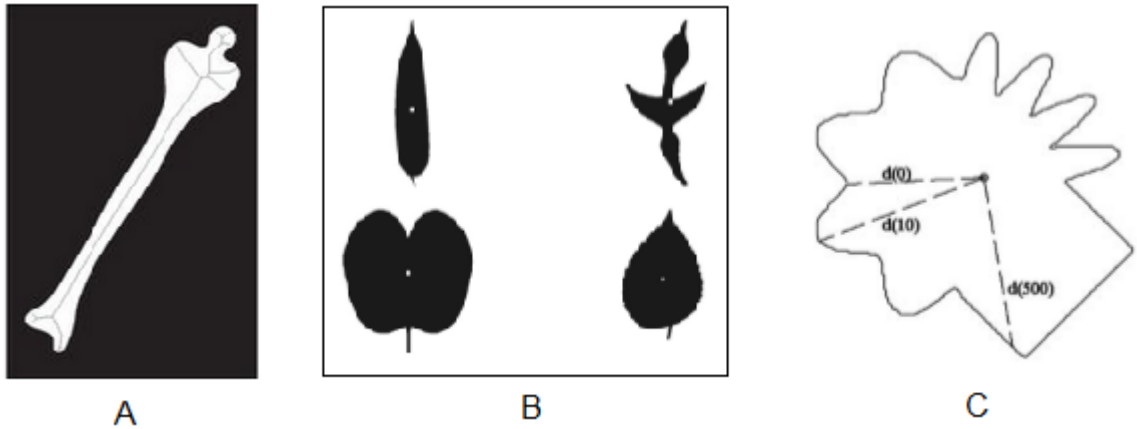
$$d_8(x,y) = \text{MAX}(|x_1 - y_1|, |x_2 - y_2|) \quad (3)$$

2.3.3 Análise da Imagem

A análise consiste em técnicas computacionais, normalmente baseada em métodos estatístico, com o objetivo de extrair características de uma imagem, tais como o perímetro, a área, a centroide ou centro de massa, a distância máxima e/ou mínima entre a centroide e a borda, o maior diâmetro entre dois pontos da imagem e as assinaturas, esta consiste na representação unidimensional de uma borda podendo assim ser representadas por meio dos eixos x e y separadamente criando perspectivas de cada eixo, mediante a distância entre a centroide e as coordenadas da borda e a curvatura ocorrida na borda (COSTA; CERSAR JUNIOR, 2001). Outro aspecto que pode ser explorado é a representação estrutural da região, este processo pode ser realizado por meio de algoritmos de afinamento ou esqueletização, no entanto, para que a imagem mantenha sua estrutura estes algoritmos devem respeitar alguns critérios sendo eles: manter os pontos extremos, manter a conectividade entre os pontos e não gerar erosões excessivas na região (GONZALEZ; WOODS, 2010). A Figura 12 demonstra a aplicação dos algoritmos em imagens.

Uma das abordagens na análise da marcha é proposta por Wang et al. (2003), nela é realizado o cálculo do centroide da silhueta, e por meio deste é feita uma análise de assinatura tendo como base o centroide no sentido anti-horário, para que isso seja possível antes de representar a imagem em uma forma unidimensional é necessário aplicar um algoritmo de borda, após realizar estes processos é realizado a média das projeções da representação unidimensional.

Figura 12: Aplicação de A – algoritmo de esqueletização, B – centroide e C – assinatura tendo como referência a centroide



Fonte: Adaptação de GONZALEZ, WOODS (2010, p. 537); COSTA; CERSAR JUNIOR (2001, p. 408, 410).

Outra abordagem, segundo Han e Bhanu (2016), consiste na técnica denominada *Gait Energy Image* (GEI), por meio dela é gerada uma imagem em tons de cinza que indica o acúmulo de energia em uma sequência de imagens, sendo os pixels com maior intensidade responsável por representar regiões com maior frequência do movimento. A intensidade do pixel é definida pela equação (4), sendo N o número de quadro no ciclo completo e t o número do quadro na sequência, já x e y são coordenadas de uma imagem 2D. Os resultados desta técnica está representado na Figura 13

$$G(x,y) = \frac{1}{N} \sum_{i=1}^N B_t(x,y) \quad (4)$$

Figura 13: Exemplo de duas pessoas caminhando, e o resultado do GEI no quadro mais a direita.



Fonte: HAN; BHANU (2016)

2.4 RECONHECIMENTO DE PADRÕES

Para uma máquina identificar e reconhecer padrões é uma tarefa complexa, embora pareça simples se comparado com as habilidades do ser humano, tal complexidade dá-se por não existir um algoritmo genérico para identificar padrões de um objeto (RUSSELL; NORVING, 2013), sabendo que, um padrão pode ser definido como um vetor de características que “rotulariza” um objeto, logo, um conjunto de padrões é denominado classe de padrões (GONZALEZ; WOODS, 2010).

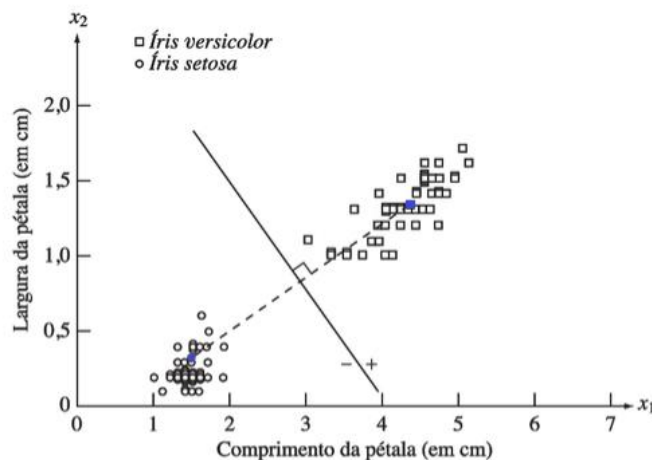
Neste contexto, um sistema inteligente pode ser classificado como supervisionado, em que um conjunto de valores são inseridos e previamente classificados para o sistema reconhece-lo, e a outra classificação trata-se de uma não-supervisionada, nesta o sistema deve identificar o padrão por meio de relações das características que foram inseridas (RUSSELL; NORVING, 2013).

Uma das aplicações que são abordados inteligência artificial é a criação de sistema especialistas, isto é, implementa-se um sistema que possa fazer algo similar a um profissional, por exemplo, treinar um sistema para identificar tumores ao mesmo nível que um médico faria (LUGER, 2004).

2.4.1 Classificador por Distância Mínima

No classificador por distância mínima a primeira etapa consiste em identificar um parâmetro inicial para cada classe que irá compor o sistema, tal padrão é calculado por meio da média dos valores associados a classe. Obtendo o valor médio ou centroide realiza-se o cálculo da distância euclidiana para cada item da classe, caso a distância e a centroide seja pequena o padrão comparado pertence aquela classe (GONZALEZ; WOODS, 2010). A Figura 14 representa uma possível situação que este classificador é utilizado.

Figura 14: Representação visual do classificador por distância mínima, em que a centroide é o ponto em destaque



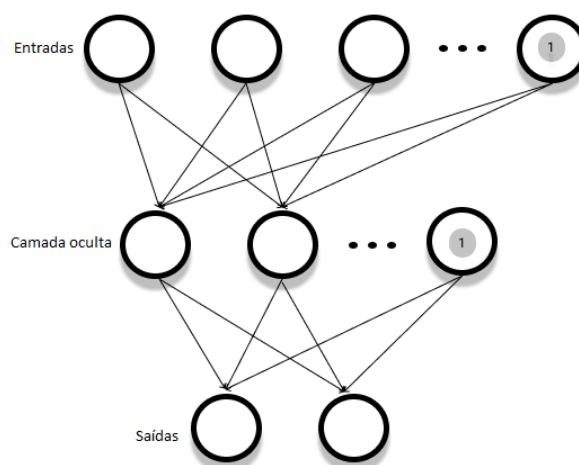
Fonte: Adaptação de GONZALEZ; WOODS (2010, p. 572).

2.4.2 Rede Neural – Multilayer Perceptron (MLP)

Uma rede neural é um sistema que tem como base o funcionamento de um neurônio e a forma como são conectados no cérebro, tal sistema é implementado por meio de componentes simples que através de um processamento e/ou treinamento suas conexões são ajustadas pelos algoritmos de retro propagação (LUGER, 2004).

Este sistema tem como estrutura uma determinada quantidade de entradas e sua respectiva resposta ou saída, sendo que, entre as entradas e saídas são adicionados camadas ocultas com uma certa quantidade de neurônios, em que cada neurônio das camadas ocultas estão relacionados com todos os neurônios das camadas que o antecede, note que, esta lógica também é aplicada à saída, criando uma estrutura similar as conexões do cérebro, além disso, no vetor de entrada e nas camadas ocultas são adicionados o rótulo de +1, denominado unidade de polarização, correspondente ao termo de intercepção. Outro fator importante para o funcionamento de uma rede neural são os ativadores que corresponde a como o neurônio irá processar ou associar as informações que chegam nele, esta operação pode ser realizada por meio da função de sigmoide, da função da tangente hiperbólica ou da função linear retificada (STANFORD, 2017). A Figura 15 representa uma possível distribuição de rede neural.

Figura 15: Representação de uma rede neural com uma camada oculta

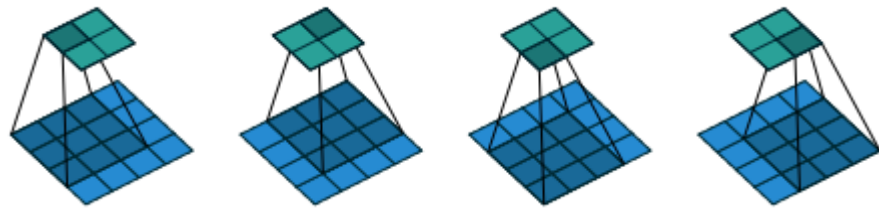


Fonte: Autoria própria.

2.4.3 Rede Neural Convolucional ou ConvNet

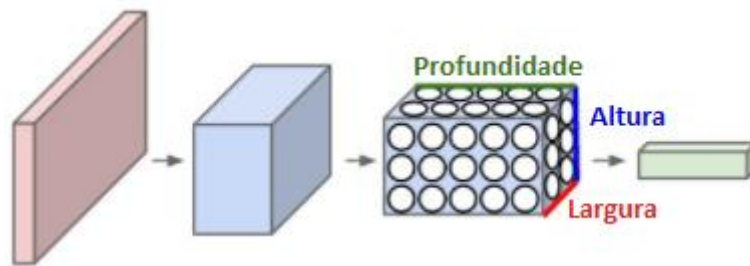
As redes neurais convolucionais são semelhantes a redes MLP, pois são constituídas de neurônios e possuem pesos e tendências aprendíveis. No entanto, diferem em sua estrutura e especialização. As ConvNet são especializadas na análise e de imagens e/ou sinais, trabalhando com estruturas tridimensionais (largura, altura e profundidade) no caso de imagens. Todavia, diferente de uma MLP, suas camadas contêm três tipos de configuração: (i) a camada convolucional, consiste na aplicação de filtros, que são obtidos por meio do deslocamento de uma matriz (*kernel*) na largura, altura e volume das entradas, relacionando os pixels por meio da transformação linear e posteriormente aplicando a função de ativação, conforme demonstrado na Figura 16; (ii) a camada *pooling* ou agrupamento, realiza um processo similar a convolução, contudo, não é aplicado a transformação linear ou ativação, mas técnicas de agrupamento de valores, tais como média ou valor máximo; (iii) e a camada *fully-connected* ou classificador, nesta, o funcionamento é o mesmo de uma MLP, pois aqui, os dados já estarão agrupados em uma única dimensão (CS231N, 2017). A Figura 17 ilustra a estrutura de uma rede convolucional.

Figura 16: Deslocamento do filtro no mapa de entrada



Fonte: DUMOULIN; VISIN (2016).

Figura 17: Estrutura de uma rede neural convolucional



Fonte: CS231N (2017).

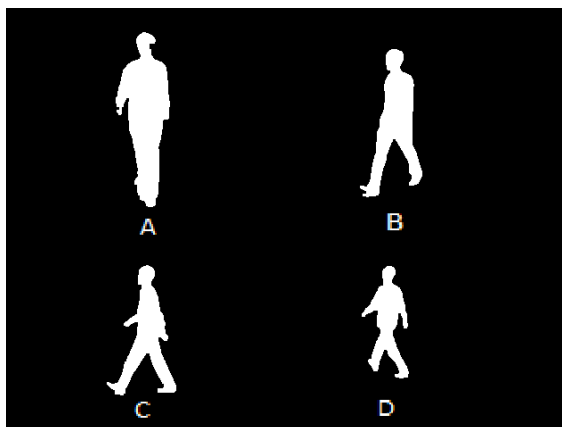
3 MATERIAIS E MÉTODOS

Neste capítulo é apresentada as técnicas de realizadas durante o processo de desenvolvimento do sistema para análise, sendo subdividida em conjunto de dados, requisitos não-funcionais e arquitetura.

3.1 Conjunto de dados

O desenvolvimento do sistema fez uso da base de dados de marcha da Casia e o conjunto B (CASIA, 2017). Este conjunto de imagens, com resolução 320×240 , foi criada tendo como base 124 indivíduos e 11 visões diferentes, contemplando, 6 sequencias para caminhada natural, 2 sequencias com vestuário mais volumoso e 2 sequencias carregando algum tipo de bolsa, tais dados foram coletados para cada indivíduo (YU; TAN; TAN, 2006). Além disso, as imagens já estão segmentadas/binarizadas destacando a silhueta dos indivíduos. A Figura 18 representa algumas visões disponíveis na base de dados.

Figura 18: A – Visão frontal 0°; B – Visão de 36°; C – Visão de 90°; D-Visão de 126°

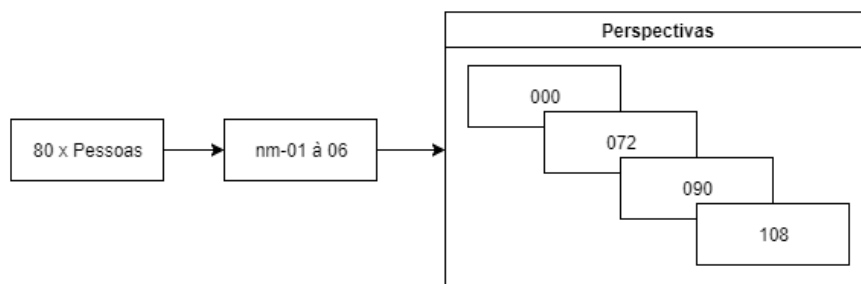


Fonte: Adaptação CASIA (2017).

No entanto, algumas imagens apresentavam imperfeições que poderiam ser significativas, por isso, das 124 amostras foram selecionadas 80 para realização do projeto, além disso, para manter a mesma quantidade de amostras para cada indivíduo, foram selecionadas 4 perspectivas. A redução dos dados, seja na escolha dos indivíduos e em suas perspectivas, ocorreu, pois, algumas amostras não continham imagens que representam uma caminhada completa e/ou não existia imagens para algumas perspectivas. Por fim, foram utilizados apenas as amostras de caminhada normal sem o envolvimento do conjunto com

vestuário volumoso ou carregando algum tipo de bolsa, está redução de dados também se sucedeu por faltar imagens em algumas perspectivas. Note que, mesmo retirando uma quantidade significativa de amostras o conjunto contempla com o total de 1920 amostras e sua distribuição é representada na Figura 19.

Figura 19: Distribuição do conjunto de dados



Fonte: Autoria própria.

3.2 Requisitos Não-Funcionais

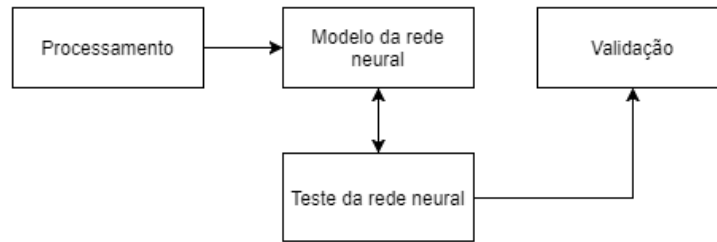
O projeto foi desenvolvido por meio da linguagem de programação Python 3, para tal, o ambiente de desenvolvimento contém o interpretador Python 3.5.2, para ser compatível com o módulo Tensorflow 1.3, biblioteca para computação numérica usando gráficos de fluxo de dados (TENSORFLOW, 2017) e Opencv 3.2, biblioteca focada em recursos para visão computacional e aprendizado de máquina. (OPENCV, 2017). Além destes também foram utilizados os módulos Numpy 1.13, pacote fundamental para computação científica (NUMPY, 2017) e Pandas 0.21.0, biblioteca que oferece estruturas de dados de alto desempenho e fácil de usar e ferramentas de análise de dados (PANDAS, 2017). Para representação e criação dos gráficos foi utilizado o módulo Matplotlib 2.1.0.

Para o desenvolvimento do projeto foi utilizado o editor de texto Atom 1.22.0, por conter *plug-ins* que facilitam a execução do projeto.

3.3 Arquitetura

Com o intuito de manter uma estrutura lógica, esquematizada na Figura 20, durante o desenvolvimento, o projeto está dividido em quatro etapas: o processamento de imagem, a modelagem da rede neural, os testes para redimensionar a rede neural e a validação.

Figura 20: Estrutura das etapas do projeto



Fonte: Autoria própria.

O processamento da imagem, tem como base a técnica de reconhecimento de marcha GEI, mas o que difere é a forma como é obtido o espectro de movimento ou energia, calculada pela equação (5), onde i denota a posição atual na lista e B indica a imagem com suas respectivas coordenadas no plano cartesiano x e y .

$$EM = \sum_{i=0} [B_{i+1}(x, y) - B_i(x, y)] \quad (5)$$

O modelo da rede neural escolhido é denominado rede neural convolucional, pois estes são especializados para realizar o processamento de imagem, sem a necessidade de extrair características por meio de técnicas de processamento de imagem.

Com a finalidade de evitar vícios em uma rede neural, impossibilitando-a de encontrar a melhor solução ou a solução real, alguns algoritmos foram desenvolvidos sendo denominados regularizadores. Um dos regularizados mais utilizados em grandes conjuntos de dados é chamado *dropout*, nele é estipulado um valor que representa a quantidade de conexões que serão anuladas recebendo o valor zero. No entanto, durante as próximas iterações do treinamento, não necessariamente serão as mesmas conexões anuladas, isto é, as conexões zeradas são escolhidas aleatoriamente.

O treinamento, no retro propagação ou *backpropagation*, tem como otimizador o algoritmo de Adam, pois, segundo Kingma e Ba (2015), é um método otimizado estocástico que requer apenas gradientes de primeira ordem com pouca necessidade de memória. Além disso, este algoritmo também é, normalmente, aplicado em treinamento com grande conjunto de dados.

A validação tem como fundamento a matriz de confusão cujo a finalidade é avaliar modelos estatísticos. Já o processo de teste, consiste em executar o modelo da rede e verificar se o resultado é ou não significativo para ser validado.

4 RESULTADOS

O desenvolvimento do espectro de movimento da caminhada contém duas funções essenciais:

- Região de interesse ou *Region Of Interest* (ROI): é extraída por meio da técnica de processamento de imagem denominada *bounding box*. Nela é calculado os quatro pontos mais extremos da figura ou os dois pontos mais extremos interligados por uma diagonal, com o intuito de montar um retângulo que contempla somente a região que contém características relevantes.
- Centralizar: nela é extraída a ROI (O) e, posteriormente, centralizada em outra imagem (D) com dimensão 240x120, com a finalidade de padronizar as proporções das imagens para inserir na rede neural convolucional. Além disso, o processo de centralização tem como referência o centro do retângulo e não o centroide correspondente à imagem. Uma amostra desta função é representada na Figura 21.

Figura 21: Amostra de uma imagem com a função de Centralizar aplicada

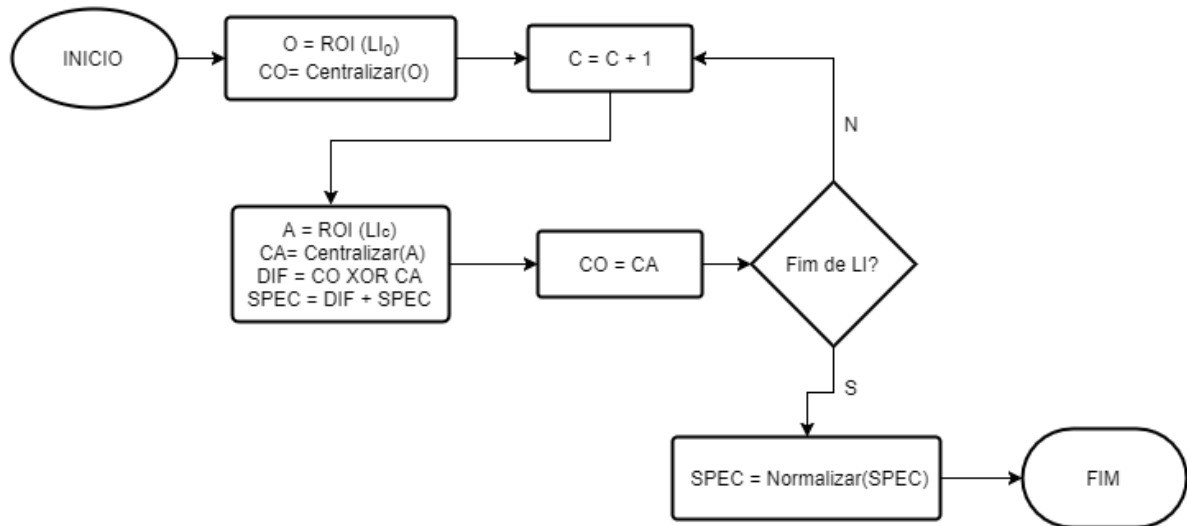


Fonte: Autoria própria.

A lógica, representada na Figura 22, para efetuar o processamento de imagem, consiste em uma lista de frames ou imagens (LI) representando o ciclo da caminhada, tal que, na primeira imagem é aplicado a função de centralização (CO) e armazenada para, posteriormente, ao aplicar a centralização na próxima imagem (CA), efetue a operação lógica XOR entre elas e seu resultado seja somado a uma imagem auxiliar (SPEC), que representa o espectro de movimento da caminha. Após esta primeira fase a imagem CO assume o valor da imagem (CA). Esta lógica será executada até que todas as imagens da lista tenham passado

pelo processo. Ao percorrer toda a lista, o resultado, ilustrado na Figura 23, contido na imagem auxiliar é normalizado com os valores entra 0 e 1.

Figura 22: Lógica realizada no processamento do conjunto de imagens



Fonte: Autoria própria.

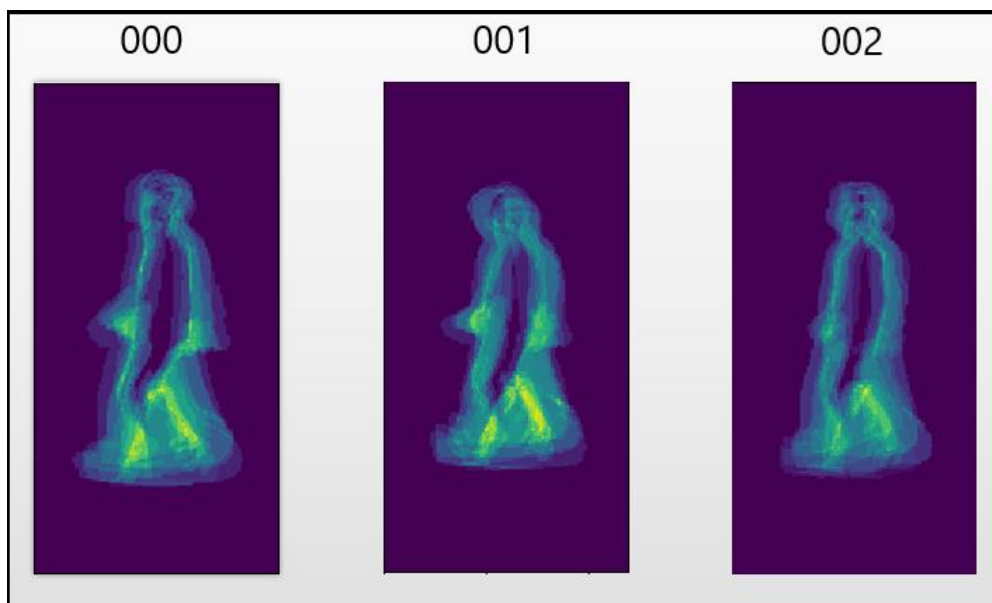
Figura 23: Resultado obtido com o processamento de imagem



Fonte: Autoria própria.

Com a aplicação desta lógica é possibilitou uma comparação visual de algumas amostras de pessoas diferentes, ilustrada na Figura 24, em que o indivíduo 000 desloca seu braço mais que o 002, além disso, o descolamento do tórax e cabeça é menor que o 001, sendo possível observar também as áreas onde estão situados os maiores acúmulos de energia.

Figura 24: Comparação dos resultados obtidos após o processamento de imagem.



Fonte: Autoria própria.

Os dados obtidos foram renomeados para possuir uma característica sequencial, isto é, caso a próxima amostra tenha o *label* 003, mas na anterior o *label* 001 é realizado a renomeação. Esta etapa tem a finalidade manter os *labels* linear, pois isto afeta o resultado da precisão durante os testes da rede, neste projeto a redução da precisão estava próximo a 10%, dependendo da amostra de teste.

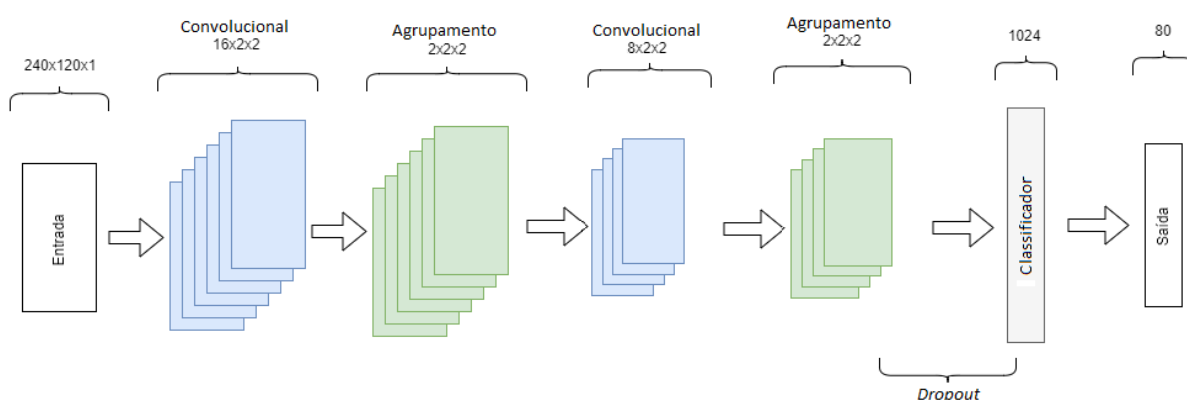
A configuração da rede neural convolucional supervisionada, ilustrada na Figura 25, dispõe de uma entrada com dimensão 240x120x1 (largura x altura x canal de cor) seguidas de duas camadas convolucionais com proporções, respectivamente, 16x2x2 e 8x2x2 (filtros x *kernel*), sendo que cada uma destas, contém uma camada de *pooling*, com dimensões 2x2x2 (passo x *kernel*), interligada em suas respectivas saídas. Além disso, o resultado das convoluções possuem as mesmas dimensões de entrada. Já as camadas de agrupamento, por conter o passo igual à 2 (distancia de um *kernel* a outro durante o deslocamento na imagem), sua saída é a metade da dimensão de entrada.

Com as convoluções definidas, a camada referente ao classificador tem 14400 entradas, este resultado é calculado dividindo a entrada (largura e altura) por 4, pois foi aplicado duas camadas – no caso a de agrupamento – em que o passo era 2, isto faz com que o resultado daquela camada seja dividido pela metade, e posteriormente, a dimensão resultante foi multiplicada por 8, sendo este o valor referente a quantidade de filtros da última camada antes do classificador. O classificador contém uma camada oculta com 1024 nós e interligado com

80 nós de saída. Além disso, entre as entradas e a camada oculta existe um regularizador *dropout*, configurado para anular 40% das conexões.

Já a saída do classificador, foi estipulado este valor por analisar 80 amostras, pois o vetor inicia-se com os valores em zero para identificar o primeiro indivíduo e, posteriormente, é rotulado o valor 1 para o primeiro índice do vetor e o restante recebe o valor de zero, logo, para identificar o próximo indivíduo a posição com o valor 1 é trocada com seu vizinho à direita, e este procedimento é realizado para rotular todas os indivíduos.

Figura 25: Representação do modelo da rede neural convolucional



Fonte: Autoria própria.

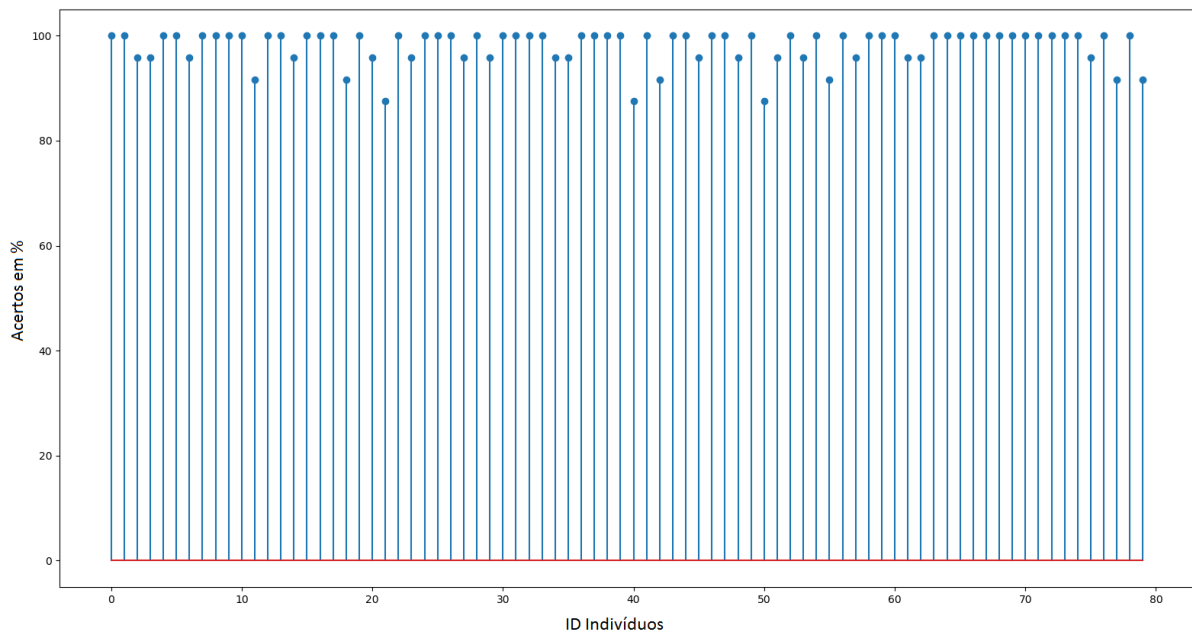
Para efetuar os testes e a validação, foram montados 24 conjuntos de dados diferentes, em que cada um contém 80 amostras para teste e 1840 amostras de treinamento. A divisão destes conjuntos tem como fundamento fazer com que todas as amostras da base de dados fossem submetidas a teste, e a quantidade de conjunto foi definida pois cada pessoa contém 24 amostras de dados diferentes. Note que, as amostras de treinamento foram embaralhadas antes de serem inseridas na rede, para prevenir que aprenda vícios da sequência.

Com os conjuntos definidos, foram realizados testes na rede com apenas um conjunto para avaliar sua precisão (maior) e o erro (menor), e assim modificar a rede até obter a configuração já descrita. Na validação todos os conjuntos foram submetidos ao treinamento e posteriormente, seus respectivos testes. Com isso os valores reais dos *labels* e as previsões dos testes foram armazenados para montar a matriz de confusão e o resultado desta é representado na Figura 26, onde os identificadores iniciam em 0 e finalizam em 79. Note que, o treinamento de cada conjunto foi realizado em redes diferentes, mas com a mesma configuração, além disso, todos os treinamentos foram realizados com 6000 iterações ou

passos e durante o treinamento o conjunto foi dividido em lotes de 100 com o intuito de reduzir o tempo de treino.

Analisando os dados obtidos na matriz de confusão é possível observar que o valor médio das precisões equivale a 97,97%, já a mediana das mesmas equivale a 100%. Além disso, 66,25% dos resultados de cada pessoa tem a precisão de 100%.

Figura 26: Resultados dos acertos obtidos na matriz de confusão



Fonte: Autoria própria.

Comparando os melhores resultados da caminhada normal obtido por o Alotaibi e Mahmood (2017) com o modelo desenvolvido, respectivamente, temos a precisão de 98,3% e 97,97%. Por estes resultados, é possível observar que a diferença é mínima, no entanto, o modelo de Alotaibi e Mahmood (2017) usufruiu das 124 amostras dispostas na base de dados.

5 CONCLUSÃO

O objetivo de desenvolver um sistema computacional de identificação biométrica para análise da caminhada de um indivíduo foi parcialmente concluído, pois os algoritmos não foram estruturados para o uso de pessoas leigas, contudo, o processo de análise obteve um ótimo resultado com precisão de 97,97%, possibilitando a identificação de indivíduos. Neste contexto, aumentar o número de iterações no processo de treinamento possivelmente aumentará, ao menos um pouco, a precisão.

A construção de um algoritmo de análise de perfil da marcha humana foi concretizada, pois já existia um fundamento teórico referente ao acúmulo e dispersão de energia no movimento descrito pela GEI, logo, o algoritmo desenvolvido tem como finalidade de analisar a variação entre um frame e o seguinte, para posteriormente analisar o acúmulo e a dispersão da energia das variações obtidas até finalizar a sequência de *frames*. Outro fator fundamental foi o uso da rede neural convolucional, responsável por extrair as características das amostras obtidas após o processamento da imagem, por isso, foi essencial treinar e balancear uma rede neural para cumprir com o objetivo do trabalho, além disso, com o uso de poucas camadas de convolução e *pooling* na estrutura da rede, possibilitando maior agilidade no processo de treinamento.

Outros algoritmos que podem ser implementados, compõem o critério de segmentação em tempo real das pessoas durante sua caminhada, com o intuito de minimizar problemas gerado nesta etapa, além disso, identificar claramente o ciclo de caminhada para que posteriormente possa ser elaborado um sistema completo de análise e identificação de pessoas por meio de seu caminhar.

REFERÊNCIAS

ALOTAIBI, M.; MAHMOOD, A. Improved Gait Recognition Based on Specialized Deep Convolutional Neural Network. **Computer Vision and Image Understanding**, 2017. 29p.

ARANTES, M. **Método de reconhecimento da marcha humana por meio da fusão das características do movimento global**. 2010. 115p. Tese (Doutorado) - Curso de Programa de Pós-graduação em Engenharia Elétrica. Universidade de São Paulo, São Carlos.

ASENCIO, G. et. al. **A marcha humana, a corrida e o salto**. 1ª ed. Barueri, SP: Manole Ltda, 2001.

BRANQUINHO, M. A. et al. **Segurança de Automação Industrial e SCADA**. 1ª ed. Rio de Janeiro: Elsevier, 2014.

BORGES, D.; SÁ, V. J.; MAGALHÃES, S. T.; SANTOS, H. Study of the Perceptron on the Biometric Technology. **7th International and 4th e-Democracy Joint Conferences, ICGS3/e-Democracy**. Thessaloniki;Grecia, Agosto/2011.

BOULGOURIS, V. N; HATZINAKOS, D.; PLATANOTIS, N. K. Gait Recognition: A challenging signal processing technology for biometric identification. **IEEE Signal Processing Magazine**, v. 22, n.6, p. 78-90, Novembro/2005.

CASIA. **Database**. Disponível em: <http://www.cbsr.ia.ac.cn/users/szheng/?page_id=71>. Acesso em: 17 Maio 2017.

COSTA, L. F.; CESAR JUNIOR, R. M. **Shape classification and analysis: theory and practice**. 2ª ed. Nova York: CRC Press, 2001. 651p.

CS231N. **Convolutional neural networks**. Disponível em: <<https://cs231n.github.io/convolutional-networks/>>. Acesso em: 15 Novembro 2017.

CHIN, R.; PONG, T. **Computer Vision - ACCV'98**. 2ª ed. Berlin: Springer, 1998. 504 p.

DEVORE, J. L. **Probability & statistics for engineering and the sciences**. 8ª ed. Boston: Cengage Learning, 2012. 687 p.

DUMOULIN, V.; VISIN, F. **A guide to convolution arithmetic for deep learning**. 2016. Disponível em: < <https://arxiv.org/abs/1603.07285>>. Acesso em 23 Novembro 2017.

GOOGLE. **Espectro de banda**. Disponível em < http://www.fisica-interessante.com/image-files/cor-espectro_eletromagnetico1.jpg>. Acesso em: 11 Maio 2017.

GONZALEZ, R. C.; WOODS, R. C. **Processamento digital de imagens**. 3ª ed. São Paulo: Pearson Prentice Hall, 2010. 500 p.

HAN, J.; BHANU, B. Individual Recognition Using Gait Energy Image. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 28, n. 2, p. 316-322, Fevereiro/2006.

KEKRE, H. B.; GHARGE, S. M. Image Segmentation using Extended Edge Operator for Mammographic Images. **International Journal on Computer Science and Engineering**, Tamil Nadu, v. 2, n. 4, p.1086-1091, 2010.

KINGMA, D. P; BA, J. L. **Adam: A Method for Stochastic Optimization**. 2014. Disponível em < <https://arxiv.org/abs/1412.6980>>. Acesso em: 23 Novembro 2017.

LUGER, G. F. **Inteligência artificial: estruturas e estratégias para a solução de problemas complexos**. 4ª ed. Porto Alegre: Bookman, 2004. 774p.

MAGEE, J. D. **Avaliação muscoesquelética**. 5ª ed. Barueri, SP: Manole Ltda, 2010, p. 940-971.

NUMPY. **Numpy**. Disponível em: <<http://www.numpy.org/>>. Acesso em: 15 Novembro 2017.

OPENCV. **Definição**. 2017. Disponível em: <<http://opencv.org/about.html>>. Acesso em: 17 Maio 2017.

OTTOBONI, C.; FONTES, V. S.; FUKUJIMA, M. M. Estudo comparativo entre a marcha normal e a de pacientes hemiparéticos por acidente vascular encefálico: aspectos biomecânicos. **Revista de Neurociências**, São Paulo, v.10, n. 1, p.10-16, 2002.

PANDAS. **Python Data Analysis Library**. Disponível em: <<http://pandas.pydata.org/>>. Acesso em: 15 Novembro 2017.

REVETT, K. **Behavioral biometrics**: a remote access approach. 1ª ed. Reino Unido: John Wiley & Sons Ltd, 2008, 244p.

RUSSELL, S.; NORVING, P. **Inteligência Artificial**. 1ª ed. Rio de Janeiro: Elsevier, 2004. 1021p.

SANTOS, A. **Gerenciamento de identidades**. 1ª ed. Rio de Janeiro: Brasport, 2007.

SAMSUNG. **Galaxy Note 7**. 2017. Disponível em: <<https://news.samsung.com/br/de-olho-na-seguranca-o-reconhecimento-de-iris-do-galaxy-note7>>. Acesso em: 05 Maio 2017.

STALLINGS, W.; BROWN, L. **Segurança de computadores**: princípios e práticas. 2ª ed. Rio de Janeiro: Elsevier, 2014.

STANFORD. **Multi-layer neural network**. Disponível em: <<http://ufldl.stanford.edu/tutorial/supervised/MultiLayerNeuralNetworks/>>. Acesso em: 15 Novembro 2017.

SZELISKI, R. **Computer vision**: algorithms and applications. 1ª ed. Nova York. Spring, 2010, 957p.

TENSORFLOW. **Tensorflow**. Disponível em: <<https://www.tensorflow.org/>>. Acesso em: 15 novembro 2017.

US DEPT. OF JUSTICE. **Science of Fingerprints**: Classification and Uses. 1ª ed. Washington, DC: Justice Dept., Federal Bureau of Investigation, 1984.

WANG, L. et al. Silhouette Analysis-Based Gait Recognition for Human Identification. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 25, n. 12, p. 1505-1518, Dezembro/2003.

WANG, L.; GENG, X. **Behavioral Biometrics for Human Identification**: Intelligent Applications. Estados Unidos: Hershey PA, 2010.

WAYMAN, J. et al. **Biometric Systems**: Technology, Design and Performance Evaluation. 1ª ed. Londres: Springer, 2005.

YOUNG, H. D.; FREEDMAN, R. A. **Física I**. 12ª ed. São Paulo: Pearson Prentice Hall, 2008.

YU, S; TAN, D; TAN, T. A Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition. In Proc. of the 18'th **International Conference on Pattern Recognition (ICPR)**. Hong Kong. Agosto/2006.

ZAMPIROLI, F. A. **Transformada de distância por morfologia matemática**. 2003. 139 p. Tese (Doutorado) - Curso de Programa de Pós-graduação em Engenharia Elétrica, Universidade Estadual de Campinas, Campinas.

APÊNDICE A – CÓDIGO DO MODELO DA REDE NEURAL CONVOLUCIONAL CRIADO UTILIZANDO O MÓDULO TENSORFLOW

```

IMG_WIDTH = 240
IMG_HEIGHT = 120
N_CLASSES = 80

def cnn_model(features, labels, mode):
    input_layer = tf.reshape(features["x"], [-1, IMG_WIDTH, IMG_HEIGHT, 1])

    cnet = tf.layers.conv2d(inputs=input_layer, filters=16, kernel_size=[2, 2], padding="same", activation=tf.nn.relu)
    pnet = tf.layers.average_pooling2d(inputs=cnet, pool_size=[2, 2], strides=2)

    cnet = tf.layers.conv2d(inputs=pnet, filters=8, kernel_size=[2, 2], padding="same", activation=tf.nn.relu)
    pnet = tf.layers.average_pooling2d(inputs=cnet, pool_size=[2, 2], strides=2)

    flat = tf.reshape(pnet, [-1, 60*30*8])
    dense = tf.layers.dense(inputs=flat, units=1024, activation=tf.nn.relu)
    dense = tf.layers.dropout(inputs=dense, rate=0.4, training=mode == tf.estimator.ModeKeys.TRAIN)

    logits = tf.layers.dense(inputs=dense, units=N_CLASSES)

    predictions = {
        "classes": tf.argmax(input=logits, axis=1),
        "probabilities": tf.nn.softmax(logits, name="softmax_tensor")
    }

    if mode == tf.estimator.ModeKeys.PREDICT:
        return tf.estimator.EstimatorSpec(mode=mode, predictions=predictions)

    onehot_labels = tf.one_hot(indices=tf.cast(labels, tf.int32), depth=N_CLASSES)
    loss = tf.losses.softmax_cross_entropy(onehot_labels=onehot_labels, logits=logits)

    if mode == tf.estimator.ModeKeys.TRAIN:
        optimizer = tf.train.AdamOptimizer(learning_rate=0.001)
        train_op = optimizer.minimize(loss=loss, global_step=tf.train.get_global_step())
        return tf.estimator.EstimatorSpec(mode=mode, loss=loss, train_op=train_op)

    eval_metric_ops = { "accuracy": tf.metrics.accuracy(labels=labels, predictions=predictions["classes"])}
    return tf.estimator.EstimatorSpec(mode=mode, loss=loss, eval_metric_ops=eval_metric_ops)

```