TUANY DIAS PINHEIRO

Classificação de imagens faciais para o auxílio ao diagnóstico de Autismo

São Paulo

2016

TUANY DIAS PINHEIRO

Classificação de imagens faciais para o auxílio ao diagnóstico de Autismo

Texto de Exame de Qualificação apresentado à Escola de Artes, Ciências e Humanidades da Universidade de São Paulo como parte dos requisitos para obtenção do título de Mestre em Ciências pelo Programa de Pós-graduação em Sistemas de Informação.

Área de concentração: Metodologia e Técnicas da Computação

Orientador: Prof. Drª. Ariane Machado Lima

São Paulo

2016

Texto de Exame de Qualificação de autoria de Tuany Dias Pinheiro, sob o título **"Classificação de imagens faciais para o auxílio ao diagnóstico de Autismo"**, apresentado à Escola de Artes, Ciências e Humanidades da Universidade de São Paulo, como parte dos requisitos para obtenção do título de Mestre em Ciências pelo Programa de Pós-graduação em Sistemas de Informação, na área de concentração Metodologia e Técnicas da Computação, aprovado em ___ de _____ de _____ pela comissão examinadora constituída pelos doutores:

Prof. Dr. _____

Presidente

Instituição: _____

Prof. Dr. _____

Instituição: _____

Prof. Dr. _____

Instituição: _____

Prof. Dr. _____

Instituição: _____

Resumo

O transtorno do espectro autista (TEA) é um transtorno de desenvolvimento que prejudica persistentemente a comunicação e a interação social e causa padrões restritos e repetitivos de comportamento, interesses e atividades no indivíduo. Esses sintomas estão presentes desde o início da infância e limitam ou prejudicam o cotidiano do indivíduo. Contudo, vários fatores impedem que seja possível diagnosticar o indivíduo antes dos três anos de idade, entre eles o fato de que o diagnóstico é essencialmente clínico e realizado com base nos critérios descritos no Manual diagnóstico e estatístico de transtornos mentais da sociedade americana de psiquiatria (DSM), entrevistas com os pais, observação do comportamento da criança e aplicação de questionários e escalas padronizadas como ADI-R (*Autism Diagnostic Interview-Revised*), CARS (*Childhood Autism Rating Scale*), ADOS (*Autism Diagnostic Observation Schedule*) e ADOS-G (*Autism Diagnostic Observation Schedule-Generic*). Estas ferramentas e questionários para a realização do diagnóstico ainda carecem de validação e adaptação ao contexto brasileiro. O estudo das características antropométricas em indivíduos com TEA e indivíduos em desenvolvimento normal mostrou que existem diferenças significativas como distâncias entre as pupilas, formato das orelhas, estrabismo e circunferência da cabeça. Portanto, a hipótese aqui trabalhada é que seria possível classificar indivíduos com TEA e indivíduos normais com base nas medidas antropométricas faciais. Desta forma, este projeto visa à construção de um classificador que, dada uma imagem facial de uma criança, consiga discriminar entre autista e normal, auxiliando assim o diagnóstico.

Palavras-chaves: Classificação. Autismo. Antropometria.

Abstract

The autism spectrum disorder (ASD) is a developmental disorder that persistently affects the communication and social interaction and causes restricted and repetitive patterns of behavior, interests and activities in the individual. These symptoms are present since the early infancy and limit the patient. However, several factors impair the diagnosis of individuals before the three years. Such diagnosis is essentially clinical and conducted based on the criteria described in the Diagnostic and Statistical Manual of the American Psychiatric Association (DSM), interviews with parents, child behavioral observation and application of standardized questionnaires and scales such as ADI-R (Autism Diagnostic Interview-Revised), CARS (Childhood Autism Rating Scale), ADOS (Autism Diagnostic Observation Schedule), ADOS-G (Autism Diagnostic Observation Schedule-Generic). These diagnosis tools and questionnaires still require validation and adaptation to the Brazilian context. The study of anthropometric characteristics in individuals with ASD and individuals in normal development has shown the existence of significant differences between these two groups, such as distances between the pupils, ears format, strabismus and head circumference. So the hypothesis here considered is that it would be possible to classify individuals with ASD and healthy controls subjects based on the facial anthropometric measures. Thus, this project aims to build a classifier that given a facial image of a child can discriminate between autistic and control, helping the diagnosis.

Keywords: Classification. Autism. Anthropometry.

Lista de figuras

Figura 1 – Exemplo dos *Landmarks* ou pontos de controle antropométricos 18
Figura 2 – Pontos de controle 19
Figura 3 – Exemplo de características *Haar* 21
Figura 4 – Característica que corresponde à região dos olhos 22
Figura 5 – Característica que corresponde à região do nariz 22
Figura 6 – Visão geral do *Constrained Local Neural Fields* 24
Figura 7 – Áreas de interesse em uma imagem 25
Figura 8 – Exemplo de classificação binária 27
Figura 9 – Distâncias em garotos com TEA 33
Figura 10 – Comando OpenFace *FaceLandmarkImg* 39
Figura 11 – Comando OpenFace *FeatureExtraction* 39
Figura 12 – Exemplo de extração Viola-Jones 44
Figura 13 – Exemplo de extração Viola-Jones 45
Figura 14 – Regiões detectadas pelo Viola-Jones 45
Figura 15 – Exemplo da extração CLNF 46

Lista de tabelas

Tabela 1 – Resumo dos trabalhos na literatura 36
Tabela 2 – Cronograma das tarefas passadas e futuras 42

Lista de abreviaturas e siglas

ADI-R	*Autism Diagnostic Interview-Revised*
ADOS	*Autism Diagnostic Observation Schedule*
ADOS-G	*Autism Diagnostic Observation Schedule-Generic*
CARS	*Childhood Autism Rating Scale*
CLNF	Algoritmo para localização de pontos em imagens *Constrained Local Neural Fields*
CLM	Algoritmo para localização de pontos em imagens *Constrained Local Model*
DDBT	Algoritmo *Dynamic Discriminant Bounds Tree*
DSM	Manual diagnóstico e estatístico de transtornos mentais da sociedade americana de psiquiatria
ICA	Algoritmo de fusão de características Análise dos Componentes Independentes
IPq	Instituto de Psiquiatria do Hospital das Clínicas da Universidade de São Paulo
JV	Algoritmo de classificação *Jet Voting*
kNN	Algoritmo de classificação kth Nearest Neighbours
LBP	Descritor visual *Local Binary Patterns*
LDA	Algoritmo de classificação *Linear Discriminant Analysis*
mRMR	*Minimum Redundancy Maximum Relevance*
NU-RLMS	*Regularised Landmark Mean Shift Não Uniforme*
PCA	Análise dos componentes principais
PDD-NOS	Transtorno global do desenvolvimento sem outra especificação

PROTEA	Programa de Transtorno do Espectro Autista
RF	Algoritmo de classificação *Randoms Forests*
RLMS	*Regularised Landmark Mean Shift*
SVM	Algoritmo de classificação *Support Vector Machines*
TEA	Transtorno do Espectro Autista
UNIFESP	Universidade Federal de São Paulo

Sumário

1 Introdução 11
1.1 Objetivos 13
1.2 Justificativas 14
1.3 Organização do documento 15

2 Conceitos Fundamentais 17
2.1 Antropometria 17
2.2 Imagem 18
2.3 Detecção de Faces 19
2.3.1 Algoritmo Viola-Jones 20
2.3.2 Algoritmo Constrained Local Neural Fields 23
2.4 Aprendizado de máquina supervisionado 25
2.4.1 Classificação 26
2.4.2 Redução da dimensionalidade 27

3 Revisão bibliográfica 30

4 Materiais e métodos 37
4.1 Construção de uma base de imagens faciais de crianças afetadas por TEA e crianças em desenvolvimento típico 37
4.2 Processamento das imagens coletadas para extração das medidas antropométricas faciais 38
4.3 Pesquisa e aplicação de métodos de seleção de característica 40
4.4 Treinamento e comparação de classificadores 40
4.5 Plano de Trabalho e Cronograma 41

5 Resultados preliminares 43
5.1 Protocolo para coleta de dados 43
5.2 Extração dos pontos de controles 44
5.3 Cálculo das distâncias 46
5.4 Dificuldades encontradas 47

6 Considerações Finais 48

Referências[1] 49

APÊNDICES 54

Apêndice A – Termo de consentimento livre e esclarecido 55

Apêndice B – Termo de assentimento 57

[1] De acordo com a Associação Brasileira de Normas Técnicas. NBR 6023.

1 Introdução

Este capítulo tem por objetivo realizar a contextualização do problema de pesquisa e são apresentados os objetivos da realização deste trabalho na seção 1.1 e sua justificativa na seção 1.2. Ao final deste capítulo, na seção 1.3, está descrita a organização do restante do documento.

O transtorno do espectro autista (TEA) é um transtorno de desenvolvimento que prejudica persistentemente a comunicação e a habilidade para desenvolver, manter e compreender relacionamentos, além de causar padrões restritos e repetitivos de comportamento, interesses e atividades no indivíduo. Esses sintomas estão presentes desde o início da infância e limitam ou prejudicam o cotidiano do indivíduo (APA, 2014). O TEA pode ser mais ou menos grave e nos últimos anos passou a ser categorizado com um espectro de transtornos que incluem o transtorno desintegrativo da infância, a síndrome de Asperger e o transtorno global do desenvolvimento sem outra especificação (PDD-NOS) (LORD et al., 2000; JOHNSON; MYERS et al., 2007). De acordo com a gravidade, o indivíduo pode apresentar desde dificuldade na abordagem social e compartilhamento reduzido de interesses, emoções e afeto até comprometimento da fala, hiperatividade e comportamento de interesse fixo e excessivo em objetos e no ambiente (APA, 2014).

O TEA afeta, aproximadamente, 1,47% das crianças de até 8 anos de idade, nos Estados Unidos (BAIO, 2010). No Brasil, um estudo piloto estimou a prevalência de TEA sendo 0,27% das crianças, no município de Atibaia em São Paulo (PAULA et al., 2011). Porém é difícil estimar a quantidade real de pessoas afetadas no país, pois existem diversas limitações como estudos realizados com amostras pequenas, falta de conscientização da população e profissionais locais, falta de registros, diagnóstico errôneo e diagnóstico tardio (PAULA et al., 2011; RICE, 2009).

Os sintomas de TEA surgem já nos primeiros anos de vida (LEVY; MANDELL; SCHULTZ, 2009). Contudo, vários fatores impedem que seja possível diagnosticar o indivíduo antes dos três anos de idade. Essa dificuldade advém do caráter heterogêneo das causas do TEA nas crianças afetadas e a complexidade em se estabelecer uma estratégia de detecção precoce (ZWAIGENBAUM; BRYSON; GARON, 2013; MANDELL; NOVAK; ZUBRITSKY, 2005). Os sintomas não são consistentes entre todos os pacientes e alguns comportamentos podem desaparecer com o passar dos anos (LORD et al., 2000).

As causas do TEA ainda não são totalmente claras, porém existem evidências de causas genéticas e ambientais (MUHLE; TRENTACOSTE; RAPIN, 2004). Em Colvert et al. (2015) foi conduzido um estudo com irmãos gêmeos para examinar a importância de fatores genéticos e ambientais na suscetibilidade de autismo. Os resultados apresentados indicam correlação entre os valores obtidos através dos questionários ADOS e ADI-R entre irmãos. Além do fator genético, existem hipóteses de que toxinas tenham influência na ocorrência de TEA. Em Deth et al. (2008) é levantada a hipótese de que toxinas do ambiente provoquem estresse oxidativo em indivíduos geneticamente sensíveis e como consequência ocorreria a redução da atividade de metilação.

Outros fatores que dificultam o diagnóstico de TEA são a falta de marcadores biológicos definidos e as mudanças nas definições clínicas e os critérios de diagnóstico com o passar dos anos (BAIO, 2010; MCPARTLAND; REICHOW; VOLKMAR, 2012).

Em seu trabalho, DeMyer, Zeman e Palmer (1964) utilizam a frase *"A face prevê o cérebro"* para explicar o transtorno no desenvolvimento cerebral de pacientes afetados por holoprosencefalia. Para todos os pacientes que foram investigados no estudo foi atestado que anomalias faciais medianas estavam presentes em quadros de holoprosencefalia. Outros estudos também sugerem que a morfologia facial seja um reflexo de problemas cerebrais. Alguns desses trabalhos são apresentados mais detalhadamente no Capítulo 3.

Estudos das características antropométricas em indivíduos com TEA e indivíduos em desenvolvimento típico mostraram que existem diferenças significativas entre esses dois grupos, como por exemplo distâncias entre as pupilas, formato das orelhas, estrabismo, circunferência da cabeça (ALDRIDGE et al., 2011; MILES et al., 2000; RODIER; BRYSON; WELCH, 1997). Portanto, a hipótese aqui considerada é que seria possível classificar indivíduos com TEA e indivíduos normais com base em características antropométricas faciais e assim auxiliar o diagnóstico de autismo.

As características antropométricas podem ser extraídas a partir de medições com fitas métricas ou por meio do processamento de uma imagem facial (FARKAS et al., 2002). A extração por meio de uma imagem ocorre através da detecção de pontos de controle que podem ser usados para o cálculo das distâncias antropométricas. A existência de diferenças antropométricas entre crianças com TEA e em desenvolvimento típico permite que este problema seja visto como um problema de classificação, em que é possível categorizar um indivíduo em um dos grupos de acordo com suas características, nesse caso, as distâncias antropométricas. As técnicas de reconhecimento de padrões têm sido

utilizadas na resolução de diversos problemas que necessitam que dados sejam categorizados em classes identificáveis através da extração de suas características significantes. Dessa forma, um objeto pode ser classificado como parte de uma determinada população (TOU; GONZALEZ, 1974). Técnicas como *Random Forests* (RF) e *Support Vector Machines* (SVM) têm sido utilizadas para diagnóstico de síndromes e transtornos (BOEHRINGER et al., 2006; HAMMOND et al., 2004; ZHAO et al., 2014a).

Os trabalhos encontrados na literatura que analisam a medidas antropométricas faciais de pacientes com TEA foram realizados com ajuda de um equipamento de estereofotogrametria que realiza a reconstrução de uma imagem tridimensional a partir de coordenadas estimadas utilizando várias imagens bidimensionais tiradas de vários ângulos diferentes. Este tipo de equipamento não está disponível no Brasil e possui um custo alto[1] (ALDRIDGE et al., 2005; ALDRIDGE et al., 2011; HAMMOND; SUTTIE, 2012).

1.1 Objetivos

Considerando as perspectivas levantadas pela literatura assim como suas lacunas observadas, este projeto tem como objetivo geral propor e testar um classificador de imagens faciais de crianças, capturadas por uma câmera fotográfica digital comum, que discrimine-as entre as classes TEA e não TEA, assim como criar uma base de imagens faciais de crianças acometidas com o TEA e em desenvolvimento típico a ser utilizada para treinamento desse classificador e estudos futuros.

A entrada do classificador será uma estrutura de dados que já contenha os pontos de controle obtidos das imagens iniciais. Assim, não é o escopo desta pesquisa o estudo e análise dos algoritmos de processamento da imagem (detecção da face e dos componentes faciais), porém entende-se que será necessário tratar as imagens ou normalizá-las, assim como entender o funcionamento básico de tais algoritmos e realizar a adaptação dos mesmos para que seja possível a extração dos pontos e o cálculo das distâncias entre estes pontos.

Os objetivos específicos deste trabalho são:

1. construir uma base de imagens faciais de crianças afetadas por TEA e crianças em desenvolvimento típico a ser utilizada neste trabalho e em trabalhos futuros;

[1] 3dMD Face System e 3dMD Head System possuem custo estimado de U$39000 e U$89000, respectivamente. Informação obtida por e-mail através do departamento de vendas da empresa.

2. processar as imagens coletadas para extração das medidas antropométricas faciais, ajustando os algoritmos de processamento se necessário;
3. pesquisar os métodos de seleção de características que possam ser mais adequados ao problema proposto;
4. aplicar os algoritmos de seleção de características sobre os dados extraídos das imagens;
5. treinar dois ou mais classificadores, baseados em algoritmos distintos de classificação, utilizando as características selecionadas no objetivo específico anterior;
6. comparar os classificadores obtidos da combinação de diferentes métodos de seleção de características e algoritmos de classificação a fim de propor o mais adequado para o problema proposto.

1.2 Justificativas

No Brasil o diagnóstico de crianças em idade pré-escolar ainda é muito raro devido à dificuldade de acesso aos serviços de saúde, falta de suporte social aos pais e falta de conscientização da população. Além disso, não existem serviços educacionais orientados a crianças com TEA o que afeta a qualidade de vida não só das crianças como dos pais também (GOMES et al., 2015).

Segundo o modelo teórico apresentado em Dawson (2008), o diagnóstico precoce seguido de intervenção apropriada possibilitaria ao paciente se aproveitar da plasticidade cerebral para inibir as manifestações do TEA. Por meio de tratamentos adequados é possível realizar prevenção em crianças no grupo de risco antes que todos os sintomas do transtorno sejam manifestados. Em outro trabalho mais recente, Dawson et al. (2010) conduziram um estudo que mostrou ganhos através da intervenção precoce de forma que as crianças envolvidas no estudo apresentaram melhora significativa no funcionamento cognitivo e comportamento adaptativo. Em Corsello (2005) também é discutida a importância da intervenção precoce em crianças afetadas por TEA.

Sendo assim, a proposta aqui apresentada inclui capturar imagens bidimensionais utilizando uma câmera digital comum com resolução mínima de 14 megapixels[2] para a criação de um banco de dados de imagens de indivíduos com TEA a fim de realizar análise

[2] Para a coleta dos dados foi utilizada uma câmera semi-profissional descrita na seção de materiais e métodos

antropométrica. Desta forma, a classificação será baseada nas distâncias extraídas dessas imagens, oferecendo a aplicação de um método financeiramente acessível aos grupos de pesquisa e prática clínica.

A análise das características antropométricas faciais através de imagens oferece uma forma não invasiva de auxílio ao diagnóstico de TEA e, de acordo com os resultados preliminares, poderia auxiliar o diagnóstico precoce que permitiria uma intervenção antecipada e assim aumentaria a qualidade de vida das crianças afetadas pelo TEA. Para realizar esse tipo de análise em imagens é fundamental que exista uma base de dados com indivíduos portadores de TEA assim como indivíduos em desenvolvimento típico para realizar a comparação. Para isso, é necessário que seja estabelecido um protocolo de captura bem definido e o recrutamento de indivíduos em desenvolvimento típico. A construção de uma base de dados de imagens de pacientes com TEA e controles permite a evolução da pesquisa de TEA e que novas hipóteses sejam levantadas e conclusões possam ser obtidas.

Este trabalho faz parte de um projeto maior enviado à FAPESP e coordenado pela Profª Drª Helena Brentani, que propõe o estudo do TEA em conjunto com modelos computacionais. O módulo ao qual este trabalho faz parte tem como objetivo estudar coortes de alto risco e possíveis mecanismos biológicos associados, relacionando fatores ambientais ao risco de transtornos do neurodesenvolvimento, e estudar sistemas computacionais e sua contribuição para classificadores fenotípicos e de diagnóstico precoce para TEA. Assim, apesar deste trabalho focar em características antropométricas faciais para a classificação de crianças autistas e não autistas, não tem a pretensão de realizar o diagnóstico puramente baseado nestas medidas, mas irá compor, futuramente, um classificador multimodal que será baseado em vários aspectos envolvidos no TEA.

1.3 Organização do documento

Este documento está organizado em capítulos descritos a seguir.

No Capítulo 2 são apresentados os conceitos fundamentais ao desenvolvimento desta pesquisa como antropometria, pontos de controle em imagens, algoritmos de detecção de faces, classificação e redução de dimensionalidade.

No Capítulo 3 é apresentada a revisão bibliográfica de métodos de análise antropométrica realizadas para diagnóstico de transtornos e síndromes.

No Capítulo 4 são apresentados os materiais e métodos que foram ou serão utilizados na resolução do problema de pesquisa e coleta dos dados, assim como o plano de trabalho e cronograma.

No Capítulo 5 são apresentados os resultados preliminares e as dificuldades encontradas.

No Capítulo 6 são apresentadas as considerações finais e é feita a conclusão deste documento.

2 Conceitos Fundamentais

Este capítulo apresenta os conceitos fundamentais envolvidos nesta pesquisa. Na seção 2.1 é apresentada a definição de antropometria; na seção 2.2 são apresentados os conceitos de imagem e pontos de controle; na seção 2.3 são apresentados os algoritmos de detecção facial que foram testados; na seção 2.4 é apresentada a definição de aprendizado de máquina supervisionado; na subseção 2.4.1 é apresentado o conceito de classificação supervisionada e por fim, na subseção 2.4.2 é definido o problema da dimensionalidade e seleção de características.

2.1 Antropometria

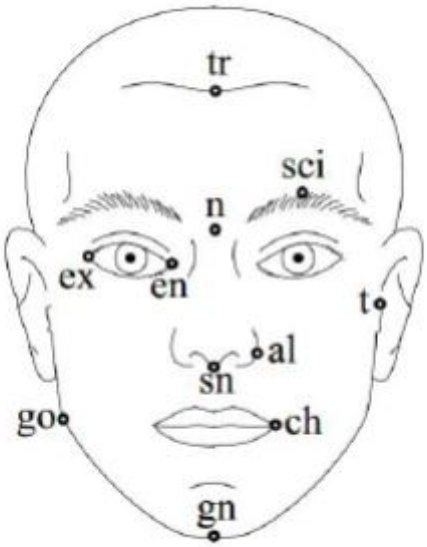
Antropometria é a ciência que estuda as medidas do corpo humano incluindo altura, peso, tamanho das partes componentes do corpo (FARKAS, 1994).

Através da antropometria é possível realizar uma análise da morfologia do corpo de forma não invasiva. A obtenção das medidas pode ser feita diretamente, com equipamentos de medição como fitas e trenas, ou indiretamente, utilizando imagens de radiografias, imagens bidimensionais ou tridimensionais por estereofotogrametria (WEINBERG et al., 2006).

Para assegurar a confiabilidade, os dados devem ser obtidos de forma sistemática tanto direta como indiretamente, pois variações podem ocorrer nos dois tipos de abordagem (FARKAS et al., 2002). Para isso, é necessária a utilização de um sistema de medida meticuloso baseado em pontos de referência (DEUTSCH et al., 2012). FARKAS propôs uma manual antropométrico para descrever medidas e definir normas para captura de dados antropométricos.

As medições antropométricas são feitas por meio de pontos bem definidos sobre o corpo humano chamados de *landmarks* ou pontos de controle. Neste trabalho é usada a nomenclatura ponto de controle. Na figura 1 são apresentados exemplos de ponto de controle sobre a face humana.

Figura 1 – Exemplo dos *Landmarks* ou pontos de controle antropométricos



Fonte: Adaptado de Farkas (1994)

2.2 Imagem

A seguir, está descrita a definição de imagem digital, pixels e pontos de controle baseada em (GONZALEZ; WOODS, 2002).

Uma imagem pode ser definida como um função de duas dimensões $f(x, y)$ em que x e y são as coordenadas espaciais do plano e o valor de f de cada par ordenado de coordenadas (x, y) é chamado intensidade da imagem neste ponto ou nível de cinza (no caso de imagens em escala de cinza). Uma imagem é chamada de imagem digital quando os valores de x, y e a intensidade de f são valores finitos e discretos.

No contexto de fotografias, a imagem é gerada por meio da composição de uma fonte de iluminação e os elementos da cena que podem refletir ou absorver a luz emitida pela fonte. A função $f(x, y)$ é caracterizada por dois aspectos: a intensidade de luz na cena e a intensidade da iluminação sendo refletida pelos objetos da cena. Esses conceitos são de extrema importância para algoritmos de detecção de objetos, pois a iluminação pode influenciar a acurácia na detecção dos pontos de controle que se desejam extrair da imagem.

Para criar uma representação digital da imagem é necessário converter o dado capturado para sua forma digital. Essa conversão envolve dois processos que são amostragem

e quantização. O processo de discretização das coordenadas (x, y) é chamado amostragem e dos valores de f é chamado quantização.

Uma imagem $f(x, y)$ digitalizada é uma matriz $M \times N$, em que M é o número de linhas e N o número de colunas. Escrevendo a matriz $M \times N$ que representa uma imagem digital têm-se:

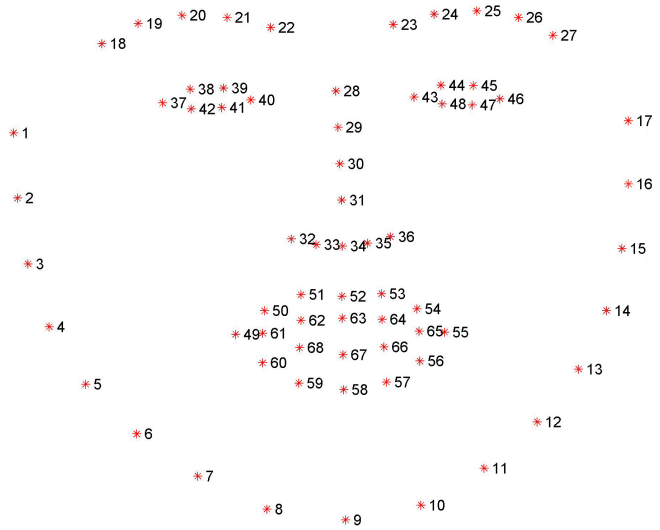
$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & \ddots & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix} \tag{1}$$

Em uma imagem digital a origem, por convenção é o primeiro elemento da matriz $M \times N$. O lado direito da equação por definição é a imagem digital e cada elemento da matriz do lado esquerdo é chamado elemento da imagem ou pixel (do inglês, *picture element*.

2.3 Detecção de Faces

Para extrair medidas antropométricas de imagens que representam faces humanas, é necessário identificar os pontos de controle, conforme mostrado na figura 2. Um ponto de controle é composto pelas coordenadas *(x,y)* do pixel que representa tal ponto.

Figura 2 – Os 68 pontos de controle que são extraídos das imagens



Dada uma imagem frontal de uma face humana, a viabilização da identificação dos pontos de controle requer a execução prévia de dois passos: detecção da área da

face e detecção dos componentes faciais dentro da área facial identificada. Existem três abordagens para a realização da detecção da área facial e componentes faciais: técnicas baseadas em modelos ou *templates*, técnicas baseadas em características e técnicas que usam uma combinação das duas primeiras (HSU; ABDEL-MOTTALEB; JAIN, 2002). As técnicas baseadas em modelos assumem uma posição inicial para a face na imagem e corrigem a posição dos pontos através de funções. Já as técnicas baseadas em características tratam a detecção como um problema de classificação em que as regiões da imagem são classificadas em face e não face.

Nesta seção são apresentados os algoritmos de detecção de faces em imagens. Na subseção 2.3.1 é apresentado o algoritmo Viola-Jones e na subseção 2.3.2 é apresentado o algoritmo *Constrained Local Neural Fiels*. O algoritmo Viola-Jones utiliza classificadores e características para detectar os componentes faciais enquanto que o *Constrained Local Neural Fields* baseia-se em modelos.

2.3.1 Algoritmo Viola-Jones

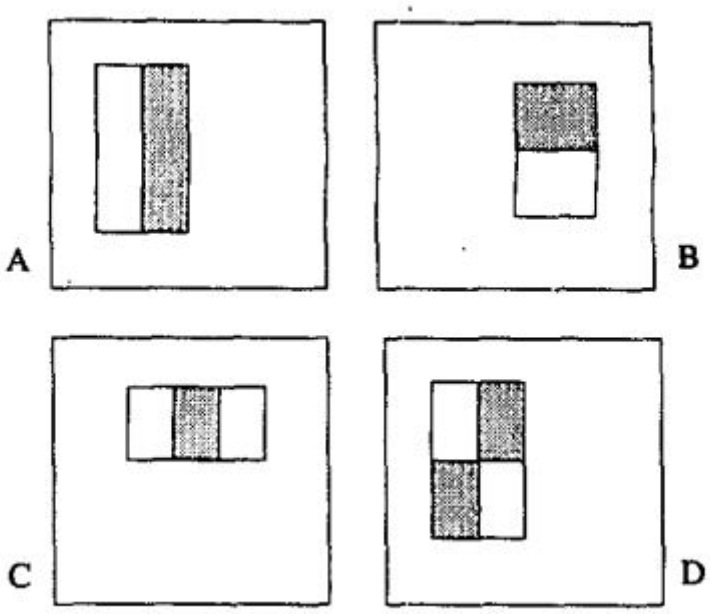
As definições nesta subseção foram escritas com base no trabalho de Viola e Jones (2001).

O algoritmo Viola-Jones tem como objetivo detectar faces e componentes faciais que estejam em posição frontal. Isso é feito através de classificadores que separam as regiões da imagem em dois grupos: faces e não faces. O algoritmo tem quatro etapas principais descritas a seguir:

1. seleção das características *Haar*;
2. transformação da imagem para representação de imagem integral;
3. treinamento do AdaBoost;
4. aplicação dos classificadores em cascata.

As características utilizadas pelo Viola-Jones são características *Haar* baseadas nas características propostas por Papageorgiou, Oren e Poggio (1998). Essas características representam diferenças nas intensidades médias entre duas regiões retangulares. As variações no contraste entre conjuntos de pixels são usadas para determinar regiões escuras e claras na imagem. Existem três tipos de características *Haar*, que são mostradas na figura 3.

Figura 3 – Exemplo de características Haar: (A) e (B) características com 2 retângulos, (C) característica com 3 retângulos e (D) característica com 4 retângulos



Fonte: Adaptado de Viola e Jones (2001)

O valor da característica composta por dois retângulos é a soma das intensidades dos pixels que estão sob o retângulo branco subtraída da soma das intensidades dos pixels que estão sob o retângulo preto; o valor para a característica composta por três retângulos é a soma das intensidades dos pixels sob os dois retângulos pretos subtraída da soma das intensidades dos pixels sob o retângulo branco e finalmente, o valor da característica composta por quatro retângulos é a subtração entre a soma das intensidades dos pixels dos pares das diagonais.

A face humana apresenta certas similaridades nas quais é possível fazer correspondência com as características *Haar*. Um exemplo é a região dos olhos que é mais escura que a região das bochechas. A figura 4 exemplifica a característica *Haar* que corresponde à região dos olhos e a figura 5 exemplifica a característica que corresponde à região do nariz, em que a região do nariz é mais clara do que a região dos olhos.

Para avaliar as características e realizar as somas de forma mais rápida, Viola e Jones (2001) propuseram o conceito de imagem integral. A imagem integral é uma representação intermediária em que a coordenada (x, y) contém a soma dos pixels acima e

Figura 4 – Característica *Haar* que corresponde à região dos olhos

Figura 5 – Característica *Haar* que corresponde à região do nariz

à esquerda de (x, y) conforme a equação 2, em que $g(x, y)$ é a imagem integral e $f(x, y)$ é a imagem original.

$$g(x, y) = \sum_{x' \leq x, y' \leq y} f(x', y') \tag{2}$$

Usando as recorrências descritas nas equações 3 e 4, a imagem integral é calculada em apenas uma varredura sobre a imagem original.

$$s(x, y) = s(x, y - 1) + f(x, y) \tag{3}$$

$$g(x, y) = g(x - 1, y) + s(x, y) \tag{4}$$

Apesar do Viola-Jones utilizar somas simples para o cálculo das características, o número total de características *Haar* em uma imagem é grande. Viola e Jones (2001) explicam que em uma imagem existem sub-janelas ou regiões em que é possível encontrar as características *Haar*.

Por este motivo é usado uma variação do algoritmo AdaBoost (FREUND; SCHAPIRE, 1995) com o objetivo de selecionar as características que melhor separam os exemplos positivos e negativos. O algoritmo AdaBoost constrói um classificador forte através da combinação linear de classificadores fracos e seus pesos.

Ao final do treinamento, são selecionadas as características que definem as sub-janelas positivas e as sub-janelas que não correspondem a componentes faciais são descartadas. Em cascata, cada estágio é composta de um classificador forte. Assim, todas as características são agrupadas em várias fases, em que cada fase tem certo número de características.

A forma geral do processo de detecção pode ser vista como uma árvore de decisão degenerada. Um resultado positivo a partir do primeiro classificador forte desencadeia a avaliação de um segundo classificador; um resultado positivo a partir do segundo classificador desencadeia um terceiro classificador, e assim sucessivamente. Um resultado negativo em qualquer ponto leva à rejeição imediata da sub-janela.

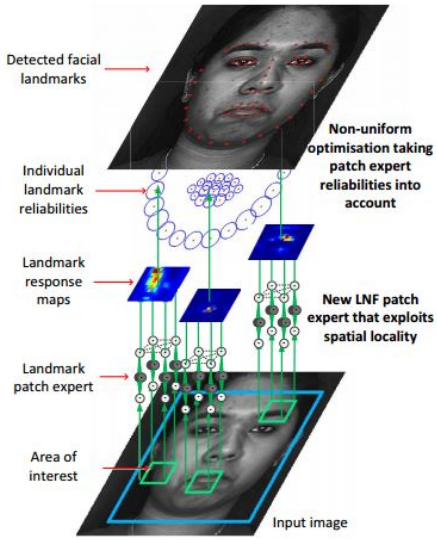
2.3.2 Algoritmo Constrained Local Neural Fields

As definições nesta subseção foram escritas com base no trabalho em Baltrusaitis, Robinson e Morency (2013).

O algoritmo *Constrained Local Neural Fields* (CLNF) segue uma abordagem baseada em modelo e tenta resolver os problemas encontrados na detecção de características em cenas complexas, como baixa luminosidade e oclusões nas faces a serem detectadas. O

CLNF utiliza um modelo de distribuição dos pontos e realiza a detecção adaptando os pontos de controle iniciais do modelo a face na imagem. A figura 6 apresenta a visão geral do algoritmo CLNF.

Figura 6 – Visão geral do *Constrained Local Neural Fields*



Fonte: Adaptado de Baltrusaitis, Robinson e Morency (2013)

O algoritmo CLNF é composto pelas três etapas principais descritas a seguir:

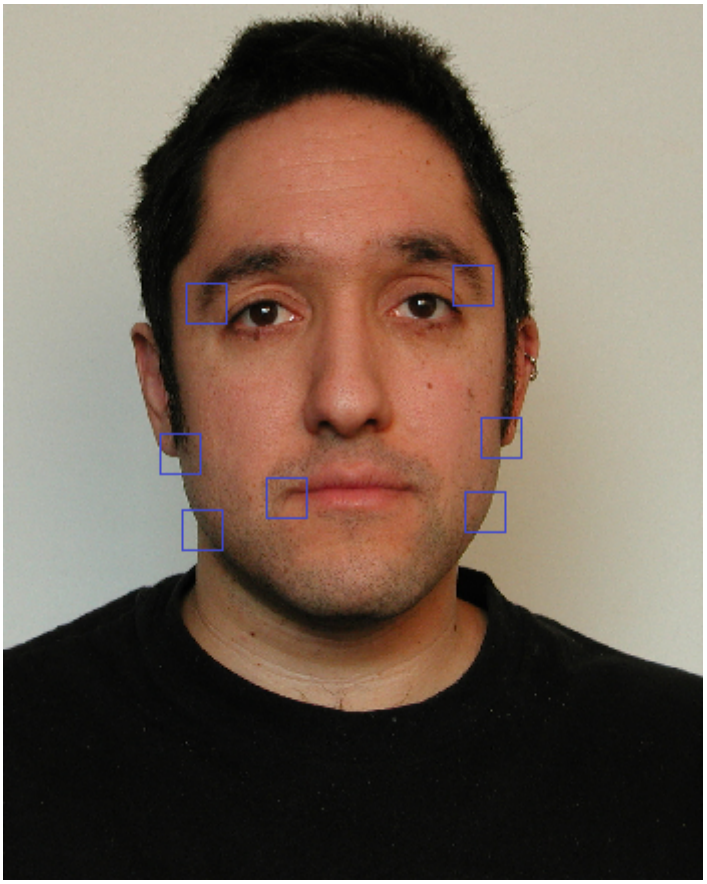
1. aplicação do modelo de distribuição dos pontos;
2. detecção de regiões de interesse;
3. adaptação do modelo.

O modelo de distribuição dos pontos é construído utilizando imagens em que os pontos de controle foram anotados manualmente. É realizado o treinamento no qual os pontos do modelo de distribuição são estimados a partir dessas imagens anotadas. Após obter o modelo de distribuição são definidas áreas de interesse em que o algoritmo realiza otimização para aproximar os pontos do modelo de distribuição na imagem alvo. Uma área de interesse é uma região da imagem em que existe um ponto de controle em potencial.

No CLNF, as áreas de interesse são ajustadas usando parâmetros. Esses parâmetros indicam áreas ao redor de um ponto de controle e são usadas para avaliar se um ponto de controle na imagem facial foi detectado ou não. A figura 7 mostra exemplos de áreas de interesse na imagem.

Durante o passo de detecção é avaliada a probabilidade de um ponto do modelo de distribuição estar alinhado ao pixel da imagem que corresponde ao ponto de controle

Figura 7 – Áreas de interesse em uma imagem



desejado. Essa probabilidade é calculada com base nos valores de intensidade dos pixels na região de interesse.

A partir desta avaliação pode ser realizado o *fitting*. Tendo uma estimativa de parâmetro p_0 deseja-se encontrar um parâmetro de atualização Δp que aproxime uma solução do valor ótimo p^* (parâmetro p que indica se o ponto foi detectado) de forma que $p^* = p_0 + \Delta p$. O processo de *fitting* é iterativo e realizado através de uma função objetivo. Depois de encontrados os parâmetros Δp o modelo é adaptado a imagem e os pontos de controle são detectados. Por se basear em um modelo e utilizar estimativas, o algoritmo CLNF consegue lidar de maneira eficiente com oclusões na imagem.

2.4 Aprendizado de máquina supervisionado

As definições nesta seção foram extraídas de Mitchell (1997), Duda, Hart et al. (1973) e Dash e Liu (1997).

Aprendizado de máquina é o estudo de algoritmos que aperfeiçoam o desempenho na realização de uma tarefa através do aprendizado automatizado. Uma das categorias

desses algoritmos é conhecida como aprendizado supervisionado, na qual são apresentados exemplos de entradas e as saídas desejadas correspondentes às entradas, sendo o objetivo do algoritmo aprender a regra geral que mapeia entradas e saídas.

No contexto de aprendizado de máquina, uma entrada é representada como o conjunto de atributos ou características que a definem, sendo que uma característica é uma variável aleatória. Esse conjunto de características é um vetor de valores numéricos de dimensão n.

2.4.1 Classificação

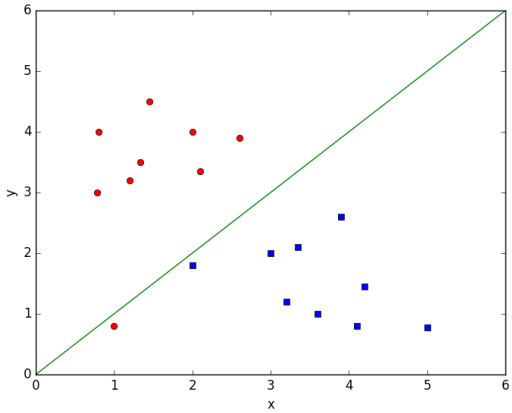
Uma tarefa de classificação em sua essência tem como objetivo aprender o modelo que permite o reconhecimento de um padrão. O reconhecimento de um padrão é definido pela categorização de dado de entrada em uma classe identificável por meio de suas características significantes.

Reconhecimento de padrões tem sido empregado em diversas tarefas de aprendizado de máquina como reconhecimento de fala, identificação de digitais, reconhecimento óptico de caracteres, análise de cadeias de DNA, etc. Estes problemas, em sua maioria, envolvem a categorização dos dados de acordo com suas características comuns, de forma que a máquina consiga identificar as classes das entradas. O reconhecimento ocorre baseando-se na inferência indutiva realizada através do aprendizado de conceitos ou características aprendidas no passado.

Dentro do contexto de aprendizado supervisionado, para a realização de uma tarefa de classificação são fornecidos ao algoritmo dados de treinamento e a partir destes dados um modelo que discrimina as classes é aprendido. A classificação é dita binária quando há apenas duas classes envolvidas ou multiclasse quando há mais de duas classes envolvidas.

Considerando a classificação binária, o modelo aprendido pelo algoritmo de classificação é uma função que separa os dados em um plano n-dimensional, no qual n é a dimensão do vetor de características dos elementos de treinamento e dos elementos a serem classificados. Na figura 8 a função que separa as classes é linear e é baseada em duas características que são representadas pelo eixo x e pelo eixo y. A linha que separa as duas classes é chamada de fronteira de decisão.

Figura 8 – Exemplo de classificação binária



2.4.2 Redução da dimensionalidade

Um algoritmo de classificação baseia-se nas características para reconhecer um padrão. Quanto maior a dimensão dos dados, isto é, quanto maior a quantidade de características, maior é a possibilidade de separação entre as classes. Logo, seria de se esperar que um conjunto grande de características permitisse que um modelo tenha desempenho melhor em distinguir um padrão de outro. Porém, quanto maior a dimensão dos dados maior é a quantidade de parâmetros necessários para definir a fronteira de decisão (classificador), o que exige maior quantidade de dados de treinamento para estimá-los. Em problemas reais, a quantidade de dados de treinamento é limitada. Sendo assim, para uma amostra de treinamento de tamanho constante, o aumento da dimensão é acompanhado do aumento do erro de estimação dos parâmetros do modelo classificador. Assim, ao aumentar a dimensionalidade, mesmo com o aumento da separação entre as classes, há um ponto em que o aumento do erro de estimação faz com que a taxa de erro de classificação aumente, degradando o desempenho da classificação. Esse dilema é chamado de maldição da dimensionalidade (HUGHES, 1968). Portanto, é necessário que a dimensão dos dados seja adequada para permitir a construção de um modelo eficaz que consiga classificar padrões utilizando uma amostra de treinamento de tamanho limitado. Para isso, muitas vezes é necessária uma redução de dimensionalidade.

Além disso, em certos problemas de classificação, mesmo que a dimensão não seja tão grande, é necessário eliminar características redundantes ou irrelevantes para a classe alvo, evitando assim ruídos, o que também pode ser realizado a partir da redução da dimensionalidade.

Uma das formas de redução da dimensionalidade é a seleção de características. Um método de seleção de características tipicamente realiza uma busca no espaço de características para encontrar o subconjunto ótimo segundo algum critério pré-definido. Porém essa busca pode ser custosa de acordo com o tamanho do conjunto, pois o número de soluções candidatas é 2^N, sendo N o número original de características. Por este motivo foram criados métodos que se baseiam em heurísticas ou busca aleatória para reduzir a complexidade computacional.

Sendo assim, os vários métodos de seleção de características podem se distinguir de acordo com como definem os seguintes itens:

1. forma de percorrer o espaço de busca;
2. função de avaliação de cada solução candidata;
3. critério de parada;
4. relação com o classificador.

Os métodos para percorrer o espaço de busca utilizam abordagens diferentes que incluem realizar busca completa, utilizar heurísticas ou de forma aleatória.

Um método que utiliza busca completa garante que o subconjunto ótimo de características será selecionado e a busca pode ser exaustiva ou não. Os algoritmos de busca completa utilizam a função de avaliação para garantir a otimalidade do subconjunto final. Alguns exemplos de algoritmos desse tipo são *branch and bound* (B&B) (NARENDRA; FUKUNAGA, 1977), *Best first search* (BFF) (XU; YAN; CHANG, 1988) e busca *Beam* (DOAK, 1992).

Os métodos que utilizam heurísticas são iterativos e incrementais e as características podem ser adicionadas, removidas ou uma combinação de adição e remoção do subconjunto inicial. Nesse tipo de método a ordem do espaço de busca é $O(N^2)$ ou menos, o que permite métodos mais rápidos, porém a solução ótima não é garantida. Entre os métodos que utilizam heurísticas estão Relief (KIRA; RENDELL, 1992) e método de árvore de decisão (DTM) (CARDIE, 1993).

Os métodos aleatórios avaliam subconjuntos das soluções candidatas de forma aleatória e não necessariamente percorrem o espaço de busca completo, pois utilizam um número máximo de iterações. Esses métodos não garantem que o subconjunto final é uma solução ótima. Alguns exemplos de métodos aleatórios são algoritmo genético (VAFAIE; IMAM, 1994) e *simulated annealing* (DOAK, 1992).

Quanto a relação com o classificador os algoritmos de seleção de características podem ser divididos em três categorias: filtro, camada (do inglês *wrapper*) e embutidos.

Os métodos do tipo filtro são independentes do classificador, ao passo que métodos camada utilizam o classificador como função de avaliação. Os métodos embutidos são uma combinação das outras duas categorias.

Os métodos do tipo filtro se baseiam nas características gerais do conjunto de treinamento e obtém o subconjunto de características relevantes utilizando um índice ou ranking de relevância. A desvantagem dos algoritmos do tipo filtro é que estes métodos não consideram a correlação entre as características e podem escolher características redundantes.

Os métodos do tipo camada utilizam o próprio classificador como função de avaliação para o subconjunto de características ótimo. Esse tipo de método costuma apresentar bons resultados, porém o custo computacional é alto.

Os métodos embutidos se aproveitam das vantagens dos outros dois tipos de métodos e fazem parte do próprio método de classificação.

3 Revisão bibliográfica

Neste capítulo são apresentados os trabalhos encontrados durante a revisão da literatura de análise antropométrica facial e transtornos.

Em alguns casos, é possível identificar que indivíduos acometidos com certas síndromes apresentam a morfologia facial nitidamente distinguível de indivíduos em desenvolvimento típico. Pode ser citada como exemplo a síndrome de Down, em que os indivíduos apresentam um fenótipo facial típico (ZHAO et al., 2014b). A morfologia facial poderia guardar em si reflexos sobre o desenvolvimento cerebral (ALDRIDGE et al., 2005). Dessa forma, seria possível analisar esses reflexos em um indivíduo que possui determinada síndrome de acordo com suas medidas faciais e o padrão de face. Em trabalhos correlatos é apresentada essa mesma hipótese.

Boehringer et al. (2006) propôs uma análise automatizada da face que permite classificar a imagem de entrada entre dez síndromes diferentes (Síndrome de DiGeorge, Cri-du-chat, Cornelia de Lange, X frágil, Mucopolissacarídose tipo 3, Noonan, Prader-Willi, Smith-Lemli-Opitz, Sotos, Williams-Beuren). Os dados extraídos das imagens foram analisados usando diversos classificadores, especificamente *Linear Discriminant Analysis* (LDA), *Suport Vector Machines* (SVM), *kth Nearest Neighbours* (kNN) e *Jet Voting* (JV). Foi realizada a classificação de imagens de entrada entre as síndromes de duas formas: a primeira, a imagem de entrada foi classificada entre uma das dez síndromes citadas e a segunda, a imagem foi classificada entre duas síndromes de maneira que fosse possível estabelecer padrões comuns entre síndromes. Para a validação os autores utilizaram validação cruzada 10-*fold*.

As imagens utilizadas como dados de entradas foram obtidas através de câmeras comuns e também foram gravados vídeos de onde foram extraídos *frames* estáticos. A extração dos pontos de controle nas imagens foi feita de duas formas com posicionamento do pontos automaticamente (detecção realizada através de software) e com posicionamento manual.

Os resultados obtidos pelos autores apresentam que o algoritmo LDA obteve, na média geral, 76% de acurácia na classificação. Os métodos SVM e kNN obtiveram desempenho de 70% e 68%, respectivamente. O pior desempenho na acurácia da classificação foi do algoritmo JV que atingiu 55%. Entre as síndromes, a síndrome do X frágil, Mucopolissacarídose tipo 3, Williams-Beuren, Cornélia de Lange, Noonan, Prader-Willi, DiGeorge

e Sotos obtiveram os melhores resultados na classificação sendo que, utilizando LDA, a acurácia foi de 99,9%, 94.2%, 88.1%, 83.7%, 78.4%, 77.9%, 77.4% e 72.3%, respectivamente. Os piores resultados na classificação, utilizando LDA, foram das síndromes Smith-Lemli-Opitz com 59.0% e Cri-du-chat com 27.5%.

Zhao et al. (2014a) propôs um *framework* para diagnóstico de síndrome de Down usando técnicas de processamento de imagem e aprendizado de máquina. A técnica proposta utiliza imagens faciais frontais que não precisam estar padronizadas e podem ter sido obtidas com qualquer tipo de câmera básica. Os autores utilizaram um conjunto de imagens adquiridas com vários tipos de câmera e sob diversas condições de iluminação.

Para a extração dos pontos de controle foi empregado o método *Constrained Local Model* (CLM) em conjunto com análise dos componentes independentes (ICA) e as características de textura da imagem foram avaliadas usando *local binary patterns* (LBP) e *Gabor wavelets*. Os autores também fizeram uma comparação entre os métodos de classificação SVM, kNN, *Random Forests* (RFs) e LDA. Para validação da classificação os autores utilizaram o método *leave-one-out*.

Os resultados da classificação utilizando SVM com função radial e LDA obtiveram, em ambos, 96,7% de acurácia. SVM com função linear, kNN e RF obtiveram, respectivamente, 95%, 92,6% e 93% de acurácia na classificação.

O método proposto também foi aplicado a um outro conjunto de dados de pacientes afetados por outras síndromes que causam dismorfia para demonstrar a robustez do método. O melhor resultado na classificação foi obtido utilizando SVM com função linear em que o resultado foi 97% de acurácia. SVM com função radial, kNN, RF e LDA obtiveram, respectivamente, 94%, 96%, 91% e 96%.

Compton et al. (2007) conduziram uma análise das diferenças antropométricas entre pacientes esquizofrênicos e um grupo de controle. Os dados foram obtidos através da medição direta com equipamento métrico. Apenas sete medidas faciais foram analisadas de modo que fosse possível replicar estudos anteriores e também evitar medidas correlacionadas e redundâncias.

Foi aplicada análise de agrupamentos para examinar as diferenças morfológicas. A análise foi realizada usando amostras pareadas por idade para garantir uma combinação de medidas faciais que discriminasse os casos dos controles. As medidas foram transformadas em z-*score* e foi utilizado método de Ward com distâncias euclidianas quadradas para o agrupamento. Outras medidas de distância foram utilizadas para comparação (distância

euclidiana, Chebychev e City Block) e para determinar se a medida de distância era uma influência na análise dos agrupamentos. As médias das medidas faciais de cada agrupamento foram comparadas utilizando teste t de *Student* e, para testar a associação de cada tipo facial de casos e controles com os grupos, foi utilizado teste chi-quadrado. Não foram obtidos resultados que demonstrassem que é possível identificar um tipo facial específico para pacientes com esquizofrenia, apesar de que algumas medidas faciais fossem diferentes.

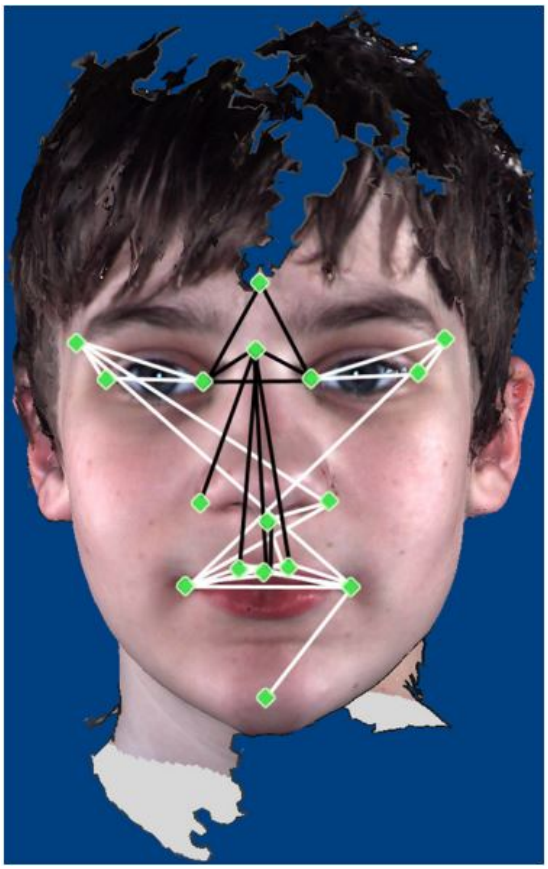
Em Aldridge et al. (2011) foi executada análise das diferenças antropométricas de indivíduos com TEA e indivíduos em desenvolvimento típico. Os autores utilizaram um equipamento de fotografia tridimensional e estereofotogrametria[1] para obter medidas faciais mais precisas e testar duas hipóteses: a primeira de que as crianças com TEA apresentam diferenças na morfologia facial em comparação com as crianças em desenvolvimento típico; e a segunda que existem subgrupos de fenótipos faciais entre a população de crianças com TEA. Todos os participantes eram do sexo masculino. Os resultados apresentados indicaram que 39 das 136 distâncias calculadas ($\approx 28,7\%$ das distâncias) eram significantemente diferentes entre garotos com TEA em comparação com garotos em desenvolvimento típico.

Na figura 9 são apresentadas as distâncias que diferem em garotos com TEA de todos grupos em comparação com garotos em desenvolvimento típico. Foi identificado que as distâncias entre o canto interno dos olhos e o ponto central entre os olhos, ponto central da testa e ponto central entre os olhos, e os pontos localizados no nariz e o filtro demonstraram ser significativamente reduzidas no grupo de garotos com TEA, enquanto que as distâncias da região nasal inferior e bucal, cantos internos e externos dos olhos e a face superior, face superior lateral com os olhos e lado da boca demonstraram ser significativamente maiores.

Além da comparação entre as distâncias, foi realizada análise de agrupamentos na qual a maioria dos garotos com TEA foram agrupados com os garotos em desenvolvimento típico. No entanto, dois subgrupos de garotos com TEA se diferenciaram de ambos os grupos no que diz respeito a morfologia da face. Em tais subgrupos os garotos possuíam características morfológicas que puderam ser correlacionadas com a severidade do autismo. Também foi identificado que a idade e o tamanho da cabeça não influenciaram no agrupamento.

[1] *3dMD Cranial System*

Figura 9 – Distâncias que diferem em garotos com TEA - Visão geral



Fonte: Adaptado de Aldridge et al. (2011)

No trabalho de Ozgen et al. (2011) foi realizada a análise das diferenças morfológicas entre crianças com TEA e em desenvolvimento típico. Os dados foram obtidos através de medição direta com equipamento métrico.

Foi criada uma estrutura de árvore compreendendo 29 áreas anatômicas, cada área foi subdividida em 98 estruturas e cada estrutura contém 638 anormalidades morfológicas padronizadas. As anormalidades foram subdividas em anomalias maiores, causadas por desenvolvimento anormal, e variações menores. O estudo foi realizado de forma a identificar qual dos dois tipos de anomalias ocorrem em crianças com TEA: anomalias maiores ou variações morfológicas menores. Para realizar a comparação, foi utilizado pareamento caso-controle no qual era considerado etnia, gênero e idade.

Não foram utilizados métodos de aprendizado de máquina, foi feita apenas uma análise estatística das anormalidades entre crianças com TEA e controles. Na comparação com o grupo de controle, o grupo com TEA apresentou uma prevalência maior de anomalias maiores.

O trabalho de Obafemi-Ajayi et al. (2015) é uma continuação do trabalho de Aldridge et al. (2011) utilizando o mesmo equipamento. Foi realizada uma análise de agrupamentos utilizando medidas faciais de garotos autistas com o objetivo de identificar os grupos de garotos com características morfológicas semelhantes e correlacionar com os traços comportamentais e clínicos. Para isso, foram utilizadas imagens tridimensionais em que eram calculadas as distâncias geodésicas. O aparelho utiliza estereofotogrametria para gerar uma imagem facial tridimensional e extrair as distâncias entre as superfícies.

A análise de agrupamentos foi realizada para separar os grupos de garotos com TEA e em seguida foi realizada classificação para determinar qual grupo possuía características faciais clinicamente discriminantes. Foram escolhidos quatro métodos de análise de agrupamentos para comparação: *Expectation Maximization* (EM), *Self-Organizing Maps* (SOM), *K-Means* e *Partitioning Around Medoids* (PAM).

Os autores utilizaram critérios internos de cada algoritmo para avaliar o desempenho de cada um e baseado nesses critérios selecionaram os agrupamentos obtidos através do algoritmo *K-means* para realizar a classificação. O resultado da análise de agrupamentos gerou três agrupamentos sendo que as características dos garotos do agrupamento 1 se distinguiam parcialmente do grupo de controle, do agrupamento 2 se distinguiam totalmente do grupo de controle e do agrupamento 3 não se distinguiam do grupo de controle.

Foram aplicadas os algoritmos SVM, *Multi Layer Perceptron* (MLP) e RF para avaliar a distinção entre os grupos. A classificação foi feita entre garotos com TEA de cada agrupamento e grupo de controle. Para o agrupamento 1, a precisão e sensibilidade da classificação foi de 90% e 93% usando SVM, 91% e 97% usando MLP e 81% e 100% usando RF. Para o agrupamento 2 a precisão e sensibilidade da classificação foi de 92% e 86% usando SVM, 100% e 86% usando MLP e 100% e 79% usando RF. E finalmente, para o agrupamento 3 a precisão e sensibilidade da classificação foi de 94% e 94% usando SVM, 94% e 94% usando MLP e 100% e 78% usando RF. A acurácia média geral foi de 91,94% com SVM, 93,55% com MLP e 88,71% com RF.

Gilani et al. (2015) executou uma análise em relação às características sexualmente dismórficas em adultos com TEA, também utilizando equipamento de estereofotogrametria. Em seu trabalho, mostrou que os indivíduos com um nível mais severo de TEA possuem menos características discriminantes entre os sexos.

Foram extraídas 23 medidas faciais e transformadas em características. Foi utilizado o algoritmo *Minimum Redundancy Maximum Relevance* (mRMR) para a seleção de características e foi feita classificação utilizando o algoritmo LDA. Para a validação do classificador foi utilizada validação cruzada 10-*fold*.

A partir das 23 distâncias, o algoritmo LDA identificou um subconjunto de seis distâncias que forneceram a maior acurácia de classificação por sexo sendo que para o grupo masculino a acurácia foi de 97,19% e para o grupo feminino a acurácia foi de 95,04%. Segundo os autores, as características mais discriminantes entre os sexos são largura da testa, largura entre os cantos externos dos olhos, comprimento do nariz, filtro e ponta do nariz, comprimento do filtro e largura do nariz.

Em todos os trabalhos estudados foi identificada a existência de diferenças na morfologia facial entre os casos e o grupo de controle. Dessa forma, o estudo das características antropométricas a partir de ferramentas computacionais permite a identificação desses padrões de características diferentes entre os grupos. A tabela 1 apresenta um resumo dos trabalhos encontrados.

Tabela 1 – Resumo dos trabalhos na literatura

Trabalho	Síndromes tratadas	Métodos de análise	Tipo de imagem	Resultados
Boehringer et al. (2006)	DiGeorge, Cri-du-chat, Cornelia de Lange, X frágil, Mucopolissacarídose tipo 3, Noonan, Prader-Willi, Smith-Lemli-Opitz, Sotos, Williams-Beuren	LDA, SVM, kNN, JV	Bidimensional com câmera comum	76% de acurácia com LDA
Zhao et al. (2014a)	Down, Klinefelter, Charge, alcoólica fetal, Goldenhar, Stickler, Treacher Collins, Noonan,Williams, DiGeorge, Trissomia 18, Beckwith Wiedemann, Rubinstein-Taybi, Waardenburg, Sotos	LDA, SVM, kNN, RF	Bidimensional sem especificação de câmera	96,7% de acurácia com SVM-RBF e LDA
Compton et al. (2007)	Esquizofrenia	Análise de agrupamentos Ward com distâncias euclidiana, Chebychev e City Block	Não se aplica	Não foi possível identificar a regra de pertencimento dos agrupamentos
Aldridge et al. (2011)	TEA	Análise de agrupamentos	Tridimensional com estereofotogrametria	Foram obtidos 3 agrupamentos
Ozgen et al. (2011)	TEA	Análise estatística	Não se aplica	Crianças com TEA apresentaram anomalias morfológicas maiores
Obafemi-Ajayi et al. (2015)	TEA	EM, SOM, K-Means, PAM, SVM, MLP, RF	Tridimensional com estereofotogrametria	3 agrupamentos com K-Means e acurácia média geral 93,55% com MLP
Gilani et al. (2015)	TEA	mRMR, LDA	Tridimensional com estereofotogrametria	Acurácia de 97,19% para o grupo masculino e 95,04% para o grupo feminino

4 Materiais e métodos

Como dito anteriormente, este projeto visa à construção de um classificador que, dada uma imagem facial de uma criança, classifique-a entre autista e normal, fornecendo assim um apoio ao diagnóstico. A seguir são descritos os materiais e métodos para cada objetivo específico a fim de alcançar esse objetivo geral.

4.1 Construção de uma base de imagens faciais de crianças afetadas por TEA e crianças em desenvolvimento típico

A fim de conhecer melhor o problema da antropometria e transtornos de desenvolvimento foi realizada a revisão da literatura descrita no Capítulo 3. Durante a revisão não foram encontradas bases de dados de imagens faciais de crianças com TEA que pudessem ser utilizadas neste trabalho.

Para a construção da base de dados de crianças com TEA e controles está sendo realizada a coleta de dados. Os voluntários do grupo com TEA vieram do ambulatório do Instituto de Psiquiatria do Hospital das Clínicas da Universidade São Paulo.

Para realizar a coleta foi definido um protocolo que ajudasse a padronizar da melhor forma as imagens. Todas as fotos estão sendo capturadas com a câmera compacta Coolpix L120 Nikon. A câmera é acoplada a um tripé que é fixado a uma distância de 50cm da cadeira onde a criança deve se sentar. A parede que serve de fundo às imagens é branca e foi fixada uma faixa vermelha de 50cm por 5cm que serve como referência para que se possa depois determinar as distâncias em unidade métrica e não trabalhar apenas com medidas em pixels.

O protocolo, juntamente com a especificação do projeto e os devidos termos de consentimento e assentimento que os voluntários devem assinar antes da coleta, foram submetidos à análise do Comitê de Ética em Pesquisa em Seres Humanos onde obtiveram parecer de aprovação.

O protocolo define que, para a coleta, as crianças devem ser instruídas a permanecer com a expressão o mais neutra possível. São capturadas cinco fotos faciais frontais além da gravação de um vídeo curto para que seja possível extrair *frames* estáticos. O vídeo só é utilizado caso não seja possível obter uma foto em que a expressão está neutra ou caso o rosto esteja em uma posição muito inclinada.

É esperado que o número de amostras de treinamento seja maior para crianças com o transtorno do que controles, pois as imagens serão obtidas de crianças que fazem parte do Programa de Transtorno do Espectro Autista (PROTEA) e o número de voluntários para o grupo de controle é reduzido. Dessa forma, será necessário buscar métodos que tratem o desbalanceamento entre as classes.

Além das imagens, também estão sendo coletadas outras informações como idade e pontuação na escala *Childhood Austim Rating Scale* (CARS) (SCHOPLER et al., 1980) que indica a gravidade do autismo.

Nos apêndices A e B estão descritos os termos de consentimento livre e esclarecido e termo de assentimento que foram submetidos ao Comitê de Ética.

4.2 Processamento das imagens coletadas para extração das medidas antropométricas faciais

Para cada criança participante da coleta há uma ou mais imagens frontais e possivelmente um vídeo. O primeiro passo consiste em selecionar a melhor imagem ou o melhor $frame$, considerando que a criança deva estar com uma expressão neutra e a imagem seja o mais frontal possível.

Para realizar o processamento das imagens coletadas nesta pesquisa foram utilizadas as bibliotecas e ferramentas OpenFace (BALTRU et al., 2016), OpenCV (BRADSKI, 2000) e Dlib (KING, 2009). Foi feita a escolha por estas bibliotecas e ferramentas por serem bem estabelecidas na literatura e por serem disponibilizadas sob licenças de uso gratuito para fins científicos e educacionais. No Capítulo 5 são apresentados os resultados e problemas encontrados durante a etapa de processamento das imagens. A ferramenta OpenFace está disponível em versão compilada como um arquivo de *batch* para sistema operacional Windows[1]. Neste trabalho está sendo utilizada a versão 0.2.2 64 bits.

Após a detecção da face a mesma deve ser rotacionada, pois apesar das imagens serem frontais podem existir pequenas inclinações na pose. A rotação será realizada da seguinte forma:

1. as pupilas são detectadas e é traçada uma reta entre elas;
2. é calculado o ângulo entre o eixo horizontal e a reta traçada;
3. a imagem é rotacionada com base no ângulo obtido no passo anterior.

[1] Disponível em https://github.com/TadasBaltrusaitis/OpenFace/wiki/Windows-Installation

Para extração dos pontos de controle está sendo utilizado o comando *FaceLandmarkImg*; também foi feito teste com o comando *FeatureExtraction*. As figuras 10 e 11 apresentam os comandos utilizados. A saída dos comandos é um arquivo texto com as coordenadas de 68 pontos de controle em pixels e se especificado o parâmetro -oi também é gravada uma imagem com os pontos marcados.

Figura 10 – Comando OpenFace *FaceLandmarkImg*

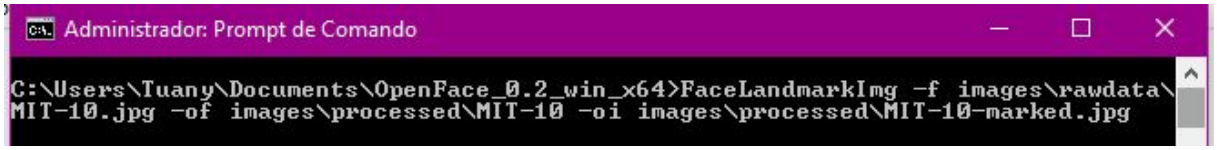
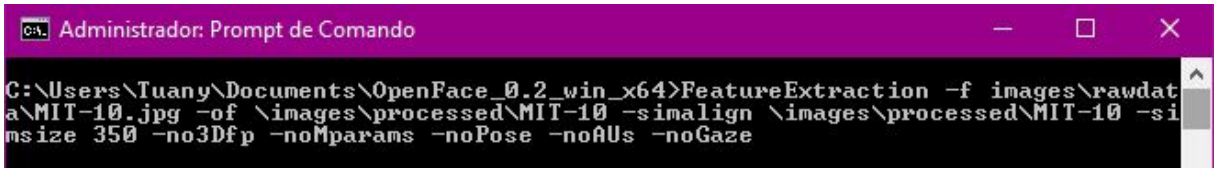


Figura 11 – Comando OpenFace *FeatureExtraction*



A partir dos pontos de controle extraídos das imagens será necessário calcular as distâncias, porém os pontos extraídos são coordenadas de uma imagem digital e as distâncias precisarão ser convertidas para sistema métrico. Existem algumas formas de realizar essa conversão. A biblioteca OpenFace oferece um método que realiza a estimativa dos pontos em um espaço tridimensional, porém este método não será utilizado porque se baseia em coordenadas globais estimadas.

Para resolver essa conversão de medidas será utilizada uma faixa de referência no fundo da imagem. Esta faixa tem tamanho conhecido e foi posicionada na parede de fundo onde as imagens são capturadas. Dessa forma, é possível encontrar as distâncias entre os pontos por meio de um cálculo de proporção e ao mesmo tempo normalizar as imagens de interesse tendo como referência esse objeto. O cálculo das distâncias será feito conforme descrito a seguir:

1. são extraídos os pontos de controle da imagem;
2. são calculadas as distâncias entre os pontos em pixels;
3. a faixa de referência é detectada na imagem e é determinado seu tamanho em pixels;

4. é feita a conversão das distâncias proporcionalmente em relação à faixa de referência de tamanho conhecido.

Está sendo analisada a possibilidade de converter essas distâncias em z-scores calculados tomando como base curvas antropométricas que descrevem certas medidas em função da idade da criança. As curvas disponíveis em BRUNONI e PEREZ (2013) foram traçadas com dados levantados pela equipe de genética médica do Hospital São Paulo da UNIFESP, e são discriminadas por sexo para diversas medidas do corpo. Assim, para uma dada criança de idade i e sexo s, o z-score de uma dada medida de distância d será calculado com base na curva antropométrica daquela medida para o sexo s, que fornecerá o valor médio (μ_i) e o desvio padrão (σ_i) daquela medida para a idade i:

$$z = \frac{d - \mu_i}{\sigma_i}.$$

Uma limitação dessa abordagem é que parece não haver dados para muitas distâncias da antropometria facial.

4.3 Pesquisa e aplicação de métodos de seleção de característica

Está sendo executada uma revisão da literatura a fim de conhecer com profundidade os métodos de seleção de características para problemas de classificação de imagem. Também será realizada a comparação de métodos de seleção de características mais promissores e em seguida, os métodos escolhidos serão aplicados e comparados.

4.4 Treinamento e comparação de classificadores

Serão utilizados e comparados alguns métodos de classificação supervisionada, a princípio SVM e *Random Forest* para o treinamento de classificadores com o fim proposto. É esperado que o número de amostras de treinamento seja maior para crianças com transtorno do que controles. Logo, um tratamento para dados desbalanceados também poderá ser necessário.

Para cada um dos métodos de classificação selecionados para comparação, será feita uma escolha de seus parâmetros, via validação cruzada, para melhor discriminarem os dados de treinamento. Deverá ser aplicado, nos métodos, um tratamento para classes

desbalanceadas. Particularmente para *Random Forests*, pretende-se analisar a utilização do algoritmo *Dynamic Discriminant Bounds Tree* (DDBT) (FRIZZARINI; LAURETTO, 2013), que tem como objetivo a indução de árvores de classificação para tratamento de dados desbalanceados. Testes preliminares indicaram boa adaptação do DDBT a problemas de conjuntos de dados desbalanceados, que em geral produz classificadores com bom equilíbrio entre as taxas de erro de classificação nas classes. Assim, esse algoritmo também será considerado em nossos experimentos.

Será utilizado o método de validação cruzada para avaliação dos classificadores. Os métodos serão comparados utilizando-se o melhor classificador obtido. Ou seja, para cada método será considerado, para fins de comparação com o outro método, apenas os resultados oriundos do classificador obtido com o conjunto de parâmetros associados com o menor erro (na etapa de escolha de parâmetros via validação cruzada). Para essa comparação serão utilizadas as taxas de erro, precisão e sensibilidade. A taxa de erro é calculada somando o número de falsos positivos com o número de falsos negativos e dividindo pelo total da amostra. A precisão é obtida dividindo o número de verdadeiros positivos pelo total de preditos positivos. A sensibilidade é calculada dividindo o número de verdadeiros positivos pela soma de verdadeiros positivos e falsos negativos.

4.5 Plano de Trabalho e Cronograma

A seguir estão descritas as atividades que já foram concluídas, as que estão em desenvolvimento e também as que serão realizadas futuramente. A tabela 2 apresenta o cronograma dessas atividades.

1. **Revisão da literatura:** nesta atividade foi realizada uma revisão da literatura para identificar o estado da arte quanto a análise de síndromes transtornos por meio de antropometria. Os trabalhos encontrados foram descritos no Capítulo 2.
2. **Delimitação do problema:** nesta atividade foi realizado o estudo das estratégias que poderiam ser utilizadas para resolver o problema e também a delimitação do escopo do projeto.
3. **Avaliação pelo comitê de ética:** nesta atividade o projeto de pesquisa foi redigido e apresentado para a aprovação do comitê de ética e obteve parecer de aprovação.

4. **Coleta de dados:** nesta atividade está sendo realizada a coleta de imagens de crianças autistas (casos) e em desenvolvimento típico (controles) em parceria com o Instituto de Psiquiatria (IPq) do Hospital das Clínicas da Universidade de São Paulo (HCFMUSP).
5. **Processamento das imagens:** nesta atividade estão sendo realizados testes de extração dos pontos de controle das imagens obtidas. Nessa etapa foi escolhida a ferramenta OpenFace para extração dos pontos de controle e definido como será realizado o cálculo das distâncias entre os pontos. Essa etapa se prolonga até o fim da coleta, já que todas as imagens coletas devem ser processadas.
6. **Definição dos métodos de seleção de características:** nesta atividade uma segunda revisão da literatura sobre algoritmos de seleção de características está sendo executada. A intenção é conhecer o estado da arte e escolher os métodos que tenham desempenho eficiente e sejam promissores. Em seguida serão testados os métodos escolhidos.
7. **Estudo e aplicação dos métodos de classificação:** nesta atividade os métodos de classificação serão estudados e serão treinados os modelos. Em seguida, para cada modelo será realizada a escolha dos parâmetros que melhor discriminam as imagens por meio de validação cruzada.
8. **Avaliação dos classificadores:** Para cada método será considerado, para fins de comparação com os demais métodos, apenas os resultados oriundos do classificador obtido com o conjunto de parâmetros associados com o menor erro.
9. **Análise dos resultados e divulgação:** Com os resultados e análises documentadas será escrito um artigo para divulgação em conferências e periódicos da área e será feita a composição do relatório final e da dissertação do mestrado.

Tabela 2 – Cronograma das tarefas passadas e futuras

Cronograma																									
	2015					**2016**												**2017**							
Etapas	**8**	**9**	**10**	**11**	**12**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1																									
2																									
3																									
4																									
5																									
6																									
7																									
8																									
9																									

5 Resultados preliminares

Neste capítulo são apresentados os resultados preliminares do processamento realizado com as imagens até o momento. Também são descritas as dificuldades encontradas.

5.1 Protocolo para coleta de dados

Na seção 4.1 foi descrito o protocolo de dados utilizado para a captura de imagens. O protocolo inicial não continha a faixa de referência, pois inicialmente a normalização das medidas seria feita em relação as proporções do próprio rosto da criança, no caso seriam utilizadas as distâncias entre as extremidades do rosto.

Porém, este método de normalização não pareceu o mais adequado, pois a ideia é comparar as medidas antropométricas em medida métrica para identificar as diferenças entre as medidas de uma criança em desenvolvimento típico e uma criança com TEA e não utilizar uma medida relativa. Além disso, cada criança tem um tamanho de rosto diferente com diferenças de largura e comprimento. Por estes motivos, foi decidido que a normalização será realizada com base em um objeto de referência no fundo da imagem. As imagens coletadas sem a faixa de referência poderão ser descartadas neste trabalho, exceto as imagens em que existe outro objeto de referência na imagem que poderá ser, por exemplo, a cadeira caso as duas extremidades laterais da mesma esteja visível na imagem.

A coleta de dados ainda não foi encerrada, porém já foram coletadas imagens de 70 crianças com TEA com idade entre 4 e 17 anos e 8 controles. Dessas 70 crianças, 37 foram fotografadas com a faixa de referência ao fundo e todos os controles foram fotografados depois da inclusão da faixa no protocolo.

Apesar de que neste trabalho algumas imagens poderão ser descartadas, foi possível melhorar o protocolo de coleta e os dados coletados. A coleta ainda está sendo realizada de forma a obter o maior número de imagens possíveis para a construção do banco de dados. O prolongamento foi necessário devido à dificuldade em encontrar voluntários em desenvolvimento típico.

5.2 Extração dos pontos de controles

De início foram realizados testes com o método Viola-Jones, descrito na seção 2.3.1. As figuras 12 e 13 mostram os resultados da extração dos pontos de controle utilizando a implementação do método Viola-Jones e *Haar cascades* disponíveis na biblioteca OpenCV (BRADSKI, 2000). As imagens usadas para exemplificar os testes são da base *CBCL Face database* (SUNG, 1996) e estão disponíveis gratuitamente para fins educacionais e de pesquisa.

Figura 12 – Exemplo da extração Viola-Jones: pontos da boca não foram detectados

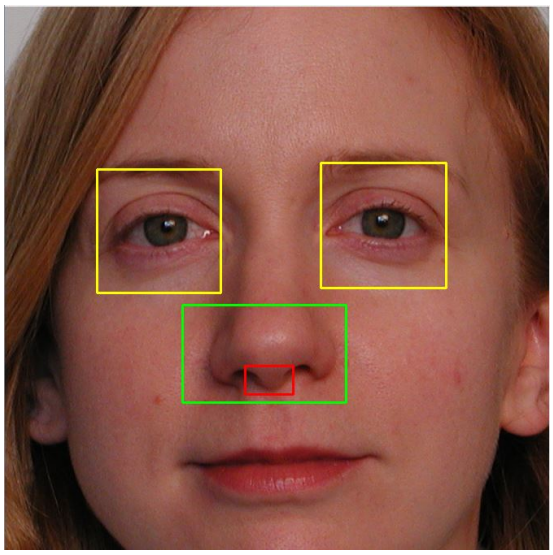
Na imagem 12 pode-se observar que os pontos da região da boca não foram encontrados de forma adequada, porém na imagem 13 notamos que alguns dos pontos foram encontrados. Como explicado na seção 2.3.1, o método Viola-Jones utiliza arquivos para descrever os as características Haar de cada região que deve ser identificada na imagem. Porém essas características podem ser encontradas em outras partes da imagem que se assemelhem a parte que se deseja detectar.

Na figura 14 pode-se observar que qualquer objeto na imagem que seja semelhante a curvatura da boca foi detectado. Na tentativa de contornar este problema outros classificadores Haar cascade para boca foram testados, porém nenhum demonstrou ser eficiente na detecção dos pontos da boca. Os classificadores Haar das outras regiões do rosto detectaram os pontos de forma adequada sendo que apenas o classificador da

Figura 13 – Exemplo da extração Viola-Jones: alguns pontos da boca foram detectados

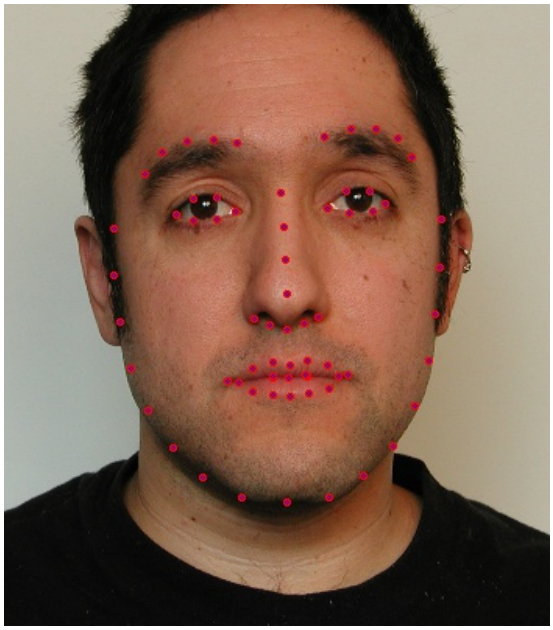
região da boca não apresentou resultados eficientes. De forma a resolver esta dificuldade foi cogitado treinar um classificador Haar para a boca. Apesar de ser possível treinar o classificador Haar essa alternativa é custosa e demorada, pois são necessários muitas imagens de exemplos positivos e negativos. Além de que seriam necessários testes para determinar se os pontos estão sendo adequadamente detectados. Portanto, essa solução foi descartada e outro método de extração dos pontos de controle foi testado.

Figura 14 – Regiões detectadas pelo Viola-Jones: os retângulos vermelhos indicam as regiões detectadas como boca, os retângulos verdes indicam as regiões detectadas como nariz e os retângulos amarelos indicam as regiões detectadas como olhos.



A figura 15 mostra o resultado da extração dos pontos de controle utilizando o método CLNF descrito na seção 2.3.2. Além de ter uma precisão maior na detecção dos pontos da região da boca o CLNF consegue detectar um total de 68 pontos e dessa forma têm-se mais distâncias e pontos de controle que podem ser analisados e comparados. O método CLNF utilizado faz parte de uma ferramenta *opensource* de visão computacional e análise de de comportamento facial criada por pesquisadores do laboratório de computação da Universidade Carnegie Mellon.

Figura 15 – Exemplo da extração CLNF: todos os pontos foram detectados



5.3 Cálculo das distâncias

A partir dos pontos de controle extraídos das imagens, para o objetivo proposto neste estudo, é necessário realizar o cálculo das distâncias faciais. Existem algumas formas de realizar este cálculo como distâncias euclidianas ou geodésicas. Para o cálculo das distâncias geodésicas são necessárias imagens de vários ângulos ou escaneamento do objeto, como é realizado no trabalho de Obafemi-Ajayi et al. (2015).

Neste estudo são utilizadas imagens bidimensionais, portanto será realizado o cálculo das distâncias euclidianas. Para isso, será necessário implementar um método de triangulação para determinar as distâncias não em pixels, mas em centímetros ou milímetros. Por este motivo, foi acrescentada a faixa de referência na captura das imagens conforme explicado na seção 5.1.

5.4 Dificuldades encontradas

Por se tratar de uma pesquisa relacionada a área de saúde uma das principais dificuldades encontradas foi o recrutamento de indivíduos para o grupo de controle. Por este motivo, é esperado que existam mais dados de crianças com TEA do que em desenvolvimento típico. A coleta está sendo conduzida no IPq e a localização também é um fator dificultante no recrutamento.

Outra dificuldade é a conscientização da população em relação à pesquisa e ao diagnóstico do transtorno. Muitos participantes do estudo não tinham diagnóstico formal e foi necessário realizar uma fase de triagem para confirmar se as crianças possuíam TEA ou não. Outro fato importante é que algumas crianças podem ser portadores de mais de uma síndrome além do TEA. Nos casos em que a criança possuía mais de uma síndrome, os dados foram excluídos.

O protocolo estabelecido para a coleta de dados também apresentou desafios na sua elaboração. Por se tratar de captura de imagens de crianças foram necessárias algumas adaptações do protocolo pensado inicialmente para garantir a qualidade dos dados e padronização das imagens.

6 Considerações Finais

Neste documento foi apresentado o problema de auxílio ao diagnóstico de autismo com base nas características antropométricas faciais. Foi apresentada a revisão da literatura na qual foram encontrados dois tipos de trabalho: os que propõem analisar a morfologia facial de indivíduos afetados por síndromes e transtornos e os que propõem métodos que realizam classificação da imagem entre as síndromes.

Os trabalhos que estão relacionados a TEA limitam-se a analisar as diferenças entre as distâncias ou correlacionar a gravidade do transtorno com os dismorfismos em crianças com TEA, como em Aldridge et al. (2011), Gilani et al. (2015), Obafemi-Ajayi et al. (2015). Nestes trabalhos foi usado um equipamento de estereofotogrametria que tem um custo alto, o que dificulta o seu acesso.

Esta pesquisa se propõe a comparar técnicas de seleção de características e classificação em imagens faciais bidimensionais, capturadas por uma câmera fotográfica comum, para separar as imagens em casos de TEA e controles. Sendo assim, o trabalho proposto poderá causar impactos científicos, sociais e econômicos, pois atualmente no Brasil não existem métodos computacionais ou equipamentos além dos questionários supracitados que não são totalmente adaptados para auxiliar o diagnóstico de autismo. Existem poucos profissionais e hospitais especializados para a realização deste diagnóstico que muitas vezes é realizado tardiamente (BOSA, 2006; GOMES et al., 2015). A depender dos resultados obtidos, o classificador resultante poderá ser adaptado para diagnosticar outros tipos de síndromes e transtornos baseando-se em imagens faciais obtidas com câmeras digitais simples.

Referências[1]

ALDRIDGE, K. et al. Precision and error of three-dimensional phenotypic measures acquired from 3dmd photogrammetric images. American Journal of Medical Genetics Part A, Wiley Online Library, v. 138, n. 3, p. 247–253, 2005. Citado 2 vezes nas páginas 13 e 30.

ALDRIDGE, K. et al. Facial phenotypes in subgroups of prepubertal boys with autism spectrum disorders are correlated with clinical phenotypes. Molecular autism, BioMed Central, v. 2, n. 1, p. 15, 2011. Citado 7 vezes nas páginas 12, 13, 32, 33, 34, 36 e 48.

APA. Manual diagnóstico e Estatístico de Transtornos Mentais-: DSM-5. [S.l.]: Artmed Editora, 2014. Citado na página 11.

BAIO, J. Prevalence of autism spectrum disorders: Autism and developmental disabilities monitoring network, 11 sites, united states, 2010. morbidity and mortality weekly report. surveillance summaries. volume 63, number 2. Centers for Disease Control and Prevention, ERIC, 2010. Citado 2 vezes nas páginas 11 e 12.

BALTRU, T. et al. Openface: an open source facial behavior analysis toolkit. In: IEEE. 2016 IEEE Winter Conference on Applications of Computer Vision (WACV). [S.l.], 2016. p. 1–10. Citado na página 38.

BALTRUSAITIS, T.; ROBINSON, P.; MORENCY, L.-P. Constrained local neural fields for robust facial landmark detection in the wild. In: The IEEE International Conference on Computer Vision (ICCV) Workshops. [S.l.: s.n.], 2013. Citado 2 vezes nas páginas 23 e 24.

BOEHRINGER, S. et al. Syndrome identification based on 2d analysis software. European Journal of Human Genetics, Nature Publishing Group, v. 14, n. 10, p. 1082–1089, 2006. Citado 3 vezes nas páginas 13, 30 e 36.

BOSA, C. A. Autismo: intervenções psicoeducacionais autism: psychoeducational intervention. Rev bras psiquiatr, SciELO Brasil, v. 28, n. Supl I, p. S47–53, 2006. Citado na página 48.

BRADSKI, G. Opencv library. Dr. Dobb's Journal of Software Tools, 2000. Citado 2 vezes nas páginas 38 e 44.

BRUNONI, D.; PEREZ, A. B. A. Guia de genética médica: Série Guias de medicina ambulatorial e hospitalar da EPM–UNIFESP. [S.l.]: São Paulo: Manole, 2013. Citado na página 40.

CARDIE, C. Using decision trees to improve case-based learning. In: Proceedings of the tenth international conference on machine learning. [S.l.: s.n.], 1993. p. 25–32. Citado na página 28.

COLVERT, E. et al. Heritability of autism spectrum disorder in a uk population-based twin sample. JAMA psychiatry, American Medical Association, v. 72, n. 5, p. 415–423, 2015. Citado na página 12.

[1] De acordo com a Associação Brasileira de Normas Técnicas. NBR 6023.

COMPTON, M. T. et al. Facial measurement differences between patients with schizophrenia and non-psychiatric controls. Schizophrenia research, Elsevier, v. 93, n. 1, p. 245–252, 2007. Citado 2 vezes nas páginas 31 e 36.

CORSELLO, C. M. Early intervention in autism. Infants & Young Children, LWW, v. 18, n. 2, p. 74–85, 2005. Citado na página 14.

DASH, M.; LIU, H. Feature selection for classification. Intelligent data analysis, Elsevier, v. 1, n. 1, p. 131–156, 1997. Citado na página 25.

DAWSON, G. Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. Development and psychopathology, Cambridge Univ Press, v. 20, n. 03, p. 775–803, 2008. Citado na página 14.

DAWSON, G. et al. Randomized, controlled trial of an intervention for toddlers with autism: the early start denver model. Pediatrics, Am Acad Pediatrics, v. 125, n. 1, p. e17–e23, 2010. Citado na página 14.

DEMYER, W.; ZEMAN, W.; PALMER, C. G. The face predicts the brain: diagnostic significance of median facial anomalies for holoprosencephaly (arhinencephaly). Pediatrics, Am Acad Pediatrics, v. 34, n. 2, p. 256–263, 1964. Citado na página 12.

DETH, R. et al. How environmental and genetic factors combine to cause autism: A redox/methylation hypothesis. Neurotoxicology, Elsevier, v. 29, n. 1, p. 190–201, 2008. Citado na página 12.

DEUTSCH, C. K. et al. The farkas system of craniofacial anthropometry: methodology and normative databases. In: Handbook of Anthropometry. [S.l.]: Springer, 2012. p. 561–573. Citado na página 17.

DOAK, J. An evaluation of feature selection methods and their application to computer security. [S.l.]: University of California, Computer Science, 1992. Citado na página 28.

DUDA, R. O.; HART, P. E. et al. Pattern classification and scene analysis. [S.l.]: Wiley New York, 1973. v. 3. Citado na página 25.

FARKAS, L. G. Anthropometry of the Head and Face. [S.l.]: Raven Pr, 1994. Citado 2 vezes nas páginas 17 e 18.

FARKAS, L. G. et al. Differences between direct (anthropometric) and indirect (cephalometric) measurements of the skull. Journal of Craniofacial Surgery, LWW, v. 13, n. 1, p. 105–108, 2002. Citado 2 vezes nas páginas 12 e 17.

FREUND, Y.; SCHAPIRE, R. E. A desicion-theoretic generalization of on-line learning and an application to boosting. In: SPRINGER. European conference on computational learning theory. [S.l.], 1995. p. 23–37. Citado na página 23.

FRIZZARINI, C.; LAURETTO, M. S. Proposta de um algoritmo para indução de árvores de classificação para dados desbalanceados. 2013. Citado na página 41.

GILANI, S. Z. et al. Sexually dimorphic facial features vary according to level of autistic-like traits in the general population. Journal of neurodevelopmental disorders, BioMed Central Ltd, v. 7, n. 1, p. 14, 2015. Citado 3 vezes nas páginas 34, 36 e 48.

GOMES, P. et al. Autism in brazil: a systematic review of family challenges and coping strategies. Jornal de pediatria, SciELO Brasil, v. 91, n. 2, p. 111–121, 2015. Citado 2 vezes nas páginas 14 e 48.

GONZALEZ, R. C.; WOODS, R. E. Digital image processing. [S.l.]: Prentice-Hall, 2 edition, 2002. Citado na página 18.

HAMMOND, P. et al. 3d analysis of facial morphology. American journal of medical genetics Part A, Wiley Online Library, v. 126, n. 4, p. 339–348, 2004. Citado na página 13.

HAMMOND, P.; SUTTIE, M. Large-scale objective phenotyping of 3d facial morphology. Human mutation, Wiley Online Library, v. 33, n. 5, p. 817–825, 2012. Citado na página 13.

HSU, R.-L.; ABDEL-MOTTALEB, M.; JAIN, A. K. Face detection in color images. IEEE transactions on pattern analysis and machine intelligence, IEEE, v. 24, n. 5, p. 696–706, 2002. Citado na página 20.

HUGHES, G. On the mean accuracy of statistical pattern recognizers. IEEE transactions on information theory, IEEE, v. 14, n. 1, p. 55–63, 1968. Citado na página 27.

JOHNSON, C. P.; MYERS, S. M. et al. Identification and evaluation of children with autism spectrum disorders. Pediatrics, Am Acad Pediatrics, v. 120, n. 5, p. 1183–1215, 2007. Citado na página 11.

KING, D. E. Dlib-ml: A machine learning toolkit. Journal of Machine Learning Research, v. 10, p. 1755–1758, 2009. Citado na página 38.

KIRA, K.; RENDELL, L. A. The feature selection problem: Traditional methods and a new algorithm. In: AAAI. [S.l.: s.n.], 1992. v. 2, p. 129–134. Citado na página 28.

LEVY, S. E.; MANDELL, D. S.; SCHULTZ, R. T. Autism. Lancet, v. 374, p. 1627–1638, 2009. Citado na página 11.

LORD, C. et al. Autism spectrum disorders. Neuron, Elsevier, v. 28, n. 2, p. 355–363, 2000. Citado na página 11.

MANDELL, D. S.; NOVAK, M. M.; ZUBRITSKY, C. D. Factors associated with age of diagnosis among children with autism spectrum disorders. Pediatrics, Am Acad Pediatrics, v. 116, n. 6, p. 1480–1486, 2005. Citado na página 11.

MCPARTLAND, J. C.; REICHOW, B.; VOLKMAR, F. R. Sensitivity and specificity of proposed dsm-5 diagnostic criteria for autism spectrum disorder. Journal of the American Academy of Child & Adolescent Psychiatry, Elsevier, v. 51, n. 4, p. 368–383, 2012. Citado na página 12.

MILES, J. et al. Head circumference is an independent clinical finding associated with autism. American journal of medical genetics, Wiley Online Library, v. 95, n. 4, p. 339–350, 2000. Citado na página 12.

MITCHELL, T. M. Machine learning. New York, 1997. Citado na página 25.

MUHLE, R.; TRENTACOSTE, S. V.; RAPIN, I. The genetics of autism. Pediatrics, Am Acad Pediatrics, v. 113, n. 5, p. e472–e486, 2004. Citado na página 12.

NARENDRA, P. M.; FUKUNAGA, K. A branch and bound algorithm for feature subset selection. IEEE Transactions on Computers, IEEE, v. 100, n. 9, p. 917–922, 1977. Citado na página 28.

OBAFEMI-AJAYI, T. et al. Facial structure analysis separates autism spectrum disorders into meaningful clinical subgroups. Journal of autism and developmental disorders, Springer, v. 45, n. 5, p. 1302–1317, 2015. Citado 4 vezes nas páginas 34, 36, 46 e 48.

OZGEN, H. et al. Morphological features in children with autism spectrum disorders: a matched case–control study. Journal of autism and developmental disorders, Springer, v. 41, n. 1, p. 23–31, 2011. Citado 2 vezes nas páginas 33 e 36.

PAPAGEORGIOU, C. P.; OREN, M.; POGGIO, T. A general framework for object detection. In: IEEE. Computer vision, 1998. sixth international conference on. [S.l.], 1998. p. 555–562. Citado na página 20.

PAULA, C. S. et al. Brief report: prevalence of pervasive developmental disorder in brazil: a pilot study. Journal of Autism and Developmental Disorders, Springer, v. 41, n. 12, p. 1738–1742, 2011. Citado na página 11.

RICE, C. Prevalence of autism spectrum disorders: Autism and developmental disabilities monitoring network, united states, 2006. morbidity and mortality weekly report. surveillance summaries. volume 58, number ss-10. Centers for Disease Control and Prevention, ERIC, 2009. Citado na página 11.

RODIER, P. M.; BRYSON, S. E.; WELCH, J. P. Minor malformations and physical measurements in autism: data from nova scotia. Teratology, Wiley Online Library, v. 55, n. 5, p. 319–325, 1997. Citado na página 12.

SCHOPLER, E. et al. Toward objective classification of childhood autism: Childhood autism rating scale (cars). Journal of autism and developmental disorders, Springer, v. 10, n. 1, p. 91–103, 1980. Citado na página 38.

SUNG, K.-K. Learning and Example Selection for Object and Pattern Recognition. Tese (Doutorado) — MIT, Artificial Intelligence Laboratory and Center for Biological and Computational Learning, Cambridge, MA, 1996. Citado na página 44.

TOU, J. T.; GONZALEZ, R. C. Pattern recognition principles. Pattern Recognition in Physics, v. 1, 1974. Citado na página 13.

VAFAIE, H.; IMAM, I. F. Feature selection methods: genetic algorithms vs. greedy-like search. In: Proceedings of the International Conference on Fuzzy and Intelligent Control Systems. [S.l.: s.n.], 1994. v. 51. Citado na página 28.

VIOLA, P.; JONES, M. Rapid object detection using a boosted cascade of simple features. In: IEEE. Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. [S.l.], 2001. v. 1, p. I–511. Citado 3 vezes nas páginas 20, 21 e 23.

WEINBERG, S. M. et al. Anthropometric precision and accuracy of digital three-dimensional photogrammetry: comparing the genex and 3dmd imaging systems with one another and with direct anthropometry. Journal of Craniofacial Surgery, LWW, v. 17, n. 3, p. 477–483, 2006. Citado na página 17.

XU, L.; YAN, P.; CHANG, T. Best first strategy for feature selection. In: IEEE. Pattern Recognition, 1988., 9th International Conference on. [S.l.], 1988. p. 706–708. Citado na página 28.

ZHAO, Q. et al. Digital facial dysmorphology for genetic screening: Hierarchical constrained local model using ica. Medical image analysis, Elsevier, v. 18, n. 5, p. 699–710, 2014. Citado 3 vezes nas páginas 13, 31 e 36.

ZHAO, Q. et al. Ensemble learning for the detection of facial dysmorphology. In: IEEE. 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. [S.l.], 2014. p. 754–757. Citado na página 30.

ZWAIGENBAUM, L.; BRYSON, S.; GARON, N. Early identification of autism spectrum disorders. Behavioural Brain Research, Elsevier, v. 251, p. 133–146, 2013. Citado na página 11.

Apêndices

Apêndice A – Termo de consentimento livre e esclarecido

Eu, _____, estudante do programa de pós-graduação em Sistemas de Informação da EACH-USP, estou realizando uma pesquisa intitulada de Classificação de imagens faciais para diagnóstico de autismo, sob a orientação de Profª Drª Ariane Machado Lima, docente do Curso de Pós-graduação em Sistemas de Informação. Essa pesquisa tem como objetivos:

Realizar o estudo comparativo de métodos computacionais de classificação utilizando imagens faciais de crianças autistas e em desenvolvimento típico. Esse estudo comparativo permitirá construir um protótipo de ferramenta que auxiliará médicos especialistas a realizarem o diagnóstico de autismo. Para tanto, gostaria que você participasse desta pesquisa, na qual me comprometo a seguir a Resolução CNS 96/1996, relacionada à Pesquisa com Seres Humanos, respeitando o seu direito de:

1. ter liberdade de participar ou deixar de participar do estudo, sem que isso lhe traga algum prejuízo ou risco;
2. manter o seu nome em sigilo absoluto, sendo que o que disser não lhe resultará em qualquer dano à sua integralidade;
3. interromper a participação na pesquisa caso se sinta incomodado (a) com a mesma;
4. responder as questões levantadas pela pesquisadora caso seja solicitado (a) para uma entrevista, para a qual será marcado um local na instituição, horário e data em que possa se sentir mais confortável;
5. garantia de receber uma resposta a alguma dúvida durante ou após a entrevista.

Esclareço-lhe que ao participar dessa pesquisa, há risco mínimo de cansaço visual devido a iluminação ou flash utilizados para realiza a captura das imagens.

As imagens faciais serão capturadas com uma câmera fotográfica alocada a um tripé e o participante deverá estar em posição indicada pelo pesquisador(a) na sala. Também será gravado um vídeo de no máximo 15 segundos, em posição frontal, perfil direito e perfil esquerdo. O responsável pelo participante poderá acompanhá-lo(a) durante toda a pesquisa. É assegurada a assistência ao participante e ao responsável durante toda a pesquisa, bem como é garantido ao responsável o livre acesso a todas as informações e esclarecimentos adicionais sobre o estudo e suas consequências. Esclareço-lhe ainda que o tempo estimado de sua participação será de 10 a 25 minutos.

A sua participação nesta pesquisa ajudará a construção de uma ferramenta que auxiliará os médicos especialistas a realizarem o diagnóstico de autismo. Também permitirá que sejam identificadas técnicas computacionais que auxiliem o diagnóstico.

Este Termo de Consentimento será emitido em duas vias, sendo que uma via ficará em poder do pesquisador e a outra em poder do participante. Deixo e-mail para contato: tuany.pinheiro@usp.br – Tuany Dias Pinheiro e o endereço do CEP-EACH-USP no rodapé deste impresso, para que possa obter mais esclarecimentos ou informações sobre o estudo e sua participação.

Grato (a) pela atenção

_____	_____
Assinatura do (a) pesquisador (a)	Assinatura do (a) orientador (a)

Declaro que fui suficientemente informado a respeito do estudo Classificação de imagens faciais para diagnóstico de autismo. Eu discuti as informações acima com o Pesquisador Responsável Tuany Dias Pinheiro ou pessoa(s) por ele delegada(s) (_____) sobre a minha decisão em participar nesse estudo. Ficaram claros para mim os objetivos, os procedimentos, os potenciais desconfortos e riscos e as garantias. Concordo voluntariamente em participar deste estudo, assino este termo de consentimento e recebo uma via rubricada pelo pesquisador.

São Paulo, ____________.

_____

Assinatura do sujeito de pesquisa (ou responsável legal)

Apêndice B – Termo de assentimento

Você está sendo convidado(a) como voluntário(a) a participar da pesquisa Classificação de imagens faciais para diagnóstico de autismo. Neste estudo pretendemos realizar o estudo comparativo de métodos computacionais de classificação utilizando imagens faciais de crianças autistas e em desenvolvimento típico. Esse estudo comparativo permitirá construir um protótipo de ferramenta que auxiliará médicos especialistas a realizarem o diagnóstico de autismo.

Para este estudo adotaremos o(s) seguinte(s) procedimento(s): As imagens faciais serão capturadas com uma câmera fotográfica alocada a um tripé e o participante deverá estar em posição indicada pelo pesquisador(a) na sala. Também serão gravados três vídeos de no máximo 15 segundos cada, um em posição frontal, um em posição de perfil direito e um em posição de perfil esquerdo. O participante poderá permancecer no colo do responsável se assim desejar.

Para participar deste estudo, o seu responsável deverá autorizar e assinar um termo de consentimento. Não existirá nenhum custo, nem qualquer vantagem financeira. O seu responsável será esclarecido(a) em qualquer aspecto que desejar e estará livre para consentir sua participação ou não. O responsável poderá retirar o consentimento ou interromper a sua participação a qualquer momento. A sua participação é voluntária e a recusa em participar não acarretará qualquer penalidade ou modificação na forma em que é atendido(a) pelo pesquisador que irá tratar a sua identidade com padrões profissionais de sigilo. A sua identificação não ocorrerá em nenhuma publicação. Para este estudo há um risco mínimo de cansaço visual devido a iluminação e flash utilizados para a captura das imagens.

Os resultados estarão à sua disposição quando finalizada. Seu nome ou o material que indique sua participação não será liberado sem a permissão do responsável legal. Os dados e instrumentos utilizados na pesquisa ficarão arquivados com o pesquisador responsável. Este termo de assentimento encontra-se impresso em duas vias, sendo que uma cópia será arquivada pelo pesquisador responsável, e a outra será fornecida ao seu responsável.

Em caso de dúvidas com respeito aos aspectos éticos deste estudo, deixo e-mail para contato: tuany.pinheiro@usp.br – Tuany Dias Pinheiro e o endereço do CEP-EACH-USP no rodapé deste impresso, para que possa obter mais esclarecimentos ou informações sobre o estudo e sua participação.

Eu, _____, portador(a) do documento de identidade _____ (se já tiver documento), fui informado(a) dos objetivos do presente estudo de maneira clara e detalhada e esclareci minhas dúvidas. Sei que a qualquer momento poderei solicitar novas informações, e o meu responsável poderá modificar a decisão de participar se assim o desejar. Tendo o consentimento do meu responsável já assinado, declaro que concordo em participar desse estudo. Recebi uma cópia deste termo assentimento e me foi dada a oportunidade de ler e esclarecer as minhas dúvidas.

São Paulo, ____________.

_____	_____
Assinatura do(a) menor	Assinatura do(a) pesquisador(a)