UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ARTES, CIÊNCIAS E HUMANIDADES
PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS DE INFORMAÇÃO

TUANY DIAS PINHEIRO

**Classificação de imagens faciais para o auxílio ao diagnóstico do transtorno
do espectro autista**

São Paulo

2018

TUANY DIAS PINHEIRO

**Classificação de imagens faciais para o auxílio ao diagnóstico do transtorno
do espectro autista**

Dissertação apresentada à Escola de Artes,
Ciências e Humanidades da Universidade de
São Paulo 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

Versão corrigida contendo as alterações
solicitadas pela comissão julgadora em 27 de
março de 2018. A versão original encontra-se
em acervo reservado na Biblioteca da
EACH-USP e na Biblioteca Digital de Teses
e Dissertações da USP (BDTD), de acordo
com a Resolução CoPGr 6018, de 13 de
outubro de 2011.

Orientador: Prof. Dr^a. Arianne Machado Lima

São Paulo

2018

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

CATALOGAÇÃO-NA-PUBLICAÇÃO

(Universidade de São Paulo. Escola de Artes, Ciências e Humanidades. Biblioteca)
CRB-8 4936

Pinheiro, Tuany Dias

Classificação de imagens faciais para o auxílio ao diagnóstico do transtorno do espectro autista / Tuany Dias Pinheiro ; orientadora, Ariane Machado Lima. – 2018.

89 p. : il

Dissertação (Mestrado em Ciências) - Programa de Pós-Graduação em Sistemas de Informação, Escola de Artes, Ciências e Humanidades, Universidade de São Paulo.

Versão corrigida

1. Reconhecimento de padrões. 2. Inteligência artificial. 3. Percepção da face. 4. Autismo. I. Lima, Ariane Machado, orient. II. Título.

CDD 22 ed.- 006.4

Dissertação de autoria de Tuany Dias Pinheiro, sob o título “**Classificação de imagens faciais para o auxílio ao diagnóstico do transtorno do espectro autista**”, apresentada à Escola de Artes, Ciências e Humanidades da Universidade de São Paulo, 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, aprovada em 27 de março de 2018 pela comissão julgadora constituída pelos doutores:

Prof. Dra. Arianne Machado Lima

Universidade de São Paulo

Presidente

Prof. Dr. Carlos Bandeira de Mello Monteiro

Universidade de São Paulo

Prof. Dr. Roberto Hirata Junior

Universidade de São Paulo

Prof. Dr. Leandro Luís Galdino de Oliveira

Universidade Federal de Goiás

*A meu pai e minha mãe que não puderam estudar, mas se esforçaram para que eu o
pudesse fazer por eles*

Agradecimentos

Quero agradecer à minha família que mesmo sendo humilde me apoiou e incentivou ao estudo. Aos meus pais, Maria e José, que sacrificaram muitas coisas para que eu pudesse estudar e aos meus irmãos, Isis e Anderson, que ofereceram sempre apoio e orações por mim.

À minha orientadora Ariane Machado Lima, pela oportunidade, orientação, conselhos e exemplo acadêmico. À professora Fátima Nunes pela colaboração em diversos momentos durante a elaboração dessa pesquisa e à professora Helena Brentani pelo apoio e parceria no desenvolvimento deste trabalho.

À equipe da escola de ensino infantil e fundamental WISH pela colaboração durante a coleta de dados.

Aos meus colegas de laboratório Jéssica dos Santos de Oliveira, Renan Vinicius Aranha, Rafael Luiz Testa e Thiago Nobayashi pela colaboração e sugestões e por sempre estarem disponíveis para me ajudar.

A todos os professores e educadores que acreditaram no meu potencial e dedicaram seu tempo tirando minhas dúvidas.

E principalmente, a Deus, sem o qual eu não estaria aqui.

“E sabemos que todas as coisas contribuem juntamente para o bem daqueles que amam a Deus, daqueles que são chamados segundo o seu propósito.”
(Romanos 8:28)

Resumo

PINHEIRO, Tuany Dias. **Classificação de imagens faciais para o auxílio ao diagnóstico do transtorno do espectro autista**. 2018. 89 f. Dissertação (Mestrado em Ciências) – Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, São Paulo, 2018.

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. 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 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 e aplicação de questionários e escalas padronizadas. 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 típico mostrou que podem existir diferenças como distâncias entre as pupilas, formato das orelhas, estrabismo e circunferência da cabeça. A hipótese é que seria possível classificar indivíduos com TEA e indivíduos em desenvolvimento típico com base nas medidas antropométricas faciais. Desta forma, este trabalho teve como objetivo a construção de um classificador que, dada uma imagem facial de uma criança, consiga discriminar entre os dois grupos, auxiliando assim o diagnóstico. A fim de testar a hipótese, foram coletadas imagens bidimensionais de crianças e adolescentes com TEA e em desenvolvimento típico para a construção de uma base de dados. As imagens foram processadas por meio de um *pipeline* definido neste trabalho e foram testados e comparados diferentes métodos de redução de dimensionalidade e classificação e como resultado obteve-se acurácia de 80% na classificação com *Random Forests*.

Palavras-chaves: Classificação. Transtorno do espectro autista. Antropometria. Processamento de Imagens.

Abstract

PINHEIRO, Tuany Dias. **Classification of facial images to aid the diagnosis of Autism Spectrum Disorder**. 2018. 89 p. Dissertation (Master of Science) – School of Arts, Sciences and Humanities, University of São Paulo, São Paulo, 2018.

Autism Spectrum Disorder (ASD) is a developmental disorder that persistently impairs communication and social interaction and causes restricted and repetitive patterns of behavior, interests, and activities. These symptoms are present from the beginning of childhood and limit or impair the daily life of the individual. However, several factors prevent it from being possible to diagnose before the age of three, including the fact that the diagnosis is essentially clinical and performed based on the criteria described in the Diagnostic and Statistical Manual of Mental Disorders of the American Psychiatry Association (DSM) , interviews with parents, observation of behavior and application of questionnaires and standardized scales. These tools and questionnaires to carry out the diagnosis still lack validation and adaptation to the Brazilian context. The study of anthropometric features in individuals with ASD and individuals in typical development showed that there may be differences such as distances between the pupils, ear format, strabismus and head circumference. The hypothesis is that it would be possible to classify individuals with ASD and individuals in typical development based on anthropometric facial measures. Therefore, this work aimed to construct a classifier that, given a child's facial image, can discriminate between the two groups, thus helping the diagnosis. In order to test the hypothesis, two-dimensional images of children and adolescents with ASD and in typical development were collected for the database construction. The images were processed in a pipeline defined in this work and different methods of dimensionality reduction and classification were tested and compared and as a result 80% accuracy was obtained in the classification with Random Forests algorithm.

Keywords: Classification. Autism Spectrum Disorder. Anthropometry. Image processing.

Lista de figuras

Figura 1 – Exemplo dos <i>Landmarks</i> ou pontos de controle antropométricos	23
Figura 2 – Pontos de controle	24
Figura 3 – Exemplo de características <i>Haar</i>	26
Figura 4 – Característica que corresponde à região dos olhos	27
Figura 5 – Característica que corresponde à região do nariz	27
Figura 6 – Visão geral do algoritmo <i>Constrained Local Neural Fields</i>	29
Figura 7 – Áreas de interesse em uma imagem	30
Figura 8 – Exemplo de matriz de confusão de um classificador.	39
Figura 9 – Distâncias em garotos com TEA	43
Figura 10 – Etapas do processamento da imagem	50
Figura 11 – Método de rotação baseado nas pupilas	51
Figura 12 – Ajuste da imagem após rotação	51
Figura 13 – Exemplos de distorção: à esquerda, original; no centro, distorção radial; à direita, distorção tangencial	52
Figura 14 – Comparação entre imagens considerando etapa de calibração	52
Figura 15 – Imagem após detecção de bordas com Canny	53
Figura 16 – Resultado da detecção da faixa	54
Figura 17 – Exemplo de extração Viola-Jones	55
Figura 18 – Exemplo de extração Viola-Jones	55
Figura 19 – Regiões detectadas pelo Viola-Jones	56
Figura 20 – Exemplo da extração CLNF	57
Figura 21 – Redução de dimensionalidade e classificação	60
Figura 22 – Acurácia média do classificador <i>Naive Bayes</i> para cada método de redução de dimensionalidade	64
Figura 23 – Acurácia média do classificador <i>Naive Bayes</i> para cada método de redução de dimensionalidade	65
Figura 24 – Acurácia média do classificador SVM linear para cada método de redução de dimensionalidade	66
Figura 25 – Acurácia média do classificador SVM polinomial para cada método de redução de dimensionalidade	67

Figura 26 – Acurácia média do classificador SVM radial para cada método de redução de dimensionalidade	67
Figura 27 – Acurácia média do classificador SVM sigmóide para cada método de redução de dimensionalidade. Os métodos mRMR, RFS, FCBF e CFS obtiveram exatamente os mesmos valores em todas medidas	68
Figura 28 – Acurácia, precisão, revocação, f1-score e erro médios agrupados por método de redução de dimensionalidade	70
Figura 29 – Comparação entre classificadores por conjunto de características	72
Figura 30 – Acurácia média do classificador <i>Naive Bayes</i> para cada método de redução de dimensionalidade.	87
Figura 31 – Acurácia média do classificador SVM linear para cada método de redução de dimensionalidade.	87
Figura 32 – Acurácia média do classificador SVM polinomial para cada método de redução de dimensionalidade.	88
Figura 33 – Acurácia média do classificador SVM radial para cada método de redução de dimensionalidade.	88
Figura 34 – Acurácia média do classificador SVM sigmóide para cada método de redução de dimensionalidade. Alguns métodos obtiveram os mesmos resultados e suas curvas se sobrepõem (CFS, FCBF, mRMR, RFS) . .	89

Lista de tabelas

Tabela 1	–	Resumo dos trabalhos na literatura	46
Tabela 2	–	Otimização dos parâmetros SVM	61
Tabela 3	–	Otimização dos parâmetros <i>Random Forests</i>	62
Tabela 4	–	Resumo considerando melhores acurácias médias dos quatro conjuntos de distâncias por quantidade de características. A coluna método é descrita pela sigla do algoritmo de seleção de características (CFS, FCBF, mRMR, PCA, ReliefF, RFS) e a sigla do tipo de classificador (<i>Naive Bayes</i> (NB), SVM linear, SVM polinomial (poli), SVM radial (rbf) e SVM sigmóide).	69
Tabela 5	–	Melhores resultados na classificação por conjunto baseados na acurácia	71

Lista de abreviaturas e siglas

ADI-R	<i>Autism Diagnostic Interview-Revised</i>
ADOS	<i>Autism Diagnostic Observation Schedule</i>
ADOS-G	<i>Autism Diagnostic Observation Schedule-Generic</i>
B&B	<i>Branch and Bound</i>
BFF	<i>Best first search</i>
BioInfoMed	Laboratório de Bioinformática e Informática Médica
CARS	<i>Childhood Autism Rating Scale</i>
CFS	<i>Correlation-based Feature Selection</i>
CLNF	Algoritmo para localização de pontos em imagens <i>Constrained Local Neural Fields</i>
CLM	Algoritmo para localização de pontos em imagens <i>Constrained Local Model</i>
DCM	Descrição de comprimento mínimo
DSM	Manual diagnóstico e estatístico de transtornos mentais da sociedade americana de psiquiatria
EACH-USP	Escola de Artes, Ciências e Humanidades da Universidade de São Paulo
FCBF	<i>Fast Correlation-based Filter</i>
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 <i>Jet Voting</i>
kNN	Algoritmo de classificação <i>k</i> th Nearest Neighbours

LApIS	Laboratório de Aplicações de Informática em Saúde
LBP	Descritor visual <i>Local Binary Patterns</i>
LDA	Algoritmo de classificação <i>Linear Discriminant Analysis</i>
LIM/23	Laboratório de Psicopatologia e Terapêutica Psiquiátrica
mRMR	<i>Minimum Redundancy Maximum Relevance</i>
NB	<i>Naive Bayes</i>
PCA	<i>Principal Component Analysis</i>
PDD-NOS	Transtorno global do desenvolvimento sem outra especificação
PROTEA	Programa de Transtorno do Espectro Autista
RF	Algoritmo de classificação <i>Randoms Forests</i>
RFS	<i>Robust Feature Selection</i>
SVM	Algoritmo de classificação <i>Support Vector Machines</i>
TEA	Transtorno do Espectro Autista
UNIFESP	Universidade Federal de São Paulo

Sumário

1	Introdução	16
1.1	<i>Objetivos</i>	18
1.2	<i>Justificativas</i>	19
1.3	<i>Organização do documento</i>	20
2	Fundamentação teórica	22
2.1	<i>Antropometria</i>	22
2.2	<i>Processamento de imagens</i>	22
2.2.1	Definição de imagem	22
2.2.2	Detecção de Faces	24
2.3	<i>Aprendizado de máquina supervisionado</i>	30
2.3.1	Redução da dimensionalidade	31
2.3.2	Classificação	36
2.3.3	Avaliação de classificadores	38
3	Revisão bibliográfica	40
4	Metodologia	47
4.1	<i>Construção de uma base de imagens faciais de crianças afetadas por TEA e crianças em desenvolvimento típico</i>	47
4.2	<i>Processamento das imagens coletadas</i>	49
4.2.1	Pré-processamento	50
4.2.2	Detecção da faixa de referência	51
4.2.3	Detecção dos pontos de controle e extração de características	54
4.3	<i>Redução de dimensionalidade</i>	58
4.4	<i>Treinamento e comparação de classificadores</i>	60
5	Resultados e discussão	63
5.1	<i>Comparação dos métodos de redução de dimensionalidade</i>	63
5.2	<i>Comparação dos métodos de classificação</i>	71
6	Considerações Finais	73

Trabalhos futuros	75
-----------------------------	----

Referências ¹	76
------------------------------------	----

APÊNDICES	82
------------------	-----------

Apêndice A – Termo de consentimento livre e esclarecido . .	83
---	----

Apêndice B – Termo de assentimento	85
--	----

Apêndice C – Gráficos de acurácia média por subconjunto de características dos métodos de redução de di- mensionalidade	87
---	----

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

1 Introdução

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 do transtorno. Os resultados apresentados

indicam correlação entre os valores obtidos por meio 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 manuais com fitas métricas ou por meio do processamento de uma imagem facial (FARKAS et al., 2002). A extração a partir de uma imagem ocorre por meio de detecção de pontos de controle que podem ser utilizados 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 por meio de 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 as 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 alto¹ (ALDRIDGE et al., 2005; ALDRIDGE et al., 2011; HAMMOND; SUTTIE, 2012). Desta forma, é também de interesse investigar se técnicas de menor custo são capazes de auxiliar no diagnóstico de TEA por meio da análise de medidas antropométricas.

1.1 Objetivos

Considerando as perspectivas e lacunas observadas levantadas pela literatura apresentadas no capítulo 3, este projeto tem como objetivo geral propor e testar um classificador de imagens faciais de crianças e adolescentes, capturadas por uma câmera fotográfica digital comum, que discrimine-as entre as classes TEA e desenvolvimento típico, assim como criar uma base de imagens faciais utilizada para treinamento desse classificador e estudos futuros.

A entrada do classificador é um conjunto que contém distâncias entre pontos de controle obtidos das imagens. 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;
2. processar as imagens coletadas para extração das medidas antropométricas faciais, ajustando os algoritmos de processamento, se necessário;

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

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 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.

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 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 por meio de da intervenção precoce de forma que os voluntários envolvidos 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.

Os instrumentos disponíveis atualmente para realização do diagnóstico são em sua maioria questionários e rotinas de observação do comportamento da criança, de forma que muitas vezes este diagnóstico é adiado até a idade escolar (MANDELL; NOVAK; ZUBRITSKY, 2005). Dessa forma, utilizar técnicas computacionais para processar imagens de pacientes e classificá-los de acordo com medidas antropométricas pode ser uma alternativa para auxiliar o diagnóstico precoce do TEA.

Portanto, a proposta aqui apresentada inclui capturar imagens bidimensionais utilizando uma câmera digital comum² para a criação de um banco de dados de imagens de indivíduos com TEA a fim de realizar análise 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, principalmente, prática clínica.

A análise das características antropométricas faciais por meio de imagens oferece uma forma não invasiva de auxílio ao diagnóstico de TEA que poderia auxiliar o diagnóstico precoce e assim permitiria uma intervenção antecipada que 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 (controles) 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 coordenado pelo laboratório de Psicopatologia e Terapêutica Psiquiátrica (LIM/23) do Instituto de Psiquiatria do Hospital das Clínicas (IPq) com a colaboração do laboratório de Bioinformática e Informática Médica (BioInfoMed) e do laboratório de Aplicações de Informática em Saúde (LApIS) da Escola de Artes, Ciências e Humanidades da Universidade de São Paulo (EACH-USP). O objetivo deste projeto maior é realizar o estudo do diversos aspectos 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. Sendo assim, não é pretensão deste trabalho realizar o diagnóstico puramente baseado nas medidas antropométricas, mas compor um classificador multimodal que será baseado em vários aspectos do TEA.

1.3 Organização do documento

Este documento está organizado em capítulos descritos a seguir.

² Para a coleta dos dados foi utilizada uma câmera semi-profissional descrita na seção de metodologia

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

No capítulo 3 é apresentada a revisão bibliográfica de métodos de análise antropométrica aplicados ao diagnóstico de transtornos e síndromes.

No capítulo 4 é apresentada a metodologia de desenvolvimento deste trabalho na resolução do problema de pesquisa e coleta dos dados.

No capítulo 5 são apresentados os resultados obtidos e a discussão.

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

2 Fundamentação teórica

Este capítulo apresenta os conceitos fundamentais envolvidos nesta pesquisa. Na seção 2.1 é apresentada a definição de antropometria; na seção 2.2 são descritos conceitos envolvidos no processamento das imagens; na seção 2.3 são apresentados conceitos de aprendizado de máquina supervisionado, incluindo os métodos de redução de dimensionalidade e de classificação utilizados neste trabalho.

2.1 Antropometria

Antropometria é a ciência que estuda as medidas do corpo humano incluindo altura, peso e tamanho dos componentes do corpo (FARKAS, 1994). 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, pois variações podem ocorrer nos dois tipos de abordagem (direta ou indireta) (FARKAS et al., 2002). Para isso, é necessária a utilização de um sistema de medida metuculoso baseado em pontos de referência (DEUTSCH et al., 2012). FARKAS propôs um manual antropométrico descrevendo tais pontos e definindo cada medida antropométrica como sendo a distância entre dois desses pontos, além de definir normas para a captura dos dados antropométricos. Esses pontos de referência são chamados de *landmarks* ou pontos de controle. Neste trabalho é usada a nomenclatura ponto de controle. Na figura 1 são apresentados exemplos de pontos de controle sobre a face humana.

2.2 Processamento de imagens

2.2.1 Definição de imagem

As definições a seguir são baseadas em (GONZALEZ; WOODS, 2002). Os dados capturados por sensores de diversas naturezas podem ser armazenados para formar imagens. Esses dados podem ser processados no domínio da frequência ou no domínio espacial. No presente documento, a imagem é definida e processada no domínio espacial.

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

Fonte: (FARKAS, 1994)

Imagem digital é o termo utilizado para se referir à função bidimensional de intensidade de luz $f(x, y)$ em que x e y são as coordenadas espaciais do plano, sendo que o valor de f em x, y é proporcional ao brilho da imagem.

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 de pontos importantes, tais como bordas.

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 tem-se:

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

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 é chamado elemento da imagem ou pixel (do inglês, *picture element*).

2.2.2 Detecção de Faces

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

Fonte: Adaptado de (BALTRUSAITIS; ROBINSON; MORENCY, 2013)

Para extrair medidas antropométricas de imagens que representam faces humanas, é necessário identificar os pontos de controle definidos na seção 2.1. Assim, no contexto de

uma imagem bidimensional, um ponto de controle é composto pelas coordenadas (x,y) do pixel que representa tal ponto. A figura 2 exemplifica 68 pontos de controle da face.

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. De acordo com Hsu, Abdel-Mottaleb e Jain (2002), existem algumas 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 utilizam uma combinação das duas primeiras. As técnicas baseadas em modelos assumem uma posição inicial para a face na imagem e corrigem a posição dos pontos por meio 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.

Cada uma das abordagens citadas executa diversos algoritmos, cada um deles com uma finalidade específica, a fim de oferecer um resultado final com os componentes faciais devidamente identificados. A seguir são apresentados os principais algoritmos relacionados ao desenvolvimento do presente trabalho.

Algoritmo Viola-Jones

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

O algoritmo Viola-Jones utiliza classificadores e características para detectar os componentes faciais. Tem como objetivo detectar faces e componentes faciais que estejam em posição frontal. Isso é feito por meio 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 algoritmo AdaBoost;
4. aplicação dos classificadores em cascata.

As características utilizadas pelo Viola-Jones são características *Haar* baseadas na proposta de 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: (VIOLA; 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. 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.

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

Fonte: Adaptado de (VIOLA; JONES, 2001)

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

Fonte: Adaptado de (VIOLA; JONES, 2001)

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

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.

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

Usando as recorrências descritas nas equações 3 e 4, onde $s(x, y)$ é a soma cumulativa por linha, $s(x, -1) = 0$ e $g(-1, y) = 0$, a imagem integral é calculada em apenas uma varredura sobre a imagem original.

$$s(x, y) = s(x, y - 1) + f(x, y) \quad (3)$$

$$g(x, y) = g(x - 1, y) + s(x, y) \quad (4)$$

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 por meio de uma 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. Cada estágio executado em cascata é composto 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.

Algoritmo *Constrained Local Neural Fields*

As definições a seguir 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. Por meio da utilização de um modelo de distribuição dos pontos, realiza a detecção adaptando a possível face na imagem aos pontos de controle do modelo. A figura 6 apresenta a visão geral do algoritmo CLNF.

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

Fonte: 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 de treinamento em que os pontos de controle foram anotados manualmente. Esta fase de treinamento é executada previamente e uma vez obtido o modelo, não é necessário relizar o treinamento outra vez. 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 sendo que uma área de interesse é uma região da imagem em que existe um ponto de controle em potencial.

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

Figura 7 – Áreas de interesse em uma imagem

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

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 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 por meio de uma função objetivo. Depois de encontrados os parâmetros Δp o modelo é adaptado à 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.

2.3 Aprendizado de máquina supervisionado

As definições nesta seção são baseadas em Mitchell (1997), Duda, Hart et al. (1973), Dash e Liu (1997), Vapnik e Chapelle (2000), Chapelle et al. (2002) e Breiman (2001).

Aprendizado de máquina realiza o estudo e proposta de métodos e algoritmos que desempenham uma tarefa de forma que o aprendizado da tarefa é realizado de forma automatizada. 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 um conjunto de n atributos ou características que a definem. Esse conjunto de características é representado por um vetor de dimensão n .

2.3.1 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.

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.

Para realizar a redução da dimensionalidade existem duas estratégias comumente utilizadas: fusão e seleção de características.

Métodos baseados na estratégia de fusão criam um novo conjunto de características a partir do conjunto inicial (DASH; LIU; YAO, 1997) de acordo com algum critério de transformação. Métodos baseados na estratégia de seleção de características tipicamente realizam uma busca no espaço de forma a encontrar o subconjunto ótimo seguindo 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 em busca aleatória para reduzir a complexidade computacional.

Sendo assim, os vários métodos de seleção de características podem se distinguir de acordo com:

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 que percorrem o espaço de busca utilizam abordagens diferentes que incluem realizar busca completa, utilizar heurísticas ou realizar busca 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.

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 baseados em algoritmo genético (VAFAIE; IMAM, 1994) e *simulated annealing* (DOAK, 1992).

Quanto à 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. Já os métodos embutidos são aqueles nos quais a seleção de características ocorre durante o treinamento do classificador, ou seja, são embutidos no processo de aprendizado do mesmo.

Os métodos do tipo filtro se baseiam nas características gerais do conjunto de treinamento e obtém o subconjunto de características que maximizam a função de avaliação utilizada. Porém, por não utilizar o classificador como função de avaliação, por vezes podem apresentar um desempenho inferior aos métodos do tipo camada, já que a combinação entre um determinado classificador e o conjunto de características específico para ele costuma apresentar melhor desempenho. Os métodos *Minimum Redundance Maximum Relevance* (mRMR), *Correlation-based Feature Selection* (CFS), *Fast Correlation-based Filter* (FCBF), ReliefF e *Robust Feature Selection* (RFS) são exemplos de métodos de seleção de características do tipo filtro. Relief (KIRA; RENDELL, 1992) realiza seleção de características e é mais rápido em comparação com métodos camada, porém não considera a correlação e algumas características redundantes podem ser selecionadas.

Os métodos do tipo camada utilizam o próprio classificador como função de avaliação para identificar o subconjunto de características ótimo. Esse tipo de método costuma apresentar bons resultados, porém o custo computacional é alto dado que, para cada subconjunto possível, um classificador deve ser treinado e avaliado.

Os métodos embutidos possuem a vantagem de selecionarem o subconjunto de características mais adequado ao classificador em questão sem o custo computacional apresentado pelos métodos do tipo camada, já que a seleção é realizada durante o único treinamento do classificador.

A seguir são descritos sucintamente alguns métodos de redução de dimensionalidade que utilizam fusão e seleção de características.

Principal component analysis

O método *Principal Component Analysis* (PCA) é um exemplo de fusão por meio de uma transformação ortogonal de forma a obter os componentes principais linearmente não correlacionados. O método PCA condensa a variação de todas as características do conjunto inicial em apenas algumas características de forma que o primeiro componente é o que apresenta a maior variância, o segundo componente apresenta a segunda maior variância, e assim sucessivamente.

Minimum redundancy maximum relevance

O método *Minimum Redundance Maximum Relevance* (mRMR) (DING; PENG, 2005) tem como proposta selecionar um subconjunto em que as características sejam diferentes o máximo possível umas das outras. São estabelecidos dois critérios, o critério de redundância mínima e o critério de relevância máxima.

Por meio da informação mútua, o método possui uma medida de avaliação da similaridade entre as características. O critério de redundância é utilizado para selecionar características que sejam o máximo possível distantes em relação à medida de similaridade, enquanto que o critério de relevância procura maximizar a informações mútua entre cada característica e a classe alvo.

Diversas medidas podem ser utilizadas para determinar a correlação ou dependência das características com a classe alvo (relevância) e a distância entre as características relevantes (redundância), como por exemplo a medida de informação mútua. Neste caso, a ideia é maximizar a dependência entre a distribuição conjunta das características selecionadas e a classe alvo.

Correlation-based feature selection

Correlation Feature Selection (CFS) (HALL, 1999) utiliza uma medida de correlação para avaliar o subconjunto de características seguindo a premissa de que um bom subconjunto contém características altamente correlacionadas com a classe alvo e minimamente correlacionadas entre si.

A medida das correlações entre característica-classe alvo e característica-característica podem ser obtidas de diversas formas. Em Hall (1999) são utilizadas medidas com base em descrição de comprimento mínimo (DCM), incerteza simétrica e o resultado do método Relief. Utilizando as medidas de correlação característica-classe alvo e característica-característica, obtém-se uma função de pertencimento da característica ao subconjunto e, por meio da otimização desta função, obtém-se o subconjunto de características.

Fast correlation-based filter

Fast Correlation-Based Filter (FCBF) (YU; LIU, 2003) baseia-se no mesmo princípio do método CFS, porém as características relevantes e as redundâncias entre elas são calculadas sem realizar análise da correlação característica-característica. A medida de correlação considerada é a incerteza simétrica. Yu e Liu (2003) definem os conceitos de correlação predominante e característica predominante.

A correlação predominante SU é o maior valor de correlação de uma característica com a classe alvo considerando um valor *threshold*, e as características que possuem valores iguais à correlação predominante são ditas pares redundantes. A característica predominante é definida como a característica que possui correlação predominante SU sem que haja outras características pares redundantes. Com base nestes conceitos, o método FCBF seleciona as características predominantes enquanto elimina as características redundantes.

ReliefF

O método ReliefF (KONONENKO; ŠIMEC; ROBNIK-ŠIKONJA, 1997) é baseado no método Relief (KIRA; RENDELL, 1992). A ideia principal em ambos é estimar as características de acordo com o quanto os valores se distinguem das instâncias que estão próximas ou vizinhas. No método Relief, é realizada a busca por dois vizinhos: um que pertença à mesma classe, chamado *nearest hit*, e outro de uma classe diferente, chamado *nearest miss*.

No método Relief, as características redundantes e ruidosas podem afetar fortemente a seleção dos vizinhos mais próximos e, portanto, a estimativa de probabilidades torna-se não confiável. Para aumentar a precisão da aproximação de probabilidade, o método

ReliefF busca os *near-hits/near-misses* mais próximos em vez de apenas os dois vizinhos que o método Relief encontra e mede a contribuição de todos os *near-hits/near-misses* mais próximos.

Robust feature selection

Robust Feature Selection (RFS) (NIE et al., 2010) utiliza uma estratégia em que os dados são representados por meio de um dicionário esparso construído utilizando combinação linear. O aprendizado por dicionário esparso refere-se a métodos que tem por objetivo encontrar e interpretar o *trade-off* entre alguma medida de qualidade e a dispersão do resultado.

O método RFS realiza a minimização conjunta de distâncias euclidianas e Manhattan nas funções de regularização e custo para selecionar as características mais significativas. A minimização destas distâncias na função de custo permite com que o método seja eficiente em relação a conjuntos de dados com *outliers*. A minimização das distâncias euclidianas e Manhattan na função de regularização tem como objetivo selecionar as características entre todas com base na medida de *joint sparsity*.

2.3.2 Classificação

Uma tarefa de classificação em sua essência tem como objetivo aprender o modelo que permite o reconhecimento de um padrão, de forma a possibilitar a categorização de um 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 por meio 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. A partir destes, 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.

Support Vector Machine

O método *Support Vector Machine* (SVM) (VAPNIK, 1963) realiza a classificação separando as classes por meio de um hiperplano em um espaço de alta dimensionalidade. A menor distância entre os exemplos do conjunto de treinamento e o hiperplano de separação entre as classes é a margem sendo que quanto maior a margem, melhor é a separação entre as classes.

Em um problema de classificação, nem sempre é possível separar as classes linearmente. Nestes casos, é possível utilizar uma SVM não linear que faz uso de uma função *kernel* capaz de manipular os vetores de características em um novo espaço, de alta dimensão, no qual as classes sejam linearmente separáveis. Exemplo desta função *kernel* são chamadas de *kernel* radial, polinomial, sigmóide, dentre outros.

Random Forests

Random Forests (HO, 1995) é um método de aprendizado em conjunto que utiliza uma combinação de árvores preditoras para realizar classificação. Cada árvore é criada a partir de um subconjunto de treinamento selecionados do conjunto de dados original. A partir do subconjunto de treinamento, para cada árvore, são selecionadas aleatoriamente k características sendo que m é o total de características e $k < m$. Os nós são construídos a partir da melhor divisão do subconjunto de características k e a árvore cresce o máximo possível.

As importâncias das características no método *Random Forests* são calculadas com base em alguma medida de importância como critério baseado em erro, critério baseado em índice Gini, entre outros. Após a execução do método é possível obter um ranking das características mais importantes para a classificação.

Gaussian Naive Bayes

Naive Bayes (NB) é um método de classificação probabilístico baseado no Teorema de Bayes que assume as características são fortemente independentes (DEVROYE; GYÖRFI; LUGOSI, 2013). O método NB prediz as probabilidades de pertencimento para cada classe tal como as probabilidades de um exemplo pertencer a uma classe específica.

Pode-se assumir uma certa distribuição para os valores das características. Um dos tipos de classificador NB é o *Gaussian Naive Bayes* em que é assumido que os valores das características para cada classe seguem uma distribuição gaussiana ou normal.

2.3.3 Avaliação de classificadores

As definições desta seção baseiam-se em Metz (1978), Powers (2011) e Fawcett (2006).

Ao realizar uma tarefa de classificação é necessário avaliar a performance do classificador. Esta avaliação pode ser realizada com base em várias medidas entre elas acurácia, precisão, revocação, f1-score, erro, entre outras.

Em um problema de classificação binária, os exemplos de treinamento possuem os rótulos de sua respectiva classe (positiva ou negativa) e estes exemplos rotulados são utilizados para treinar o classificador. Após o treinamento são apresentados novos exemplos de teste de forma que o classificador deve predizer entre as classes correspondentes. A partir do conjunto de treinamento e considerando as classes preditas e reais, é construída a matriz de confusão conforme exemplo da figura 8.

Um exemplo pode ser considerado como verdadeiro positivo, falso positivo, falso negativo e verdadeiro negativo e tendo em vista cada um desses cenários as medidas de avaliação são calculadas conforme descrito a seguir:

Figura 8 – Exemplo de matriz de confusão de um classificador.

Fonte: Tuany Dias Pinheiro

$Acurácia = \frac{VP+VN}{P+N}$, em que P corresponde ao total de exemplos positivos ou $VP + FN$ e N corresponde ao total de exemplos negativos ou $VN + FP$. A acurácia apresenta uma medida de desempenho geral do classificador e representa a capacidade de prever corretamente a classe de um exemplo de entrada.

$Erro = 1 - Acurácia$. A medida de erro é redundante se utilizada em conjunto com a medida de acurácia, portanto é comum ser utilizada somente acurácia ou somente erro para avaliar o desempenho geral do classificador.

$Precisão = \frac{VP}{VP+FP}$. A medida de precisão indica a proporção de exemplos classificados como positivos que são realmente positivos.

$Revocação = \frac{VP}{P}$, em que P corresponde ao total de exemplos positivos ou $VP + FN$. A medida de revocação também é conhecida como sensibilidade e indica a proporção de reais positivos que foram corretamente classificados como positivos.

$F_1 \text{ score} = 2 \times \frac{Precisão \times Revocação}{Precisão + Revocação}$. A medida $F_1 \text{ score}$ considera precisão e revocação na forma de média harmônica e o melhor valor possível é 1 em que a precisão e a revocação são máximas.

3 Revisão bibliográfica

Neste capítulo são apresentados os trabalhos encontrados durante a revisão da literatura de análise antropométrica facial em 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, Cornélia de Lange, X frágil, Mucopolissacaridose 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), *Support 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 por meio 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 dos pontos automaticamente (detecção realizada por meio 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, Mucopolissaca-

ridose 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* 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 por meio de 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 Manhattan) 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 estereofotogrametria¹ 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 significativamente 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: (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 por meio de 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 por meio de 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 distinguiram parcialmente do grupo de controle, do agrupamento 2 se distinguiram totalmente do grupo de controle e do agrupamento 3 não se distinguiram 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, Mucopolissacaridose tipo 3, Noonan, Prader-Willi, Smith-Lemli-Opitz, Sotos, Williams-Beuren	LDA, SVM, k NN, 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, k NN, 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 Manhattan	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 Metodologia

Conforme já mencionado, o objetivo deste trabalho é propor a construção e o treinamento de um classificador que, dada uma imagem facial de uma criança ou adolescente cujo diagnóstico não seja conhecido previamente, esta imagem seja classificada entre TEA ou desenvolvimento típico, de forma a ser uma ferramenta de apoio ao diagnóstico de TEA.

As imagens para treinamento e teste do classificador foram capturadas seguindo um protocolo definido neste trabalho e uma base de dados de estudo foi construída com essas imagens, conforme descrito na seção 4.1. Por se tratarem de dados sensíveis, foram utilizadas apenas algumas das imagens de voluntários de pesquisa sendo que, para ilustrar a metodologia de trabalho foram editadas de forma a cobrir o rosto e, além disso, foram autorizadas pelos respectivos responsáveis. Para os demais exemplos, foram utilizadas imagens da base de dados pública *CBCL Face database* (SUNG, 1996).

As características exploradas neste estudo foram obtidas por meio do processamento das imagens de cada voluntário. Para a construção do classificador foram testados seis métodos de seleção de características e três métodos de classificação com variações nos parâmetros, conforme descrito nas seções 2.3.1 e 2.3.2, respectivamente. A calibração dos parâmetros e estimação de desempenho dos classificadores foi realizada via validação cruzada *k-fold*, com $k = 4$.

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.

Foi idealizada a construção de uma base de dados com imagens de crianças com TEA e desenvolvimento típico (controles) para realização de pesquisas com imagens por membros do laboratório de Aplicações de Informática em Saúde (LApIS) e do laboratório de Bioinformática e Informática Médica (BioInfoMed). O grupo de casos foi formado por voluntários com TEA do ambulatório do Instituto de Psiquiatria do Hospital das

Clínicas da Universidade São Paulo sendo que a equipe do laboratório de Psicopatologia e Terapêutica Psiquiátrica (LIM/23) auxiliou durante a coleta nas avaliações clínicas dos voluntários. O grupo de controles foi formado por voluntários que responderam aos *e-mails* de requisição de voluntários para pesquisa e alunos voluntários da escola de ensino infantil e fundamental Wish School, em São Paulo.

Para realizar a coleta foi definido um protocolo que ajudasse a padronizar as imagens. Todas as fotografias foram capturadas com a câmera compacta Coolpix L120 Nikon. Em cada coleta, a câmera foi acoplada a um tripé que foi fixado a uma distância de 50cm da cadeira onde a criança deveria estar acomodada. Foi escolhida como fundo uma parede na cor branca, onde foi fixada a faixa vermelha de 50cm por 5cm que serve como referência para normalizar as imagens e não foi utilizada outro tipo de iluminação além da iluminação existente na sala. A cadeira utilizada era uma cadeira estofada comum sem descansos de braço e foi fixada rente a parede.

Inicialmente, os dados antropométricos coletados pela equipe de genética médica do Hospital São Paulo da Universidade Federal de São Paulo (UNIFESP) seriam utilizados para comparação com os dados obtidos das imagens coletadas. Porém estes os dados já existentes se referem, em sua maioria, à medidas antropométricas de outras partes do corpo, sendo que as medidas faciais coletadas são poucas e estão disponíveis apenas em forma dos gráficos inseridos no manual antropométrico (BRUNONI; PEREZ, 2013). Dessa forma, foram utilizadas apenas medidas em pixels para a classificação e não foi realizada a comparação com os dados existentes.

Para a coleta, as crianças foram instruídas a permanecer com a expressão o mais neutra possível. Foram capturadas cinco fotografias faciais frontais e, em alguns casos, a gravação de um vídeo curto para que fosse possível extrair *frames* estáticos. A gravação do vídeo só foi realizada quando não foi possível obter uma fotografia em que a expressão estivesse neutra ou caso o rosto estivesse em uma posição muito inclinada.

Além das imagens, foram coletadas outras informações, como idade e pontuação na escala *Childhood Autism Rating Scale* (CARS) (SCHOPLER et al., 1980), que indica a gravidade do TEA para o grupo de casos. Os controles também passaram por avaliações clínicas e preencheram os questionários indicados pela equipe do laboratório de Psicopatologia e Terapêutica Psiquiátrica (LIM/23). Ao todo foram coletadas imagens de 69 casos e 60 controles.

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 e obtiveram parecer de aprovação.

Nos apêndices A e B estão disponíveis o termos de consentimento livre e esclarecido e termo de assentimento que foram submetidos ao Comitê de Ética e utilizados nesta pesquisa.

4.2 Processamento das imagens coletadas

A figura 10 apresenta um resumo da metodologia de processamento das imagens coletadas. Para cada criança participante da coleta há uma ou mais imagens frontais. A escolha das melhores imagens ocorreu de forma semiautomatizada baseando-se em dois critérios de escolha:

Critério 1 Expressão neutra

Critério 2 Proporção da faixa

Utilizando o critério 1, foram escolhidas subjetivamente todas as imagens em que o voluntário apresentasse face neutra. Utilizando o critério 2, foram escolhidas automaticamente uma imagem para cada participante em que a faixa apresentasse proporção de largura e altura em pixels o mais próxima da proporção real em centímetros. Dessa forma, foi necessário que o processamento até a etapa de detecção da faixa fosse executado em todas imagens escolhidas por meio do critério 1, de forma a estimar as menos distorcidas.

A descrição das etapas do processamento da imagem está dividida nas subseções 4.2.1, 4.2.2 e 4.2.3 que detalham o pré-processamento, a detecção da faixa de referência e a detecção dos pontos de controle e a extração de características dos dados para a classificação, respectivamente.

Figura 10 – Etapas do processamento da imagem

Fonte: Tuany Dias Pinheiro

4.2.1 Pré-processamento

A etapa de pré-processamento consiste em rotacionar a imagem para corrigir a inclinação da face. Para realizar a rotação da imagem utilizou-se o método descrito em (CARPIO, 2016). Este método detecta a região dos olhos utilizando *haar-cascades* e em seguida marca as pupilas; uma reta é traçada entre os pontos marcados e, em seguida, é calculado o ângulo entre o eixo horizontal e a reta traçada; o ângulo obtido é então utilizado para realizar a rotação. A figura 11 exemplifica o método de rotação em que θ é o ângulo formado pela reta traçada entre os pontos correspondentes as pupilas P_1 e P_2 e o eixo horizontal.

Após a rotação, as partes da imagem afetadas pela transposição de pixels são substituídas por pixels pretos que podem causar falhas durante a etapa de detecção da faixa, como mostra a figura 12a.

Figura 11 – Método de rotação baseado nas pupilas

Fonte: Adaptado de (CARPIO, 2016)

Para evitar que essas falhas ocorram, após a rotação é calculada a área do retângulo mínimo formado pela imagem sem as partes com fundo preto, exemplificado na figura 12b e a imagem é recortada. A figura 12c apresenta o resultado final do pré-processamento.

Figura 12 – Ajuste da imagem após rotação

Fonte: Tuany Dias Pinheiro

4.2.2 Detecção da faixa de referência

Em seguida ao pré-processamento ocorre a detecção da faixa de referência na qual são calculadas medidas de largura e altura em pixels para utilização posterior.

Por se tratarem de imagens obtidas a partir de uma câmera comum, distorções consideráveis podem existir, sendo necessário realizar correção de cada imagem calculando os parâmetros da câmera. A calibração da câmera é necessária para estimar os parâmetros internos da lente da câmera e como as coordenadas de um ponto da imagem são derivadas considerando a posição do ponto em relação à câmera (WENG et al., 1992).

Existem vários métodos para realizar a estimativa destes parâmetros e corrigir as distorções, sendo que grande parte se limita a remover as distorções radiais e descentradas utilizando uma matriz de correção (PRESCOTT; MCLEAN, 1997). O método testado para estimar os parâmetros da câmera e corrigir as imagens considera as distorções

radiais e tangenciais exemplificadas na figura 13. Porém a implementação disponibilizada na biblioteca OpenCV (BRADSKI, 2000) para esta finalidade apresentou resultados insatisfatórios e as imagens resultantes desse processamento apresentaram ainda mais distorções do que nas imagens originais.

Figura 13 – Exemplos de distorção: à esquerda, original; no centro, distorção radial; à direita, distorção tangencial

Fonte: Adaptado de (PRESCOTT; MCLEAN, 1997)

Na figura 14a pode ser verificado que a faixa apresenta curvatura causada pelo processo de remoção de distorções e as regiões próximas à borda da imagem parecem ser mais largas. A figura 14b apresenta a imagem sem a realização da etapa de remoção e apresenta menos distorções.

Figura 14 – Comparação entre imagens considerando etapa de calibração

(a) Curvatura nas bordas e região inferior maior

(b) Proporções mais próximas do real

Para verificar a distorção causada nas imagens foram realizados testes de detecção da faixa de referência com e sem a etapa de remoção da distorção. Incluindo a etapa de remoção de distorção apenas 9,3% das imagens apresentaram proporção da faixa de referência correspondente à real, enquanto que excluindo a etapa de remoção de distorção 80% das imagens apresentaram a proporção próxima à real. Sendo assim, esta etapa foi removida do *pipeline* de processamento das imagens.

Para realizar a detecção da faixa, o primeiro passo é a aplicação de um filtro para suavizar a imagem. Foram testados dois tipos de filtro: bilateral e gaussiano. O filtro bilateral substitui os valores das intensidades dos pixels com base na média ponderada dos

valores de intensidade de outros pixels de sua vizinhança, preservando melhor as bordas. Os pesos são calculados com base na distância entre os pixels e outros fatores como, por exemplo, diferença entre intensidades de cor e distância de profundidade. Enquanto o filtro bilateral supõe que pixels na mesma vizinhança possuam valores de intensidade próximos, o filtro gaussiano utiliza média ponderada de valores dos pixels tal que os pesos diminuam conforme a proximidade com o centro da imagem. Essa suposição faz com que as bordas sejam borradas ao utilizar esse filtro (TOMASI; MANDUCHI, 1998). A aplicação do filtro bilateral à detecção dos pontos correspondentes aos cantos da faixa de referência apresentou menor número de erros de detecção do que a aplicação do filtro gaussiano.

Figura 15 – Imagem após detecção de bordas com Canny

Fonte: Tuany Dias Pinheiro

Após a aplicação do filtro é realizada a detecção das bordas por meio do método Canny (CANNY, 1986). O método Canny se baseia em diversos passos, entre eles remoção de ruídos, cálculo dos gradientes das intensidades da imagem, afinamento das bordas, aplicação do duplo *threshold* para remover pixels decorrentes de ruídos e diminuição da fragmentação das bordas. A figura 15 apresenta o resultado com cores invertidas da aplicação do filtro e detecção de bordas utilizando o método Canny. Em seguida, são realizadas operações de dilatação e erosão para fechar os contornos que podem estar abertos após a aplicação do método Canny.

Finalmente, para encontrar o contorno referente à faixa, é utilizado o método *findContours* da biblioteca OpenCV que implementa o algoritmo descrito em Suzuki et al.

(1985) e retorna uma lista de contornos externos da imagem. Para o *pipeline* definido neste trabalho assume-se que o contorno mais próximo à extremidade superior corresponde à faixa. A figura 16 apresenta a imagem final deste processo.

Figura 16 – Resultado da detecção da faixa

Fonte: Tuany Dias Pinheiro

Com a faixa localizada, são calculadas a largura e a altura da mesma. Estas medidas são utilizadas para determinar as imagens com menor distorção e para construir os conjuntos de dados utilizados na tarefa de classificação. Para cada participante apenas a melhor imagem foi selecionada e, ao todo, 116 imagens foram escolhidas, compreendendo casos e controles, formando o conjunto final de imagens que passaram para a etapa de detecção dos pontos de controle.

4.2.3 Detecção dos pontos de controle e extração de características

De início foram realizados testes com o método Viola-Jones, descrito na seção 2.2.2. As figuras 17 e 18 mostram os resultados da detecção utilizando a implementação com método Viola-Jones e *Haar cascades* disponíveis na biblioteca OpenCV (BRADSKI, 2000). As imagens utilizadas neste teste foram retiradas da base *CBCL Face database* (SUNG, 1996).

Figura 17 – Exemplo da extração Viola-Jones: pontos da boca não foram detectados

Fonte: Tuany Dias Pinheiro

Nota-se que os pontos da região da boca não foram detectados corretamente. O método Viola-Jones utiliza arquivos para descrever as características *Haar* de cada região a ser identificada na imagem, porém essas regiões podem ser confundidas com outras partes que se assemelhem a parte que se deseja detectar.

Figura 18 – Exemplo da extração Viola-Jones: alguns pontos da boca foram detectados

Fonte: Tuany Dias Pinheiro

A figura 19 apresenta as regiões detectadas na imagem de acordo com as características *Haar*. Em amarelo, a região dos olhos; em verde a região do nariz e em vermelho a região da boca. Neste exemplo, a área entre as narinas foi detectada como sendo a boca por ter curvas que se assemelhassem a região desejada. Na tentativa de contornar este problema outros classificadores *Haar cascade* para boca foram testados, porém nenhum demonstrou ser eficiente na detecção destes pontos. De forma a resolver esta dificuldade

foi cogitado treinar um classificador *Haar* para a boca. Apesar de ser uma alternativa possível, 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 19 – 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.

Fonte: Tuany Dias Pinheiro

A figura 20 mostra o resultado da extração dos pontos de controle utilizando o método CLNF descrito na seção 2.2.2. Este método possui uma precisão maior na detecção dos pontos de todas as regiões e detecta mais pontos e dessa forma tem-se mais distâncias que podem ser analisados e comparados. Para a realização dos testes com o algoritmo CLNF, foi utilizada a ferramenta *Openface*¹ (BALTRUŠAITIS; ROBINSON; MORENCY, 2016).

¹ Código fonte disponível por licença *open source* em: <https://github.com/TadasBaltrusaitis/OpenFace>

Figura 20 – Exemplo da extração CLNF: todos os pontos foram detectados

Fonte: Tuany Dias Pinheiro

A distância entre cada par de pontos de controle detectados forma uma característica. O conjunto de todas as distâncias forma o vetor de características, que será posteriormente usado no processo de classificação. Inicialmente seriam calculadas somente distâncias euclidianas, no entanto outros tipos de distância poderiam causar um efeito diferente na classificação. As equações 5 e 6 apresentam o cálculo das distâncias euclidiana e Manhattan para dois pontos $p(x_1, y_1)$ e $q(x_2, y_2)$.

$$d_e(p, q) = \sqrt{((x_1 - x_2)^2 + (y_1 - y_2)^2)} \quad (5)$$

$$d_m(p, q) = |x_1 - x_2| + |y_1 - y_2| \quad (6)$$

Dessa maneira, foram avaliados quatro cálculos de distância, compondo portanto quatro conjuntos de características: f_e , $f_{e \times c}$, f_m e $f_{m \times c}$. Dado um par de pontos de controle $p(x_1, y_1)$ e $q(x_2, y_2)$, calculou-se as distâncias:

f_e : Distância Euclidiana

$$f_e = \frac{d_e(p, q)}{w},$$

em que d_e é a distância euclidiana entre p e q e w é a largura da faixa em pixels.

$f_{e \times c}$: Distância Euclidiana $\times 1000$

$$f_{e \times c} = \frac{d_e(p, q)}{w} \times c,$$

em que d_e é a distância euclidiana entre p e q , w é a largura da faixa em pixels e c é uma constante inteira = 1000.

f_m : Distância Manhattan

$$f_m = \frac{d_m(p, q)}{w},$$

em que d_m é a distância Manhattan entre p e q e w é a largura da faixa em pixels.

$f_{m \times c}$: Distância Manhattan $\times 1000$

$$f_{m \times c} = \frac{d_m(p, q)}{w} \times c,$$

em que d_m é a distância Manhattan entre p e q , w é a largura da faixa em pixels e c é uma constante inteira = 1000.

Dado que existem 68 pontos combinados dois a dois sem repetição, tem-se que: $C(68, 2) = 68! / (68 - 2)! \times 2! = 2278$. Sendo assim, foram obtidos quatro conjuntos de 116 instâncias descritas por 2278 características.

4.3 Redução de dimensionalidade

Os métodos de redução de dimensionalidade são de extrema importância em tarefas de classificação, como mencionado na seção 2.3.1. Alguns destes métodos foram revisados e selecionados para comparação. Cada um deles foi executado considerando valores de tamanho de subconjunto de características escolhidos de forma a definir subconjuntos pequenos e com quantidade menor que 10% do total de características.

Por se tratar de um conjunto com grande quantidade de características, um dos critérios de escolha leva em consideração o custo computacional da execução do método. Como já mencionado anteriormente, os métodos de seleção de características do tipo filtro, tendem a ter um custo computacional mais baixo em relação a métodos do tipo camada. Outro critério de escolha considera as implementações disponíveis para os métodos revisados. Neste trabalho foram utilizadas as implementações das bibliotecas *scikit-learn* (PEDREGOSA et al., 2011) e *skfeature* (LI et al., 2016). Ao todo foram selecionadas seis diferentes métodos descritos a seguir:

1. *Principal Component Analysis* (PCA) que realiza fusão das características transformando o conjunto inicial em um novo conjunto que contém apenas os componentes principais.
2. *Minimum Redundance Maximum Relevance* (mRMR) que realiza seleção de características baseando-se em dois critérios, sendo um deles o critério mínimo de redundância entre pares de características e o outro o critério de relevância máxima por meio da medida de informação mútua.
3. *Correlation-based Feature Selection* (CFS) que realiza seleção de características considerando que um subconjunto adequado contém características altamente correlacionadas com a classe alvo, porém pouco correlacionadas entre si.
4. *Fast Correlation-based Filter* (FCBF) que realiza seleção de características considerando a ideia principal do método CFS, porém calculando a medida de correlação baseada no conceito de entropia da teoria da informação.
5. ReliefF que realiza seleção de características por meio de um *ranking* de relevância de cada característica.
6. *Robust Feature Selection* (RFS) que realiza seleção de características por meio de aprendizado por dicionário esparso.

A figura 21 apresenta o resumo das etapas de redução de dimensionalidade e classificação. Para cada conjunto de características f_e , $f_{e \times c}$, f_m e $f_{m \times c}$, foram executados cada um dos métodos de redução de dimensionalidade em conjunto com os métodos de classificação *Support Vector Machines* (SVM) e *Gaussian Naive Bayes* (NB).

Figura 21 – Redução de dimensionalidade e classificação

Fonte: Tuany Dias Pinheiro

4.4 Treinamento e comparação de classificadores

Para a realização do treinamento dos classificadores, foram escolhidos os três métodos descritos nas subseções 2.3.2, 2.3.2 e 2.3.2. A construção do classificador implica em obter uma combinação de parâmetros adequada considerando que cada um dos parâmetros pode afetar o desempenho da classificação. No caso do classificador SVM, a escolha da função *kernel*, seus parâmetros assim como de outros parâmetros de um classificador SVM, nem sempre é intuitiva e variam de acordo com a função *kernel* utilizada. Considerando as funções linear, radial, polinomial e sigmóide, a seguir estão descritos os parâmetros SVM correspondentes:

C: penalidade atribuída aos erros de classificação; representa um *tradeoff* entre a maximização da margem e a minimização do erro de classificação. Parâmetro necessário para todos os tipos de *kernel*

kernel: As mais comumente utilizadas são radial, polinomial e sigmóide.

γ : coeficiente das funções *kernel* não-lineares.

grau: grau da função *kernel* polinomial.

coef: termo independente das funções *kernel* polinomial e sigmóide.

Para o método *Random Forests*, existem parâmetros que determinam a forma como é feita a divisão dos nós, o crescimento das árvores, o número de árvores da floresta, etc. A seguir são descritos estes parâmetros:

K: quantidade de árvores da floresta

crit: critério de avaliação das importâncias das características para a melhor divisão do nó.

m: quantidade de características a ser considerada para a melhor divisão do nó

i: quantidade máxima de nós da árvore

r: quantidade mínima de amostras para um nó ser considerado uma folha

Após acrescentar a redução de dimensionalidade ao *pipeline*, foram geradas combinações de parâmetros com as quais os classificadores foram testados, como mostra a figura 21. Para realizar a otimização da escolha dos parâmetros, para cada método de classificação foi utilizada a técnica de *Grid search* de forma a selecionar a configuração de parâmetros que correspondesse ao maior valor de acurácia, estimada por validação cruzada *k-fold* estratificada com $k = 4$. As tabelas 2 e 3 apresentam os valores de parâmetros que foram testados para os métodos SVM e *Random Forests*. Para o método *Naive Bayes* não foram utilizadas probabilidades *a priori*.

Tabela 2 – Otimização dos parâmetros SVM

Parâmetro	Valores possíveis
<i>kernel</i>	linear, radial, polinomial, sigmóide
C	0,01 0,1 0,3 0,5 0,7 1 5 10 50 100
γ	0,0001 0,0004 0,001 0,01 0,1 1 3 10 100
<i>coef</i>	0 0,01 0,1 1 10 100
grau	2, 3

Tabela 3 – Otimização dos parâmetros *Random Forests*

Parâmetro	Valores possíveis
m	$\log n, \sqrt{n}$
i	50 120
K	500 1000
r	0,01 0,01325 0,0165 0,01975 0,023 0,02625 0,0295 0,03275 0,036 0,03925 0,0425 0,04575 0,049 0,05225 0,0555 0,05875

A comparação entre os classificadores gerados foi realizada com base nas medidas de acurácia, precisão, revocação, erro e f1-score obtidas na classificação. As medidas foram obtidas por meio de validação cruzada *k-fold* estratificada com $k = 4$. Na seção 5.1 são apresentados os resultados da comparação dos métodos de redução de dimensionalidade e na seção 5.2 são apresentados os resultados da comparação dos classificadores.

5 Resultados e discussão

Neste capítulo são apresentados os resultados e discussão da comparação de métodos de redução de dimensionalidade e classificação em imagens de indivíduos com TEA e em desenvolvimento típico. Na seção 5.1 são comparados e discutidos os resultados da redução de dimensionalidade e na seção 5.2 são discutidos os resultados dos métodos de classificação em cada conjunto de características. Como será mencionado nesta última seção, a maior acurácia foi obtida da classificação com *Random Forests*, que atingiu o valor de 0,801 para o conjunto de distâncias euclidianas.

5.1 Comparação dos métodos de redução de dimensionalidade

Com a finalidade de facilitar a comparação entre os métodos de redução de dimensionalidade, as figuras 22, 24, 25, 26 e 27 apresentam as acurácias médias do uso dos quatro conjuntos de características (com base nas diferentes distâncias) por método de redução após a classificação por *Naive Bayes* e SVMs com *kernels* linear, polinomial, radial e sigmóide, respectivamente, considerando a quantidade de características escolhidas.

Para cada método de redução de dimensionalidade foram testados alguns tamanhos de subconjunto de características. Os testes iniciais apresentaram resultados de acurácia abaixo de 60% para tamanhos de subconjuntos muito grandes, então foram escolhidos tamanhos de subconjunto considerando 10%, 5%, 1% e menos de 1% das características totais o que corresponde a tamanhos de subconjunto 227, 114, 22, 17, 13, 8 e 6. A escolha destes valores ocorreu de forma empírica de acordo com os resultados dos testes iniciais.

Na tabela 4 estão descritos os resultados de cada método e classificador. Os valores apresentados são a média dos resultados entre os quatro conjuntos de distâncias para cada método de redução de dimensionalidade e classificador. Em negrito estão os melhores valores para cada medida de desempenho.

Figura 22 – Acurácia média do classificador *Naive Bayes* para cada método de redução de dimensionalidade

Fonte: Tuany Dias Pinheiro

O método ReliefF apresentou performance muito abaixo dos demais métodos em todos os testes com classificadores, como pode ser observado na figura 22 e de forma a facilitar a visualização, na figura 23 é apresentado novamente os resultados para o método *Naive Bayes* desconsiderando o método ReliefF. A maioria dos métodos obteve resultados bem próximos e por isso as curvas resultantes se sobrepõem sendo que, o método PCA apresenta performance pior que os demais nos subconjuntos com mais características.

Os demais gráficos apresentados nas figuras 24, 25 e 26 também não apresentam os resultados do método ReliefF. Os gráficos considerando todos os métodos são apresentados no apêndice C.

O método RFS obteve os melhores resultados das taxas de acurácia (0,752), revocação (0,647) e f1-score (0,651) com classificador *Naive Bayes* (figura 22), sendo que apenas a taxa de precisão apresentou o melhor resultado com o método CFS (0,666). Nas figuras 22, 25 e 26, observa-se que o RFS obteve seus melhores resultados selecionando mais características em relação aos outros métodos redutores de dimensionalidade, alcançando seu pico com 227 características.

Figura 23 – Acurácia média do classificador *Naive Bayes* para cada método de redução de dimensionalidade

Fonte: Tuany Dias Pinheiro

Diferente do que aconteceu com os classificadores baseados em *Naive Bayes* (figura 22), de modo geral, todos os redutores de dimensionalidade obtiveram seus melhores resultados com SVM função linear selecionando quantidades menores de características (figura 24). Somente o método ReliefF obteve maiores valores de acurácia selecionando mais características. As maiores taxas de acurácia (0,750), revocação (0,627) e f1-score (0,644) foram obtidas com PCA enquanto que a maior precisão (0,697) foi obtida com o método FCBF. Observando as figuras 25, 26 e 27 percebe-se que o método PCA também obteve resultados mais satisfatórios com menor quantidade de características.

Utilizando SVM função polinomial (figura 25), mais uma vez o método PCA obteve os maiores valores de acurácia selecionando poucas características e piorando conforme aumenta-se a quantidade delas. Os demais métodos obtiveram resultados bem próximos, apresentando picos e decaindo nas mesmas regiões, sendo que os piores resultados foram obtidos com ReliefF que em todos os testes apresentou os valores das medidas mais baixos.

Figura 24 – Acurácia média do classificador SVM linear para cada método de redução de dimensionalidade

Fonte: Tuany Dias Pinheiro

Utilizando SVM função polinomial (figura 25), mais uma vez o método PCA obteve os maiores valores de acurácia selecionando poucas características e piorando conforme aumenta-se a quantidade delas. Os demais métodos obtiveram resultados bem próximos, apresentando picos e decaindo nas mesmas regiões, sendo que os piores resultados foram obtidos com ReliefF que em todos os testes apresentou os valores das medidas mais baixos.

Utilizando classificadores SVM com função radial (figura 26), a maioria dos métodos de redução de dimensionalidade obteve resultados próximos, sendo que o método FCBF obteve o maior valor de acurácia, obtendo os maiores valores conforme aumentou-se o número de características. O método ReliefF obteve o menor valor em todos os testes, não importando a quantidade de características selecionadas.

A figura 27 apresenta os resultados utilizando classificador SVM função sigmóide, sendo que as taxas de precisão e revocação obtidas foram 0 e estes resultados foram os piores destes métodos.

Figura 25 – Acurácia média do classificador SVM polinomial para cada método de redução de dimensionalidade

Fonte: Tuany Dias Pinheiro

Figura 26 – Acurácia média do classificador SVM radial para cada método de redução de dimensionalidade

Fonte: Tuany Dias Pinheiro

O método ReliefF obteve o seu melhor resultado com acurácia de 0,645, precisão 0,215, revocação 0,103 e f1-score 0,109 para um subconjunto com seis características. PCA

apresentou os melhores resultados para este tipo de classificador e o melhor resultado foi obtido com um subconjunto de 13 características.

Figura 27 – Acurácia média do classificador SVM sigmóide para cada método de redução de dimensionalidade. Os métodos mRMR, RFS, FCBF e CFS obtiveram exatamente os mesmos valores em todas medidas

Fonte: Tuany Dias Pinheiro

A tabela 4 apresenta as médias entre os quatro diferentes conjuntos de distâncias. Para cada tipo de classificador e método de seleção de características, foram selecionados os melhores resultados por acurácia e tamanho do subconjunto de características e então foram calculadas as médias das medidas de avaliação dentre os quatro conjuntos de distâncias. Exemplificando, para o método CFS com classificador *Naive Bayes* a melhor acurácia foi obtida com o subconjunto de 114 características e dessa forma foi calculada a média dentre os quatro conjuntos de distâncias considerando o mesmo tamanho de subconjunto de características.

O método RFS com classificador *Naive Bayes* obteve os melhores resultados de acurácia (0,752), revocação (0,647) e f1-score (0,651) considerando um subconjunto de 227 características e o método FCBF com classificador SVM linear obteve o melhor resultado de precisão (0,697) considerando um subconjunto de seis características. Os métodos CFS e FCBF obtiveram os melhores resultados de acurácia considerando 114 características enquanto que os métodos mRMR e RFS obtiveram melhores resultados considerando 227

características e os métodos PCA e ReliefF obtiveram melhores resultados considerando um subconjunto com seis características.

Tabela 4 – Resumo considerando melhores acurácias médias dos quatro conjuntos de distâncias por quantidade de características. A coluna método é descrita pela sigla do algoritmo de seleção de características (CFS, FCBF, mRMR, PCA, ReliefF, RFS) e a sigla do tipo de classificador (*Naive Bayes* (NB), SVM linear, SVM polinomial (poli), SVM radial (rbf) e SVM sigmóide).

Método	Acurácia	Precisão	Revocação	F1 score	Qtd. de caracte- rísticas
CFS+NB	0,750	0,666	0,641	0,647	114
CFS+SVM linear	0,696	0,625	0,507	0,540	13
CFS+SVM poli	0,686	0,622	0,507	0,531	114
CFS+SVM rbf	0,686	0,363	0,273	0,300	227
CFS+SVM sigmóide	0,632	0,000	0,000	0,000	6
FCBF+NB	0,748	0,665	0,635	0,644	114
FCBF+SVM linear	0,724	0,697	0,511	0,552	6
FCBF+SVM poli	0,693	0,617	0,531	0,558	227
FCBF+SVM rbf	0,693	0,373	0,279	0,309	227
FCBF+SVM sigmóide	0,632	0,000	0,000	0,000	6
mRMR+NB	0,746	0,658	0,635	0,641	227
mRMR+SVM linear	0,729	0,674	0,543	0,589	13
mRMR+SVM poli	0,693	0,626	0,501	0,534	114
mRMR+SVM rbf	0,682	0,360	0,254	0,284	227
mRMR+SVM sigmóide	0,632	0,000	0,000	0,000	6
PCA+NB	0,729	0,664	0,590	0,616	13
PCA+SVM linear	0,750	0,691	0,627	0,644	6
PCA+SVM poli	0,707	0,631	0,518	0,561	8
PCA+SVM rbf	0,687	0,343	0,279	0,302	22
PCA+SVM sigmóide	0,683	0,313	0,347	0,329	13
ReliefF+NB	0,506	0,310	0,513	0,292	114
ReliefF+SVM linear	0,608	0,044	0,085	0,056	227
ReliefF+SVM poli	0,522	0,224	0,450	0,257	227
ReliefF+SVM rbf	0,501	0,185	0,500	0,270	6
ReliefF+SVM sigmóide	0,645	0,216	0,103	0,109	6
RFS+NB	0,752	0,666	0,647	0,651	227
RFS+SVM linear	0,727	0,656	0,598	0,608	22
RFS+SVM linear	0,691	0,611	0,530	0,552	114
RFS+SVM rbf	0,680	0,359	0,242	0,274	227
RFS+SVM sigmóide	0,632	0,000	0,000	0,000	6

Fonte: Tuany Dias Pinheiro

Apesar do método RFS ter obtido a maior taxa de acurácia individual com o classificador *Naive Bayes*, o método PCA obteve os resultados médios mais satisfatórios entre todos os métodos (figura 28). Os métodos mRMR, CFS, FCBF e RFS possuem resultados muito próximos aos do PCA. Já o método ReliefF obteve os piores resultados, com maior taxa de erro. Esses resultados são provavelmente devido ao fato de que o ReliefF analisa as características individualmente ao invés de subconjuntos das mesmas e que, neste trabalho, uma correta combinação de características seja importante. Essa hipótese faz sentido ao considerarmos que certas distâncias combinadas representam as proporções existentes na face humana.

O fato do método PCA ter obtido, em geral, melhores resultados com poucas características também reflete a natureza dos dados. PCA, como já mencionado na seção 2.3.1, é um método destinado a obter novas características não correlacionadas, cada uma sendo uma combinação linear das características originais, sendo que quanto maior a correlação das características originais maior pode ser a redução. Neste trabalho, cada característica é uma distância entre cada um dos 68 pontos de controle extraídos das imagens faciais. Assim, grande parte dessas distâncias devem ser correlacionadas, permitindo portanto tamanha redução.

Figura 28 – Acurácia, precisão, revocação, f1-score e erro médios agrupados por método de redução de dimensionalidade

Fonte: Tuany Dias Pinheiro

5.2 Comparação dos métodos de classificação

A tabela 5 apresenta, para cada combinação de conjunto de características (com base no tipo de distância), método de classificação e método de redução de dimensionalidade, as medidas de desempenho do melhor classificador dentro todos os obtidos após a calibração dos parâmetros com base na acurácia média entre os quatro *folds* da validação cruzada.

Tabela 5 – Melhores resultados na classificação por conjunto baseados na acurácia

Conjunto	Classificador	Acurácia	Precisão	Revocação	F1-Score	Qtd. de características
Euclidiano X 1000	SVM linear + mRMR	0,783	0,773	0,609	0,666	8
Euclidiano X 1000	SVM poly + mRMR	0,792	0,720	0,732	0,723	13
Euclidiano X 1000	SVM rbf + CFS	0,773	0,713	0,630	0,662	13
Euclidiano X 1000	SVM sigmoid + PCA	0,755	0,647	0,755	0,696	13
Euclidiano X 1000	Random Forest	0,793	0,763	0,661	0,699	2278
Euclidiano X 1000	Naive Bayes + mRMR	0,764	0,688	0,630	0,646	13
Euclidiano	SVM linear + PCA	0,776	0,708	0,666	0,674	13
Euclidiano	SVM poly + mRMR	0,775	0,782	0,589	0,654	22
Euclidiano	SVM rbf + PCA	0,767	0,715	0,616	0,653	22
Euclidiano	SVM sigmoid + PCA	0,765	0,706	0,634	0,662	22
Euclidiano	Random Forest	0,801	0,767	0,684	0,717	2278
Euclidiano	Naive Bayes + mRMR	0,800	0,738	0,705	0,719	6
Manhattan X 1000	SVM linear + PCA	0,784	0,723	0,684	0,701	6
Manhattan X 1000	SVM poly + FCBF	0,785	0,729	0,664	0,685	22
Manhattan X 1000	SVM rbf + CFS	0,755	0,669	0,655	0,660	13
Manhattan X 1000	SVM sigmoid + PCA	0,729	0,626	0,684	0,652	22
Manhattan X 1000	Random Forest	0,783	0,722	0,682	0,699	2278
Manhattan X 1000	Naive Bayes + RFS	0,764	0,673	0,680	0,671	227
Manhattan	SVM linear + CFS	0,764	0,712	0,625	0,657	22
Manhattan	SVM poly + PCA	0,764	0,750	0,550	0,622	6
Manhattan	SVM rbf + FCBF	0,764	0,742	0,575	0,638	227
Manhattan	SVM sigmoid + PCA	0,755	0,698	0,625	0,650	6
Manhattan	Random Forest	0,755	0,700	0,600	0,641	2278
Manhattan	Naive Bayes + mRMR	0,764	0,728	0,625	0,658	6

Fonte: Tuany Dias Pinheiro

O maior valor de acurácia foi obtido da classificação com *Random Forests* que atingiu 0,801 para o conjunto de distâncias euclidianas. A classificação com *Random Forests* também obteve o maior valor de acurácia (0,793) para o conjunto de distâncias euclidianas x 1000, em contrapartida obteve o menor valor (0,755) para conjunto de distâncias Manhattan.

Os maiores valores de precisão (0,782), revocação (0,755) e f1-score (0,723) foram obtidos utilizando os conjuntos baseados em distâncias euclidianas, assim como o menor erro (0,199). Os piores resultados, considerando acurácia, foram obtidos utilizando os

conjuntos baseados em distância Manhattan sendo que, seguindo o padrão comentado na seção 5.1, o classificador SVM função sigmóide obteve a menor acurácia (0,729) dos testes.

A classificação com SVM função polinomial obteve resultados próximos ao da classificação com *Random Forests* sendo o maior valor de acurácia de 0,792. O método *Naive Bayes*, seguindo o padrão comentado na seção 5.1 obteve a segunda melhor acurácia dos testes.

Observando a figura 29 e a tabela 5 nota-se que, para os conjuntos de características baseados em distâncias Manhattan, o método SVM obteve o melhor resultado dentre todos. No entanto a classificação obteve melhores resultados utilizando os conjuntos baseados em distâncias euclidianas e para estes conjuntos, o método SVM obteve resultados piores que o método *Random Forests*.

Figura 29 – Comparação entre classificadores por conjunto de características

Fonte: Tuany Dias Pinheiro

6 Considerações Finais

Neste documento foi apresentado o problema de auxílio ao diagnóstico do transtorno do espectro autista 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, foram utilizados um equipamento de estereofotogrametria que possui um custo alto, o que dificulta a sua aplicação na prática clínica.

Esta pesquisa realizou uma comparação de técnicas de redução de dimensionalidade e classificação em imagens faciais bidimensionais, capturadas por uma câmera fotográfica comum, para classificar fotografias em casos de TEA e desenvolvimento típico. Para isto, foi proposto um protocolo de coleta de imagens. A coleta de imagens foi conduzida em duas localidades, durante um período de seis meses, de forma a aumentar a quantidade de exemplos de casos e controles, sendo ao todo obtidas fotografias de 129 indivíduos que constituem a base de dados desta pesquisa.

Em paralelo à coleta de imagens, foram testados vários métodos de processamento de imagens de forma a definir um *pipeline* que fosse capaz de detectar a faixa de referência e extrair a sua medida em pixels, escolher as melhores imagens, realizar a normalização da imagem, detectar os pontos de controle e calcular as distâncias. Durante este período, o *pipeline* foi refeito de forma a estudar a presença de distorções nas imagens e aumentar a precisão da detecção da faixa e dos pontos de controle.

Após o *pipeline* ter sido definido e as imagens terem sido processadas, foram testados métodos de redução de dimensionalidade combinados aos métodos de classificação. Para os métodos de classificação foi necessário realizar *grid search* para encontrar combinações de parâmetros adequadas. Os métodos de redução de dimensionalidade e classificação foram comparados por meio de medidas de acurácia, precisão, revocação e f1-score e os resultados obtidos das comparações foram apresentados neste trabalho.

Por se tratar de uma pesquisa relacionada à área de saúde, uma das principais dificuldades encontradas foi o recrutamento de indivíduos para a coleta de dados, principalmente para o grupo de controle. Como já mencionado anteriormente, a coleta foi conduzida em sua maioria 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 possuíam 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 portadoras 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 e adolescentes foram necessárias algumas adaptações do protocolo pensado inicialmente para garantir a qualidade dos dados e padronização das imagens. Um exemplo das adaptações necessárias foi a gravação de vídeos para que imagens pudessem ser extraídas, pois alguns dos voluntários não permaneciam com expressão neutra ou em uma posição adequada. Além disso, foi necessário acrescentar a faixa de forma a existir um objeto de referência nas imagens.

Considerando que as diferenças faciais entre indivíduos com TEA e indivíduos em desenvolvimento típico pareciam ser sutis à impressão humana, foram obtidos resultados melhores do que poderiam ser esperados na classificação, sendo que o classificador *Random Forests* obteve uma acurácia de 80%. Além disso, como já mencionando anteriormente, este classificador será utilizado em combinação com outros classificadores que analisarão outros aspectos do TEA. Dessa forma um classificador multimodal será composto na expectativa de que este valor de acurácia possa ser maior.

Sendo assim, o trabalho proposto poderá causar impactos científicos, sociais e econômicos, pois atualmente no Brasil não existem outros métodos ou equipamentos além dos questionários supracitados que não são totalmente adaptados para auxiliar o diagnóstico do TEA. Além disso, existem poucos profissionais e hospitais especializados para a realização deste diagnóstico que muitas vezes é feito tardiamente. Todavia, os resultados apresentados neste trabalho demonstram que é possível utilizar as medidas antropométricas para analisar as diferenças entre indivíduos com TEA e em desenvolvimento típico por meio de métodos de classificação e propor ferramentas que auxiliem o diagnóstico utilizando métodos computacionais.

Trabalhos futuros

Apesar da proposta neste trabalho ter seu foco em TEA, pode-se adaptar a metodologia para o estudo de outros tipos de síndromes e transtornos, desde que as fotografias capturadas possuam a faixa de referência de um tamanho conhecido e sigam o protocolo de captura de imagens.

O *pipeline* de processamento de imagens necessita de mais testes, principalmente, em relação à calibração da câmera. A execução de uma etapa de calibração poderia adicionar maior precisão nas medidas das distâncias, pois as distorções criadas pela lente da câmera seriam minimizadas e, dessa forma, o desempenho da classificação poderia ser melhorado. Além disso, a etapa de detecção dos pontos de controles pode ser substituída por outra biblioteca ou *framework* que aumente a precisão dos pontos obtidos.

Na seção 4.3 foram definidos os métodos de redução de dimensionalidade utilizados para comparação, sendo que todos os métodos que realizam seleção de características apresentados são do tipo filtro. Métodos do tipo camada podem apresentar resultados melhores, já que utilizam o classificador como função de avaliação e as características selecionadas são específicas para cada classificador. Um trabalho futuro poderia realizar comparações entre os métodos do tipo filtro utilizados e métodos do tipo camada.

Referências¹

- 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 18 e 40.
- 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 17, 18, 42, 43, 44, 46 e 73.
- APA. Manual diagnóstico e Estatístico de Transtornos Mentais-: DSM-5. [S.l.]: Artmed Editora, 2014. Citado na página 16.
- 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 16 e 17.
- BALTRUSAITIS, T.; ROBINSON, P.; MORENCY, L.-P. Constrained local neural fields for robust facial landmark detection in the wild. In: Proceedings of the IEEE International Conference on Computer Vision Workshops. [S.l.: s.n.], 2013. p. 354–361. Citado 3 vezes nas páginas 24, 29 e 30.
- BALTRUŠAITIS, T.; ROBINSON, P.; MORENCY, L.-P. Openface: an open source facial behavior analysis toolkit. In: IEEE. Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on. [S.l.], 2016. p. 1–10. Citado na página 56.
- 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 18, 40 e 46.
- BRADSKI, G. Opencv library. Dr. Dobb's Journal of Software Tools, 2000. Citado 2 vezes nas páginas 52 e 54.
- BREIMAN, L. Random forests. Machine learning, Springer, v. 45, n. 1, p. 5–32, 2001. Citado na página 30.
- 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 48.
- CANNY, J. A computational approach to edge detection. IEEE Transactions on pattern analysis and machine intelligence, Ieee, n. 6, p. 679–698, 1986. Citado na página 53.
- CARPIO, L. U. S. Geração de caricaturas para representação de emoções usando processamento de imagens faciais e grafos And-Or. Dissertação (Mestrado) — Universidade de São Paulo, 2016. Citado 2 vezes nas páginas 50 e 51.
- CHAPELLE, O. et al. Choosing multiple parameters for support vector machines. Machine learning, Springer, v. 46, n. 1, p. 131–159, 2002. Citado na página 30.

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

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 16.

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 41 e 46.

CORSELLO, C. M. Early intervention in autism. Infants & Young Children, LWW, v. 18, n. 2, p. 74–85, 2005. Citado na página 19.

DASH, M.; LIU, H. Feature selection for classification. Intelligent data analysis, Elsevier, v. 1, n. 1, p. 131–156, 1997. Citado na página 30.

DASH, M.; LIU, H.; YAO, J. Dimensionality reduction of unsupervised data. In: IEEE. Tools with Artificial Intelligence, 1997. Proceedings., Ninth IEEE International Conference on. [S.l.], 1997. p. 532–539. Citado na página 32.

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 19.

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 19.

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 17.

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 17.

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 22.

DEVROYE, L.; GYÖRFI, L.; LUGOSI, G. A probabilistic theory of pattern recognition. [S.l.]: Springer Science & Business Media, 2013. v. 31. Citado na página 38.

DING, C.; PENG, H. Minimum redundancy feature selection from microarray gene expression data. Journal of bioinformatics and computational biology, World Scientific, v. 3, n. 02, p. 185–205, 2005. Citado na página 34.

DOAK, J. An evaluation of feature selection methods and their application to computer security. [S.l.]: University of California, Computer Science, 1992. Citado 2 vezes nas páginas 32 e 33.

DUDA, R. O.; HART, P. E. et al. Pattern classification and scene analysis. [S.l.]: Wiley New York, 1973. v. 3. Citado na página 30.

FARKAS, L. G. Anthropometry of the Head and Face. [S.l.]: Raven Pr, 1994. Citado 2 vezes nas páginas 22 e 23.

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 17 e 22.

FAWCETT, T. An introduction to roc analysis. Pattern recognition letters, Elsevier, v. 27, n. 8, p. 861–874, 2006. Citado na página 38.

FREUND, Y.; SCHAPIRE, R. E. A decision-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 28.

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 44, 46 e 73.

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 na página 19.

GONZALEZ, R. C.; WOODS, R. E. Digital image processing. [S.l.]: Prentice-Hall, 2 edition, 2002. Citado na página 22.

HALL, M. A. Correlation-based feature selection for machine learning. University of Waikato Hamilton, 1999. Citado 2 vezes nas páginas 34 e 35.

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 18.

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 18.

HO, T. K. Random decision forests. In: IEEE. Document analysis and recognition, 1995., proceedings of the third international conference on. [S.l.], 1995. v. 1, p. 278–282. Citado na página 37.

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 25.

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 31.

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 16.

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 2 vezes nas páginas 33 e 35.

- KONONENKO, I.; ŠIMEC, E.; ROBNIK-ŠIKONJA, M. Overcoming the myopia of inductive learning algorithms with relieff. Applied Intelligence, Springer, v. 7, n. 1, p. 39–55, 1997. Citado na página 35.
- LEVY, S. E.; MANDELL, D. S.; SCHULTZ, R. T. Autism. Lancet, v. 374, p. 1627–1638, 2009. Citado na página 16.
- LI, J. et al. Feature selection: A data perspective. arXiv:1601.07996, 2016. Citado na página 58.
- LORD, C. et al. Autism spectrum disorders. Neuron, Elsevier, v. 28, n. 2, p. 355–363, 2000. Citado na página 16.
- 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 2 vezes nas páginas 16 e 19.
- 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 17.
- METZ, C. E. Basic principles of roc analysis. In: ELSEVIER. Seminars in nuclear medicine. [S.l.], 1978. v. 8, n. 4, p. 283–298. Citado na página 38.
- 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 17.
- MITCHELL, T. M. Machine learning. New York, 1997. Citado na página 30.
- 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 16.
- 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 32.
- NIE, F. et al. Efficient and robust feature selection via joint l2, 1-norms minimization. In: Advances in neural information processing systems. [S.l.: s.n.], 2010. p. 1813–1821. Citado na página 36.
- 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 3 vezes nas páginas 44, 46 e 73.
- 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 43 e 46.
- 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 25.

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 16.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, v. 12, p. 2825–2830, 2011. Citado na página 58.

POWERS, D. M. Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. Bioinfo Publications, 2011. Citado na página 38.

PRESCOTT, B.; MCLEAN, G. Line-based correction of radial lens distortion. Graphical Models and Image Processing, Elsevier, v. 59, n. 1, p. 39–47, 1997. Citado 2 vezes nas páginas 51 e 52.

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 16.

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 17.

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 48.

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 2 vezes nas páginas 47 e 54.

SUZUKI, S. et al. Topological structural analysis of digitized binary images by border following. Computer vision, graphics, and image processing, Elsevier, v. 30, n. 1, p. 32–46, 1985. Citado na página 54.

TOMASI, C.; MANDUCHI, R. Bilateral filtering for gray and color images. In: IEEE. Computer Vision, 1998. Sixth International Conference on. [S.l.], 1998. p. 839–846. Citado na página 53.

TOU, J. T.; GONZALEZ, R. C. Pattern recognition principles. Pattern Recognition in Physics, v. 1, 1974. Citado na página 17.

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 33.

VAPNIK, V. Pattern recognition using generalized portrait method. Automation and remote control, v. 24, p. 774–780, 1963. Citado na página 37.

VAPNIK, V.; CHAPELLE, O. Bounds on error expectation for support vector machines. Neural computation, MIT Press, v. 12, n. 9, p. 2013–2036, 2000. Citado na página 30.

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 4 vezes nas páginas 25, 26, 27 e 28.

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 22.

WENG, J. et al. Camera calibration with distortion models and accuracy evaluation. IEEE Transactions on pattern analysis and machine intelligence, v. 14, n. 10, p. 965–980, 1992. Citado na página 51.

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 32.

YU, L.; LIU, H. Feature selection for high-dimensional data: A fast correlation-based filter solution. In: Proceedings of the 20th international conference on machine learning (ICML-03). [S.l.: s.n.], 2003. p. 856–863. Citado na página 35.

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 18, 41 e 46.

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 40.

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 16.

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 Prof^a Dr^a 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 permanecer 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)

Apêndice C – Gráficos de acurácia média por subconjunto de características dos métodos de redução de dimensionalidade

Figura 30 – Acurácia média do classificador *Naive Bayes* para cada método de redução de dimensionalidade.

Fonte: Tuany Dias Pinheiro

Figura 31 – Acurácia média do classificador SVM linear para cada método de redução de dimensionalidade.

Fonte: Tuany Dias Pinheiro

Figura 32 – Acurácia média do classificador SVM polinomial para cada método de redução de dimensionalidade.

Figura 33 – Acurácia média do classificador SVM radial para cada método de redução de dimensionalidade.

Figura 34 – Acurácia média do classificador SVM sigmóide para cada método de redução de dimensionalidade. Alguns métodos obtiveram os mesmos resultados e suas curvas se sobrepõem (CFS, FCBF, mRMR, RFS)

Fonte: Tuany Dias Pinheiro