UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

Sistema para interação psicológica infantil intermediado por um robô humanóide

Diogo Henrique Godoi

Dissertação de Mestrado do Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional (PPG-CCMC)

ICMC USP
SÃO CARLOS

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Diogo Henrique Godoi

Sistema para interação psicológica infantil intermediado por um robô humanóide

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Ciências – Ciências de Computação e Matemática Computacional. *VERSÃO REVISADA*

Área de Concentração: Ciências de Computação e Matemática Computacional

Orientadora: Profa. Dra. Roseli Aparecida Francelin Romero

USP – São Carlos
Maio de 2023

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

G588s

GODOI, DIOGO HENRIQUE
Sistema para interação psicológica infantil intermediado por um robô humanóide / DIOGO HENRIQUE GODOI; orientador Roseli Aparecida Francelin Romero. -- São Carlos, .
75 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional) -- Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, .

1. Teleoperação. 2. Robôs Assistivos Sociais. 3. robô NAO. 4. Crianças. 5. Psicologia. I. Romero, Roseli Aparecida Francelin , orient. II. Título.

Bibliotecários responsáveis pela estrutura de catalogação da publicação de acordo com a AACR2:
Gláucia Maria Saia Cristianini - CRB - 8/4938
Juliana de Souza Moraes - CRB - 8/6176

Diogo Henrique Godoi

System for child psychological interaction mediated by a humanoid robot

Dissertation submitted to the Instituto de Ciências Matemáticas e de Computação – ICMC-USP – in accordance with the requirements of the Computer and Mathematical Sciences Graduate Program, for the degree of Master in Science. *FINAL VERSION*

Concentration Area: Computer Science and Computational Mathematics

Advisor: Profa. Dra. Roseli Aparecida Francelin Romero

USP – São Carlos
May 2023

Este trabalho é dedicado a todas crianças.
Em especial, aquelas que foram privadas dos seus direitos ou negligenciadas pela sociedade.

AGRADECIMENTOS

Em primeiro lugar, agradeço a minha orientadora, profa. Dra. Roseli A. F. Romero, pela confiança depositada em mim e pela oportunidade dada de participar de um projeto multidisciplinar, do qual esta Dissertação de Mestrado faz parte.

Em segundo lugar, agradeço a psicóloga Márcia Thomé e aos membros do Centro de Tecnologia da Informação Renato Archer: Dr. Hélio Azevedo, Dr. Josué Ramos e Germano Filho pela colaboração durante o desenvolvimento desta pesquisa.

Finalmente, agradeço a todos os membros do Laboratório de Aprendizado de Robôs (LAR-ICMC-USP), pela troca de ideias e conselhos dados e ao ICMC e a CNPq por propocionar os meios, ferramentas e a infra-estrutura necessária.

RESUMO

GODOI, H. D. **Sistema para interação psicológica infantil intermediado por um robô humanóide**. 2023. 73 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2023.

De acordo com a Organização Mundial da Saúde, uma em cada cinco crianças apresenta um problema de saúde mental, o que destaca a importância de desenvolver novas ferramentas para ajudar os profissionais da área. Nesse contexto, o objetivo deste trabalho é desenvolver um sistema integrado que usa um robô humanoide como recurso lúdico para diagnóstico psicológico em crianças. O sistema foi avaliado usando a metodologia *GOAL QUESTION METRIC* para análise de experimentos. O sistema foi proposto para ser utilizado na psicologia forense, onde foi avaliado com crianças que sofreram algum tipo de abuso. O sistema passou por duas avaliações realizadas por psicólogos. A primeira avaliou sua funcionalidade e a segunda avaliou sua utilização em um ambiente real. Os resultados mostram que o sistema é fácil de usar e sua aplicação no âmbito da Psicologia Forense se mostra promissora, com sessões mantendo em média 30 minutos de duração.

Palavras-chave: Teleoperação, Robôs Assistivos Sociais, robô NAO, Crianças, Psicologia.

ABSTRACT

GODOI, H. D. **System for child psychological interaction mediated by a humanoid robot**. 2023. 73 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2023.

According to the World Health Organization, one in five children experiences a mental health problem, highlighting the importance of developing new tools to assist professionals in the field. In this context, the aim of this study is to develop an integrated system that utilizes a humanoid robot as a playful resource for psychological diagnosis in children. The system was evaluated using the GOAL QUESTION METRIC methodology for experimental analysis. The system was proposed to be used in forensic psychology, where it was evaluated with children who have experienced some form of abuse. The system underwent two evaluations conducted by psychologists. The first evaluation assessed its functionality, while the second evaluated its utilization in a real-world environment. The results demonstrate that the system is easy to use, and its application in the field of forensic psychology shows promise, with sessions averaging a duration of 30 minutes.

Keywords: Teleoperation, Socially Assistive Robots, Child, Psycology.

LISTA DE ILUSTRAÇÕES

Figura 1 – Em a) Robô Paro, b) Robô Tega, c) Robô Pleo, d) Robô Huggable e e) Robô NAO. 22
Figura 2 – Visão geral do Projeto "PROTEGER" 24
Figura 3 – Ilustração do Baxter Robot 26
Figura 4 – Ilustração do robô Kaspar 28
Figura 5 – Ambiente de experimentos 34
Figura 6 – Sistema de interação psicológica infantil para sessões de terapia 37
Figura 7 – Robô humanóide NAO 38
Figura 8 – OBS Studio 39
Figura 9 – Ssuite Face Time Video Phone 40
Figura 10 – MorphVOX Pro 41
Figura 11 – Esquema de comunicação da interface GUIPsyin 42
Figura 12 – Configurações da Interface Gráfica de Interação Psicológica Infantil 43
Figura 13 – Movimentos realizados pelo NAO 44
Figura 14 – Representação das camadas de uma Rede Neural Convolucional. Em a) Convolução, em b) Max Pooling, em c) Convolução, em d) Max Pooling, em e) Flatten, em f) Camada totalmente conectada e g) Classificação da imagem 45
Figura 15 – Em a) composição da distribuição das classes antes do balanceamento e em b) composição da distribuição das classes depois do balanceamento 47
Figura 16 – Histórico do treinamento da RNC em a) e da RNCD em b) 49
Figura 17 – Resultado do conjunto de teste da RNC em a) e da RNCD em b) 49
Figura 18 – Sistema de interação psicológica infantil: GUIPsyin em (1), em (2) captura do vídeo da câmera do NAO e em (3) a captura do video através do celular . 56
Figura 19 – Disposição dos equipamentos na sala de Controle. 59
Figura 20 – Em a) temos a Disposição dos equipamentos ligados ao robô NAO na sala de entrevista e em b) a visão geral da sala. 60
Figura 21 – Correlação de Pearson entre as variáveis A, B, C e D 62

LISTA DE TABELAS

Tabela 1 – Exemplo de Diálogo: psicólogo x criança 34
Tabela 2 – Arquitetura da rede RNC proposta por Pranav *et al.* (2020) com as modificações feitas. 47
Tabela 3 – Arquitetura da rede RNCD proposta por Zadeh, Imani e Majidi (2019) com as modificações feitas. 48
Tabela 4 – Tabela com os resultadas da RNC e RNCD obtidos a partir da matriz de confusão 50
Tabela 5 – Questões utilizadas para avaliar o sistema 52
Tabela 6 – Questões utilizadas para avaliar a interação do robô com a criança 53
Tabela 7 – Relação entre as perguntas da Tabela 5 e as características da ISO/IEC 25010 estabelecidas 54
Tabela 8 – Resultados obtidos a partir das questões realizadas 58
Tabela 9 – Resultados obtidos para avaliar a funcionalidade, confiabilidade, usabilidade e eficiência do sistema 58
Tabela 10 – Resultados obtidos da análise dos vídeos obtidos nas interações 61

LISTA DE ABREVIATURAS E SIGLAS

CAL system *Computers Aided Learning Systems*

CTI Centro de Tecnologia da Informação Renato Archer

GQM *Goal Question Metric*

IPq/HCFMUSP Instituto de Psiquiatria do Hospital das Clínicas da Faculdade de Medicina da USP

MOS *Mean Opinion Score*

RNA Redes Neurais Artificias

RNC Rede Neural Convolucional

RNCD Rede Neural Convolucional Densa

SUMÁRIO

1 INTRODUÇÃO 21
1.1 Objetivos 23
1.2 Justificativa 23
1.3 Organização 24

2 TRABALHOS RELACIONADOS 25
2.1 Robôs Teleoperados 25
2.2 Robótica Assistiva 27
2.3 Considerações Finais 31

3 DESENVOLVIMENTO DO SISTEMA PARA INTERAÇÃO PSICOLÓGICA INFANTIL 33
3.1 Método de utilização do Sistema de interação psicológica infantil . 33
3.1.1 Exemplo de diálogo 33
3.1.2 Cenário de uso 34
3.2 Desenvolvimento do Sistema 35
3.2.1 Criação de Filme 35
3.2.2 A credibilidade do discurso 36
3.3 Arquitetura do Sistema para interação psicológica infantil 36
3.4 Componentes do sistema 37
3.5 Softwares utilizados 39
3.5.1 OBS Studio 39
3.5.2 Ssuite FaceTime Video Phone 40
3.5.3 MorphVOX Pro 41
3.5.4 Interface gráfica (GUIPsyin) para manipulação do NAO 42
3.6 Desenvolvimento do detector de emoções 44
3.6.1 Redes Neurais Convolucionais 45
3.6.1.1 Treinamento das ConvNets 46
3.6.1.2 Arquiteturas das RNCs utilizadas 47
3.6.1.3 Avaliação do sistema de emoção 48
3.7 Considerações Finais 50

4 MÉTODO DE AVALIAÇÃO DO SISTEMA 51
4.1 Método de avaliação do sistema de software 51

4.2 Desenvolvimento do método de avaliação 52
4.2.1 Etapa 1 - Objetivos *52*
4.2.2 Etapa 2 - Questões *52*
4.2.3 Etapa 3 - Métricas *53*
4.3 Considerações Finais 54

5 RESULTADOS E EXPERIMENTOS REALIZADOS 55
5.1 Testes Iniciais 55
5.2 Caracterização de um sistema de teleoperação gestual 56
5.3 Testes de aceitação do sistema 57
5.4 Aplicação do sistema na Psicologia Forense 59
5.5 Considerações Finais 63

6 CONCLUSÃO 65

REFERÊNCIAS 69

GLOSSÁRIO 73

CAPÍTULO

1

INTRODUÇÃO

O desenvolvimento de novas tecnologias e sua participação na sociedade vêm se destacando cada vez mais nos últimos tempos, desde a criação de robôs com propósito de entretenimento até para facilitar algumas tarefas no dia a dia. O campo de estudo da Interação Humano-Robô (IHR) procura entender, avaliar e criar ferramentas que promovam a interação entre máquinas e humanos por meio de alguma forma de comunicação (MURPHY *et al.*, 2010).

Com isso, as áreas da educação, saúde e entretenimento estão se reinventando por meio do uso dessas tecnologias. A área de IHR se preocupa em desenvolver sistemas robóticos para interagir com humanos da forma mais natural possível. Nesta área, encontram-se os Robôs Sociais Assistivos (RSA), criados para complementar os esforços feitos pelos profissionais da educação, terapeutas, cuidadores de idosos, entre outros, mediando as situações utilizando robôs (CLABAUGH; MATARIĆ, 2019).

Na saúde, um dos tópicos recorrentes que vem se destacando nos últimos tempos é a saúde mental. Segundo a Organização Mundial da Saúde (OMS), saúde mental é o estado de bem-estar que permite a uma pessoa ter uma vida completa e produtiva, contribuindo para a sociedade (ORGANIZATION, 2019). Portanto, o cuidado com a saúde mental desde a infância é de grande importância, pois é nessa fase que as crianças mais se desenvolvem. A OMS menciona que uma em cada cinco crianças no mundo apresenta algum transtorno mental (ORGANIZATION, 2019). O tratamento em crianças é dificultoso devido à barreira social ainda empregada e ao pouco conhecimento dos jovens sobre o tema, além das várias barreiras que diferem entre ambientes rurais e urbanos (GULLIVER; GRIFFITHS; CHRISTENSEN, 2010).

O desenvolvimento de novas ferramentas para ajudar a identificar e tratar transtornos mentais tem crescido nos últimos anos. Os RSA têm tido um impacto positivo no tratamento de crianças com transtornos mentais (Kabacinska et al., 2020).

Alguns dos robôs que estão sendo utilizados no mercado para auxiliar no processo

de terapia com crianças (conforme Figura 1) são: o Paro[1] (um robô bebê foca) desenvolvido pela AIST, o Tega[2] (um robô andróide) desenvolvido pela Personal Robots, o Pleo[3] (um robô dinossauro) desenvolvido pela Innovo Labs, o Huggable[4] (um robô urso) também desenvolvido pela Personal Robots e o NAO[5] (um robô humanóide) comercializado pela Softbank Robotics. Este último, por ser um robô humanóide, é o mais utilizado nas pesquisas, segundo Clabaugh e Matarić (2019). Além de ser programável, ele dispõe de diversos sensores capazes de criar uma melhor interação.

Figura 1 – Em a) Robô Paro, b) Robô Tega, c) Robô Pleo, d) Robô Huggable e e) Robô NAO.

Este estudo aborda a violência e criminalidade infantil, que prejudicam as relações sociais e familiares. A família tem um papel essencial no desenvolvimento emocional das crianças. A legislação protege os direitos das crianças, mas a violência infantil continua subnotificada, causando danos significativos. Os psicólogos, especialmente os judiciários, precisam ter habilidades específicas para lidar com crianças traumatizadas. Neste estudo, investigamos se um robô humanoide pode facilitar o diálogo com a criança, tornando-o mais agradável e interessante, e auxiliar o trabalho dos psicólogos.

[1] <http://www.parorobots.com/index.asp>
[2] <https://robotic.media.mit.edu/portfolio/tega/>
[3] <https://www.pleoworld.com/pleo_rb/eng/products.php>
[4] <http://robotic.media.mit.edu/portfolio/huggable/>
[5] <https://www.softbankrobotics.com/emea/en/nao>

1.1 Objetivos

Este trabalho teve como objetivo o desenvolvimento de uma interface para a manipulação de um robô humanóide, com o intuito de intermediar a interação entre um psicólogo e uma criança. A interface foi integrada com recursos de IoT para que o robô atue como mediador na interação entre o psicólogo e a criança. A hipótese principal desta pesquisa é que a utilização de um robô humanóide teleoperado, com voz infantil e alguns movimentos, pode ajudar a quebrar o medo ou a resistência de algumas crianças em participar ou se expressar durante as sessões com o psicólogo. Acredita-se que a criança possa se sentir mais à vontade para falar sobre seus problemas com um robô, encarado como um amigo ou um brinquedo, do que com um psicólogo em sua primeira interação.

Como objetivos específicos destacam-se: o desenvolvimento de um sistema que utilizará 4 softwares: GUIPsyin, uma interface gráfica para a manipulação do robô humanoide NAO; SSuite FaceTime Video Phone para captar e transmitir o áudio entre computadores da mesma rede wifi; OBS Studio para gravar o vídeo; e MorphVox Pro para alterar a voz do usuário. Além disso, serão utilizados recursos de hardware como microfone, computador, um robô humanoide NAO e uma caixa de som Bluetooth.

1.2 Justificativa

O Projeto *PROTEGER* é um projeto em andamento que tem como objetivo principal investigar o uso de robôs humanoides em diferentes áreas da Psicologia. O projeto é realizado em conjunto com o Centro de Tecnologia da Informação Renato Archer (CTI), localizado na cidade de Campinas, São Paulo.

Duas frentes de pesquisa são desenvolvidas no âmbito do projeto: *Confident* e *Testimonial*. Na frente *Confident*, o robô NAO é utilizado para investigar casos de Mutismo Seletivo e Fobia Social no contexto da Psicologia Clínica. Esses transtornos podem ser bastante desafiadores para os psicólogos, pois muitas vezes as crianças afetadas têm dificuldade em se comunicar e expressar seus sentimentos. Acredita-se que o uso do robô humanóide possa auxiliar nesses casos, criando um ambiente mais confortável e facilitando a interação entre a criança e o psicólogo.

Na frente *Testimonial*, o robô NAO é utilizado no contexto da Psicologia Forense. Nessa área, o objetivo é investigar como o robô pode ser usado para auxiliar o psicólogo a se comunicar com crianças vítimas de abuso ou violência. Acredita-se que o uso do robô contribui para um ambiente mais amigável e acolhedor no qual a criança pode se comunicar com o psicólogo.

A Figura 2 ilustra de forma esquemática a estrutura do sistema PROTEGER, que envolve a presença de um profissional da área da saúde (psicólogo), um robô NAO e uma criança. O psicólogo interage com o sistema computacional e envia mensagens (por texto ou voz) para o robô; o robô por sua vez repassa a mensagem recebida para a criança. A criança interage com o

robô em resposta à mensagem recebida.

O desenvolvimento do projeto PROTEGER tem grande potencial para impactar positivamente a área da Psicologia, contribuindo para aprimorar os métodos e técnicas utilizados pelos profissionais dessa área.

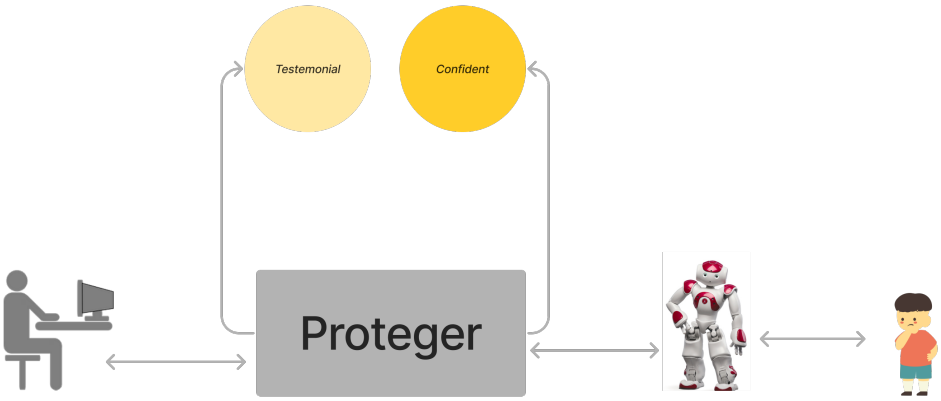


Figura 2 – Visão geral do Projeto "PROTEGER"

Conforme mencionado, o Projeto PROTEGER tem como objetivo abranger duas áreas de pesquisa: a psicologia forense e a psicologia clínica. Nesta dissertação, nos concentramos no âmbito da psicologia forense. Inicialmente, desenvolvemos uma interface denominada GUIPsyin para a manipulação de comandos do robô NAO. Em seguida, montamos um sistema integrado que engloba os elementos de software e hardware necessários para estabelecer a comunicação entre o robô e a criança, assim como entre o robô e o psicólogo. Posteriormente, foram conduzidos experimentos em Botucatu, vinculados a um Convênio de Colaboração firmado entre o ICMC-USP e a UNESP/Botucatu, que contou também com a colaboração do Tribunal de Justiça, disponibilizando um ambiente no qual o experimento foi realizado com crianças que sofreram algum tipo de violência.

Quanto à pesquisa voltada para a psicologia clínica, o sistema foi apresentado ao Instituto de Psiquiatria do Hospital das Clínicas da Faculdade de Medicina da USP (IPq/HCFMUSP). A realização de experimentos ainda não ocorreu, mas está prevista como parte dos trabalhos futuros.

1.3 Organização

Este trabalho está organizado do seguinte modo. No Capítulo 2, são apresentados alguns trabalhos relacionados ao tema abordado. No Capítulo 3, são apresentados os materiais necessários, o sistema proposto, as funcionalidades da interface. No Capítulo 4, a metodologia de avaliação do sistema é descrita. No Capítulo 5, são apresentados os experimentos realizados. Por fim, no Capítulo 6, a conclusão e os trabalhos futuros.

CAPÍTULO

2

TRABALHOS RELACIONADOS

Neste capítulo, são apresentados os trabalhos relacionados ao tema abordado. Inicialmente, são apresentados alguns trabalhos que abordam a teleoperação com foco na mimetização de movimentos humanos através de robôs humanoides. Para isso, foi realizada uma busca na base de dados IEEE e na ACM entre os anos de 2018 e 2022, com os termos "NAO AND TELEOPERATION", a fim de identificar os materiais necessários e como identificar um sistema de teleoperação. Além disso, é apresentado o método conhecido como "Mágico de Oz".

Em seguida, são apresentados os trabalhos que fazem uso do robô NAO no contexto da robótica assistiva. Esses trabalhos foram utilizados como referência para o desenvolvimento do sistema proposto neste trabalho.

2.1 Robôs Teleoperados

Os robôs teleoperados estão se tornando cada vez mais relevantes em nossa sociedade, apresentando possíveis benefícios, como a inclusão de pessoas com deficiência no mercado de trabalho e a descentralização da população urbana. O estudo de Stahl, Anastasiou e Latour (2018) destaca duas formas de como a teleoperação está ocorrendo: a primeira através da teleoperação de robôs avatares e a segunda através da telepresença da robótica móvel. Os robôs avatares são máquinas que possuem uma forma humana, enquanto a robótica móvel permite ao usuário circular com o sistema no ambiente. Esses sistemas representam um avanço na interação entre humanos e robôs, possibilitando a execução de tarefas de maneira mais eficiente e segura.

No contexto da psicologia, os robôs avatares têm desempenhado um papel importante como ferramentas de interação. Com sua forma humana, eles são capazes de expressar suas intenções através de gestos, o que ajuda a transmitir sentimentos e gerar empatia na pessoa interagindo com o robô. De acordo com Stahl, Anastasiou e Latour (2018), a linguagem corporal é responsável por 55% da transmissão do sentimento da máquina.

Para que um robô avatar possa fornecer uma linguagem corporal completa, são necessários três requisitos: o sistema deve ter meios que auxiliem a comunicação verbal, o sistema deve prover meios para que ocorra uma comunicação gestual e intencional apropriada e, por fim, o sistema deve prover meios de comunicação não intencional, como manter o contato visual.

Diversos sistemas foram propostos para mimetizar os movimentos humanos, visando tornar a interação mais natural. Um exemplo é o trabalho de Mohan e Kuchenbecker (2019), que teve como objetivo desenvolver um sistema para estimular idosos a praticarem exercícios. O sistema foi desenvolvido utilizando o robô Baxter, desenvolvido pela Rethink Robotics, que possui atuadores elásticos em série e oito graus de liberdade em cada braço. Essas características permitem que o robô interaja com um usuário humano de forma segura e execute uma grande variedade de movimentos, atendendo às necessidades do projeto. A Figura 3 apresenta uma imagem do robô Baxter.

Em resumo, o uso de robôs avatares na psicologia tem se mostrado promissor, principalmente na área de interação e comunicação com seres humanos. Com o desenvolvimento de sistemas cada vez mais sofisticados, é possível que esses robôs tenham um papel importante no futuro da psicologia clínica e forense.

Figura 3 – Ilustração do Baxter Robot

Os estudos realizados por M., Dajles e Siles (2018) e Alquisiris-Quecha *et al.* (2018) tiveram como objetivo fazer o robô NAO imitar os movimentos de uma pessoa. Para tanto, o primeiro utilizou o Mocap, uma roupa com sensores que captura as coordenadas do corpo humano em movimento e as envia para o NAO. Já o segundo se valeu do Kinect, um sensor desenvolvido pela Microsoft para o Xbox 360. Embora tenham permitido a livre movimentação do NAO, ambos métodos apresentaram limitações. No trabalho de Alquisiris-Quecha *et al.* (2018), por exemplo, os movimentos da perna devem ser suaves, pois qualquer mudança abrupta pode desequilibrar o NAO e fazê-lo cair. Já no estudo de M., Dajles e Siles (2018), o método de calibração do sistema precisava ser mais aprimorado para reduzir a possibilidade de erros durante sua operação. Além disso, seria necessário adicionar um ajuste externo de equilíbrio

para possibilitar movimentos mais complexos, como o de levantar uma perna.

Embora tais tecnologias possibilitem que o usuário manipule o NAO livremente, é importante ressaltar que elas ainda estão em fase de desenvolvimento e precisam ser mais aprimoradas.

Entre as abordagens teleoperadas, destaca-se a metodologia do Mágico de Oz, baseada na história de L. Frank Baum, em que um personagem finge ter habilidades mágicas. Essa técnica consiste em simular que o robô tem capacidade de interagir sem nenhum controlador, ou seja, que ele possui inteligência artificial. O objetivo é que a pessoa interaja com o robô para descobrir as melhores abordagens e reações durante a interação (RIEK, 2012b).

Para que o método Mágico de Oz seja bem-sucedido, são necessários três agentes: o usuário, o mágico (assistente) e o Oz (robô). O usuário é o objeto do experimento, o mágico é a pessoa que controla o robô para que ele fale, faça gestos e se expresse, simulando uma inteligência artificial, e o Oz é o robô que interage com o usuário. Este pode ser parcialmente autônomo ou controlado pelo mágico (MARGE *et al.*, 2017).

2.2 Robótica Assistiva

O uso de robôs no tratamento de crianças com autismo tem sido um dos principais usos da Robótica Assistiva. O trabalho de Aziz *et al.* (2015) teve como objetivo utilizar o NAO como um mediador na interação com crianças, a fim de despertar emoções e sentimentos pelo robô. Para isso, foram utilizados quatro módulos para despertar o interesse da criança. No primeiro módulo, o NAO se apresentava e fazia perguntas, como "Como foi seu dia?"e "Quer aprender algo comigo hoje?". O segundo módulo consistia em ensinar a criança a maneira adequada de pedir algo para seus professores ou pais. No terceiro módulo, a criança aprendia a maneira correta de jogar o lixo fora, onde a lixeira ficava próxima ao NAO. Por fim, no quarto módulo, a criança era convidada a cantar e dançar junto ao NAO. Eles concluíram que, quando o NAO convidava a criança para dançar e cantar, elas demonstravam mais emoções positivas. Também foi constatado que as crianças mostraram mais resposta emocional quando o robô falava ou fazia gestos com as mãos.

O trabalho de Huijnen *et al.* (2017) utilizou o robô KASPAR para interação com crianças autistas, como mostrado na Figura 4. O objetivo foi entender como os robôs assistivos devem ser implementados nas intervenções com crianças autistas. Os resultados mostraram que, além da aparência do robô, a voz, o som, a operação, as ferramentas disponíveis para o robô e seu comportamento devem ser levados em consideração para manter uma personalização das necessidades individuais de cada criança. Concluiu-se que a aparência e o comportamento são cruciais. No entanto, o papel do operador e do ambiente são fundamentais na implementação prática.

Figura 4 – Ilustração do robô Kaspar

Para ajudar a comunicação de crianças autistas, o trabalho de Palestra *et al.* (2017) desenvolveu uma interface gráfica para a terapia digital PECS, que é um sistema de comunicação alternativa. O objetivo era ajudar essa população a superar suas deficiências de comunicação. A ferramenta é composta por um sensor de profundidade que reconhece os gestos das mãos da criança, um monitor para exibir as ferramentas e um robô humanoide (NAO) que serve como intermediário para permitir que as crianças comuniquem suas necessidades. Com essa ferramenta, as crianças podem selecionar imagens exibidas no monitor usando o movimento das mãos e o NAO pronuncia os objetos apresentados. Essa ferramenta permite que as crianças aprendam a se comunicar de maneira mais eficiente e autônoma, o que pode ter um impacto significativo em sua qualidade de vida e na sua interação social.

O estudo de Zhang *et al.* (2019) testou como crianças com autismo e crianças com desenvolvimento normal atribuem falsas crenças em um robô social, utilizando o robô NAO e 20 crianças de 5 a 8 anos de idade. As tarefas consistiam em testes de mudança de localização e conteúdo inesperado, além de uma tarefa de categorização. Os resultados indicaram que as crianças com autismo são menos propensas a atribuir estados mentais aos robôs do que as crianças com desenvolvimento normal, mas ambas as crianças categorizam o NAO da mesma forma.

O trabalho de Martin *et al.* (2020) buscou entender quando as crianças oferecem ajuda a um robô. Para isso, utilizou-se do robô NAO controlado através de rotinas pré-programadas. O procedimento foi composto por quatro fases. A primeira, denominada *free play*, com duração de 10 minutos, consistiu em obter o consentimento do responsável pela criança e familiarizar a criança com o ambiente e o robô, este controlado por outra pessoa em uma sala separada. Na segunda fase, denominada *warm up*, eram selecionados os scripts nos quais se definia o nível de animação do robô. Esse nível consiste em dois parâmetros: o primeiro expressa o nível de autonomia do robô (alta ou baixa), enquanto o segundo expressa o nível de expressividade

(amigável ou neutra). O nível de autonomia consiste em modificar o comportamento do robô a fim de melhorar sua performance na expressividade verbal e não verbal, enquanto o nível de expressividade altera a voz do robô.

A terceira fase, chamada de *test*, começava quando um assistente trazia um xylofone para o NAO, e ele começava a tocar o instrumento para a criança. Em um certo momento, o NAO derrubava a baqueta no chão, momento em que o robô alternava seu foco, olhando para a baqueta e para a criança, com o intuito de que a criança pegasse a baqueta e devolvesse para ele. Assim, qualquer tentativa da criança de devolver a baqueta para o robô era registrada como ajuda. Se a criança não devolvia a baqueta, o experimento se encerrava. Por fim, a última fase era iniciada, chamada de *Animacy interview*, na qual a criança era entrevistada por 10 minutos, enquanto fotos de objetos animados e inanimados eram mostrados às crianças. Isso permitiu entender a percepção da criança sobre o robô como um ser animado. Os dados analisados foram o tempo que a criança levava para entregar a baqueta para o robô, a partir do momento em que ele tocava no chão, e as respostas obtidas na entrevista.

A modificação do nível de autonomia no NAO não mostrou diferença significativa em relação ao tempo de espera para que a criança entregasse a baqueta. Além disso, o nível de expressividade da voz não despertou melhora no comportamento de ajuda. Esse estudo mostrou que, independentemente do nível de autonomia do robô e da expressividade, o comportamento de ajuda de crianças de 3 anos de idade não se altera.

O artigo de Wijnen *et al.* (2019) apresenta uma investigação sobre os comportamentos explicativos de crianças entre 6 e 10 anos, utilizando um robô social como ferramenta de aprendizado. O comportamento explicativo, que se refere à habilidade da criança em expressar suas próprias ideias e conhecimentos, é fundamental para o processo de aprendizado e compreensão de novos conceitos. Os autores argumentam que o uso de um sistema de aprendizado auxiliado por computador, conhecido como *Computers Aided Learning Systems* (CAL system), pode ser uma forma efetiva de melhorar esse comportamento.

Os CAL systems fornecem aos usuários modelos de ideias, informações relevantes e podem até mesmo restringir o ambiente de aprendizado para aumentar a concentração. Ao combinar um CAL system com um robô social, os autores mostram que é possível melhorar ainda mais o processo de aprendizagem das crianças, pois o robô social pode oferecer recursos como contato visual e presença física, o que melhora a interação com o tutor.

Os resultados da pesquisa indicam que, quando as crianças são expostas a um robô social, elas verbalizam por mais tempo do que quando estão interagindo apenas com um CAL system. Além disso, as crianças forneceram mais detalhes em sua fala e usaram argumentos mais lógicos. Esses resultados sugerem que a combinação de um robô social com um CAL system pode ter um impacto positivo no aprendizado das crianças. Em suma, este estudo fornece evidências encorajadoras sobre o uso de tecnologia para melhorar a aprendizagem e o comportamento explicativo em crianças.

O estudo de Dio *et al.* (2020) teve como objetivo analisar como a confiança é adquirida, perdida e restaurada em crianças quando interagem com humanos e robôs humanoides. Para isso, foram selecionadas 94 crianças em idade escolar, divididas em quatro grupos etários: 3, 5, 7 e 9 anos. As crianças foram avaliadas em duas tarefas durante um período de duas semanas. Na primeira semana, foram submetidas à escala de AMS e ao jogo da confiança, enquanto na segunda semana passaram por um teste de falsa crença e avaliação da qualidade da relação de apego. O jogo da confiança consistiu em três agentes, dois jogadores e um mestre do jogo, onde as crianças deveriam adivinhar em qual caixa o mestre do jogo esconderia uma boneca. O jogo foi dividido em três fases: aquisição, perda e restauração da confiança da criança em relação ao seu parceiro.

Em geral, não foram encontradas diferenças significativas nos níveis de confiança entre humanos e robôs, no entanto, as crianças de 3 anos tendiam a confiar mais em humanos, enquanto as de 7 anos confiavam mais nos robôs. Além disso, os resultados indicaram que o robô foi percebido como um parceiro mais perdoável do que um humano. Conclui-se que o robô pode ser considerado um parceiro confiável e estável, com possibilidade de uso em diferentes contextos, inclusive em ambientes de educação e cuidados com crianças.

O trabalho de Logan *et al.* (2019) teve como objetivo investigar a introdução de robôs sociais na internação pediátrica. Para isso, foram utilizadas 54 crianças com idades entre 3 e 10 anos, que foram expostas a três tipos diferentes de intervenções. Na primeira intervenção, foi utilizado o robô social Huggable, na segunda, o avatar do robô social em um tablet, e na terceira, um urso de pelúcia com a presença de um humano.

O estudo avaliou os afetos positivos e negativos, a ansiedade e a intensidade da dor antes e após as intervenções. Os resultados obtidos mostraram que a utilização do robô social em comparação com o avatar e o urso de pelúcia demonstrou um nível maior de afeto positivo nas crianças, além de relatarem vários benefícios no cenário pediátrico.

Concluiu-se que a utilização de robôs sociais pode ser uma ferramenta promissora para abordar as necessidades emocionais das crianças hospitalizadas, potencializando o acesso às intervenções emocionais. Essa tecnologia pode contribuir para o bem-estar e conforto emocional das crianças, criando novas formas de interação e apoio para elas durante o processo de internação.

O trabalho de Pino *et al.* (2019) levou em consideração a utilização de robôs assistivos no âmbito da população idosa. Esse estudo avaliou a eficácia da interação humano-robô no treinamento da memória. Os resultados mostraram que, utilizando o NAO, houve um aumento no contato visual e, em alguns casos, uma redução de sintomas de depressão.

Já o estudo realizado por Soler *et al.* (2015) investigou a utilização do robô humanoide (NAO), do robô de estimação (PARO) e um animal real (cachorro) em sessões de terapia com pacientes diagnosticados com demência em dois ambientes distintos: casa de repouso e centro de

cuidados. O objetivo foi avaliar as mudanças de comportamento, apatia e qualidade de vida. Os pacientes que viviam na casa de repouso foram analisados durante 3 meses em dois grupos de testes. No primeiro grupo, foi comparada a utilização do NAO, do PARO e a terapia convencional. Ambos os robôs promoveram uma diminuição na apatia, mas houve um aumento nos delírios no grupo tratado com o NAO e em ambos os grupos com robôs houve um aumento na irritabilidade. No outro grupo, foi avaliada a utilização do robô de estimação (PARO), do animal real de estimação (cachorro) e a terapia convencional. Houve uma diminuição na qualidade de vida utilizando o robô de estimação em comparação à terapia convencional, um aumento de delírios e irritabilidade com a utilização do PARO e do cachorro em comparação à terapia convencional e um aumento na desinibição no grupo que utilizou somente o robô PARO. No estudo feito em centro de cuidados, os pacientes apresentaram uma melhora na irritabilidade e nos sintomas globais da neuropsiquiatria somente depois de participarem da sessão com o NAO.

No trabalho de Provoost *et al.* (2017) foi realizada uma busca sobre os agentes conversacionais incorporados (ECA), que são agentes animados capazes de demonstrar habilidades humanas como ter uma conversa cara a cara com a capacidade de produzir uma resposta verbal e não verbal, em aplicações psicoterapêuticas. Essa busca levou em consideração 49 estudos associados aos seguintes temas: autismo, depressão, transtorno de ansiedade, transtorno de estresse pós-traumático, transtorno psicótico e uso de substâncias. Como conclusão, foi constatado que a tecnologia ECA avançada tem um caráter futurístico e complexo, o que dificulta provar sua eficácia e segurança na utilização clínica prática. Assim, uma estratégia sugerida envolve aumentar a base de evidências por meio de intervenções com o uso de agentes com baixa tecnologia, rapidamente desenvolvidos, testados e aplicados na prática clínica.

Já em Stower *et al.* (2021), foi realizado um estudo analisando 20 artigos com o objetivo de avaliar a "confiança"na interação entre robôs e crianças. O autor ressalta que a aparência humana do robô nem sempre influencia positivamente a confiança da criança. Além disso, aponta a necessidade de melhorar os resultados e a replicação, projetando estudos com tamanhos de amostra maiores, definições consistentes e relatórios estatísticos transparentes.

O trabalho proposto aqui adota a estratégia da técnica de Mágico de Oz em classes específicas de transtornos mentais, como o Mutismo Seletivo e a Fobia, além de sua aplicação em um ambiente forense. O uso do Mágico de Oz permite a manipulação de recursos que ainda não são tecnicamente possíveis, viabilizando o levantamento de dados para que, no futuro, seja possível determinar claramente quais recursos devem ser priorizados em uma programação totalmente autônoma.

2.3 Considerações Finais

No capítulo apresentado, foram abordados diversos estudos que utilizam robôs teleoperados na área de Robótica Assistiva. Destacam-se, em especial, a metodologia do Mágico

de Oz e trabalhos voltados para o aprendizado e tratamento de crianças autistas, bem como o cuidado com idosos. Esses trabalhos apresentam uma ampla gama de possibilidades para o desenvolvimento de ferramentas capazes de auxiliar na vida diária dessas pessoas, mostrando a importância e o potencial da robótica na promoção do bem-estar e na melhoria da qualidade de vida.

CAPÍTULO

3

DESENVOLVIMENTO DO SISTEMA PARA INTERAÇÃO PSICOLÓGICA INFANTIL

Neste capítulo, o sistema proposto de interação psicológica infantil juntamente com suas funcionalidades é apresentado. O objetivo deste sistema é possibilitar a interação de um psicólogo com uma criança por meio de um robô humanoide. Os módulos que compõem o sistema são descritos detalhadamente. Em seguida, os materiais e softwares necessários são apresentados.

3.1 Método de utilização do Sistema de interação psicológica infantil

A metodologia utilizada consiste em utilizar a técnica do Mágico de Oz (RIEK, 2012a). Nessa técnica, um operador que permanece à distância teleopera o agente robótico ao receber comandos online. No nosso caso, o usuário é a criança e ela acredita que o robô atua de forma autônoma, independentemente de comandos externos. O psicólogo, que desempenha o papel de Mágico, conduz as sessões dando vida ao agente robótico. O robô NAO, apresentado na Figura (7), representa o meio pelo qual a interação entre o assistente (Mágico) e o usuário deve ocorrer.

3.1.1 Exemplo de diálogo

Na Tabela 1, é apresentado um exemplo de questões que o psicólogo pode explorar com a criança. Os primeiros três grupos de questões (*Pessoal, Escola e Família*) procuram identificar o contexto social da criança. Já o grupo *Direcionado* busca aprofundar o contato do psicólogo com a criança. Longe de estabelecer um roteiro rígido, o objetivo é enfatizar que a interação não pode ser delimitada por um questionário pré-estabelecido, mas deve fluir naturalmente de acordo com os aspectos cognitivos expressos pela criança durante a interação.

Tabela 1 – Exemplo de Diálogo: psicólogo x criança

Grupo	Pergunta
Pessoal	Qual seu nome?
	Quantos anos você tem?
	Você gosta de brincar do que?
Escola	Onde você estuda?
	Quantos amigos você tem?
	Com quem você gosta de brincar?
Família	Quem mora com você?
Direcionado	Posso ser seu amigo?
	...

3.1.2 Cenário de uso

O cenário de uso exige um ambiente em que o psicólogo possa controlar o robô NAO sem que a criança perceba essa operação remota. Porém, a criança deve permanecer em um ambiente adequado à sua idade. A Figura 5 apresenta esse ambiente, que consiste em duas salas: a Sala de Entrevista e a Sala de Controle.

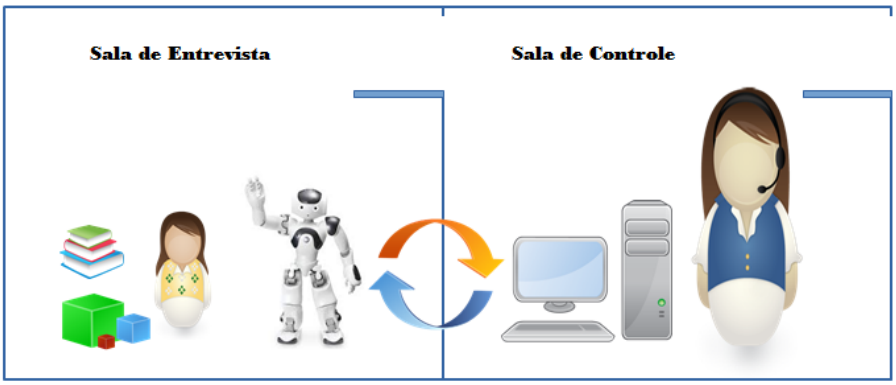


Figura 5 – Ambiente de experimentos

Na Sala de Controle, o psicólogo usa um fone de ouvido e uma interface de computador para comandar remotamente o robô NAO como uma marionete. A Sala de Entrevista inclui um ambiente adequado ao universo infantil com a presença de alguns brinquedos e do robô NAO. Essa sala também contém um sistema de câmeras e microfones para capturar a interação, dados essenciais para análise posterior da sessão pelo psicólogo.

O cenário de uso é concluído com a definição das etapas que configuram uma sessão (happy path[1]):

Para que a operação ocorra, são definidos os seguintes passos:

[1] Um cenário padrão em que não surgem condições excepcionais ou de erro

1. O psicólogo apresenta um filme que mostra o robô humanóide NAO, motivando a criança a conhecê-lo pessoalmente.

2. O psicólogo pergunta à criança se ela tem interesse em conhecer o NAO.

3. Em caso afirmativo, o psicólogo leva a criança para a Sala de Entrevista, onde o robô NAO aguarda.

4. O psicólogo se dirige à Sala de Controle.

5. O psicólogo opera remotamente o robô NAO com o apoio de uma interface que converte o som de sua voz diretamente em um alto-falante acoplado ao robô. Ele também pode ativar movimentos simples no NAO. É importante que a voz do psicólogo seja alterada para que seja compatível com as expectativas da criança em relação ao agente lúdico.

3.2 Desenvolvimento do Sistema

Nesta seção, são apresentados os componentes do sistema e a interface gráfica desenvolvida para facilitar a operação do robô NAO. Conforme apresentado na subseção 3.1.2, o desenvolvimento envolve duas frentes: a criação de um filme que convida a criança a interagir com o NAO e o desenvolvimento de uma interface que permite a operação remota do robô NAO como uma marionete pelo psicólogo.

3.2.1 Criação de Filme

Neste filme, o robô NAO se apresenta e motiva a criança a iniciar uma interação. Este filme deve ser curto com no máximo 2 minutos. O local de projeção deve ser uma sala limpa, ou a própria Sala de Entrevista. Como exemplo de storyboard para o filme, tem-se:

- Movimento: o robô NAO entra no foco da câmera caminhando de lado. Ele cessa o movimento, se vira e acena.

- Fala: “ Olá! Meu nome é NAO ”

- Fala: “ Você sabia que sou amigo de *<nome do psicólogo>* ”

- Fala: “ Hmmmm ... Acho que tenho a mesma idade que você.”

- Fala: “ Você quer me conhecer? ”

- Movimento: NAO coloca as mãos nos quadris fazendo a posição do super-homem.

O filme pode ser conferido através do link: https://youtu.be/OjhG0-YynYE

3.2.2 A credibilidade do discurso

A voz do robô representa um elemento importante para dar veracidade ao personagem interpretado através da técnica do Mágico de Oz. Idealmente, o discurso feito pelo psicólogo na Sala de Controle deveria ser reproduzido em tempo real pelo NAO, considerando não apenas o significado do discurso, mas também a prosódia (FERREIRA; CHATELARD, 2019).

No entanto, a voz não pode ser igual à do psicólogo, pois a criança pode inferir que representa a mesma pessoa. Assim, a voz deve ser alterada digitalmente antes de ser reproduzida no robô NAO. Existem muitas opções para alterar a voz, como adulto x criança, masculino x feminino ou humano x robotizado.

Ao considerar a idade da voz, é importante notar que se deseja que a criança se relacione com o robô como um amigo ou um alter ego (por exemplo, o Grilo Falante da história infantil Pinóquio). Assim, uma voz infantil é mais apropriada. Na verdade, o robô não dará conselhos como uma avó, nem dará ordens como um adulto, mas agirá mais de perto como um amigo confidente. Ao selecionar a voz infantil, a dualidade masculino x feminino perde importância devido à neutralidade da voz na faixa etária de 6 a 10 anos.

Por fim, uma pesquisa de Trovato *et al.* (2017) com adultos interagindo com uma recepcionista virtual mostrou como resultado que há uma preferência pela voz humana na interação quando comparada à voz robótica.

3.3 Arquitetura do Sistema para interação psicológica infantil

Na Figura 6, é apresentada uma visão geral do sistema, na qual todos os equipamentos e softwares necessários são representados. A Sala de Entrevista (indicada por (1)) contém um notebook, um microfone, uma caixa de som bluetooth, um celular, um tripé, um roteador e o robô NAO. O microfone é conectado ao notebook por meio de um cabo USB, que será utilizado para captar o som da sala e transmiti-lo para o notebook da Sala de Controle por meio do software *Ssuite FaceTime Video Fone*. A caixa de som, por sua vez, via *bluetooth*, transmite a voz do mágico (psicólogo) que está na Sala de Controle para a Sala de Entrevista.

Figura 6 – Sistema de interação psicológica infantil para sessões de terapia

Fonte: Elaborada pelo autor.

Na Sala de Controle (indicada por (2)), ocorre a teleoperação do NAO por meio da GUIPsyin. O OBS Studio é utilizado para realizar a gravação da sessão por meio da câmera do robô NAO e do celular localizado na Sala de Entrevista. Além disso, também é realizada a gravação do áudio do *MorphVOX Pro*, software utilizado para alterar a voz do usuário, e do microfone da Sala de Entrevista.

3.4 Componentes do sistema

Para o funcionamento do sistema são necessários os seguintes dispositivos:

- 2 Nootebooks Dell Inspiron 14
- 1 Microfone Estúdio USb Yt K669 Cardioide
- 1 Caixa de Som Bluetooth
- 1 Celular Xiamoi M3
- 1 Tripé para celular
- 1 Robô humanóide NAO
- 1 Roteador

O robô humanoide NAO foi utilizado para intermediar a interação entre o psicólogo e a criança, a fim de captar a atenção da criança e fazê-la sentir-se mais à vontade para interagir. O

Figura 7 – Robô humanóide NAO

robô possui 7 sensores, sendo um em cada pé, um na cabeça e dois em cada mão, além de duas câmeras na cabeça, dois microfones e dois alto-falantes. O robô é programável em 4 linguagens de programação: Java, Javascript, C++ e Python[2].

A empresa que desenvolve o robô, a SoftBank Robotics, disponibiliza o ambiente *Choreographe*, uma aplicação desktop multiplataforma disponível para Windows, Linux e Mac. O *Choreographe* permite criar diversos comportamentos para o robô, como interação com pessoas, criação de diálogos ou envio de e-mails, sem que seja necessário escrever uma linha de código. Além disso, é possível integrar o código Python com esses comportamentos.

Os comportamentos escritos com o *Choreographe* utilizam a linguagem gráfica de blocos, que o NAOqi, o framework do robô, interpreta e executa. É possível ter acesso a toda API do NAOqi utilizando o *Choreographe*, mas a execução pode ser um pouco mais lenta se comparada a códigos em C++ ou Python.

Neste trabalho, foi utilizada a versão 2.1.14 do NAOqi.

[2] Fonte: <http://doc.aldebaran.com/2-1/home_/nao.html>

3.5 Softwares utilizados

A seguir, serão apresentados os softwares utilizados no sistema integrado para a capitação, transmissão e gravação do sistema multimídia.

3.5.1 OBS Studio

Para a gravação da sessão de terapia proposta, utilizou-se o software Open Broadcaster Software[3] (OBS), uma ferramenta de código aberto e livre para gravação e transmissão de vídeos em tempo real. Com o OBS, é possível capturar e misturar áudio e vídeo em tempo real, além de oferecer diversas funcionalidades, como mixer de áudio com filtros por fonte, como portão de ruído, supressão de ruído e ganho, e a possibilidade de implementação de plugins.

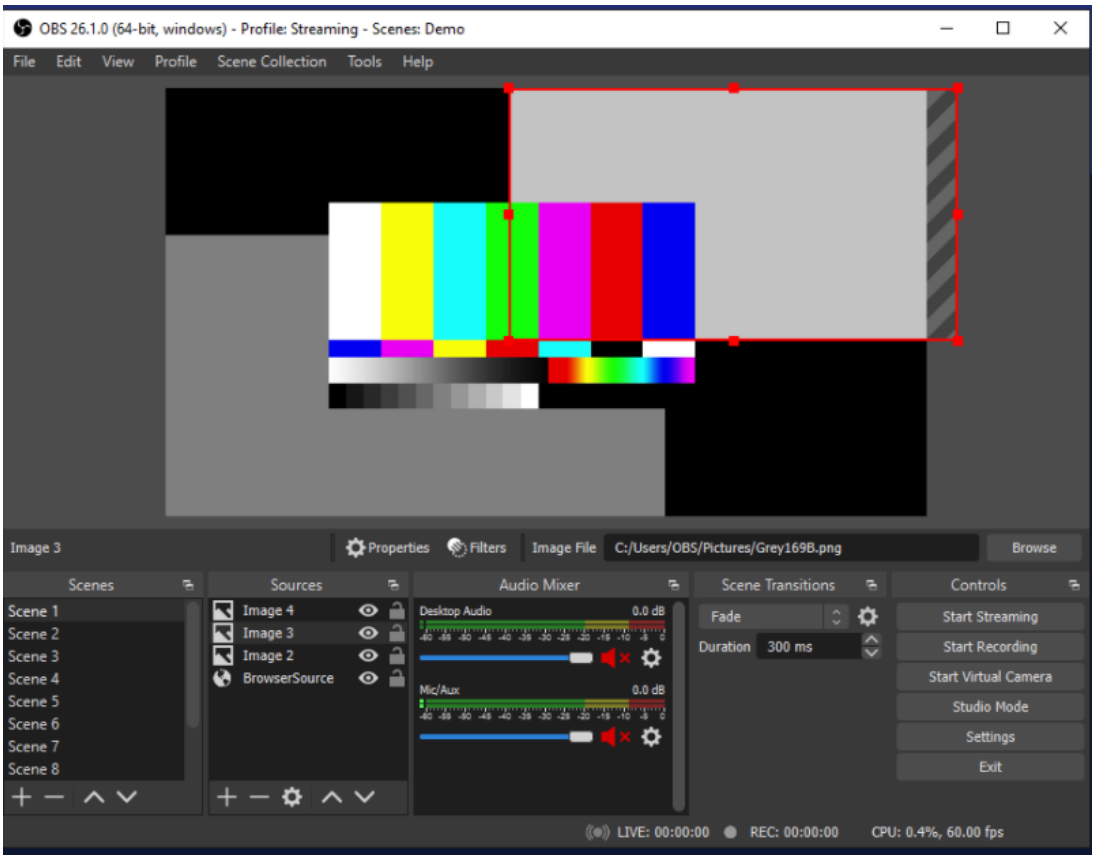


Figura 8 – OBS Studio

Para capturar a imagem da sala, utilizou-se um celular em conjunto com o aplicativo *DroidCam OBS*[4], desenvolvido pela Dev47Apps. Esse aplicativo permite que o celular transmita a imagem para o OBS por meio do plugin instalado no OBS Studio.

[3] Fonte:<https://obsproject.com/pt-br>
[4] Fonte:<https://www.dev47apps.com/obs/>

3.5.2 Ssuite FaceTime Video Phone

O *SSuite FaceTime Video Phone*, apresentado na Figura 9, é um software que permite que dois computadores conectados na mesma rede Wi-Fi realizem chamadas de áudio. Na interface do programa, o campo "Local IP Address"(1) exibe o endereço IP local do usuário, enquanto o campo "IP to Call"(2) requer que o usuário insira o endereço IP da outra pessoa com quem deseja estabelecer a conexão. Ao clicar no botão "Chat"(3), a chamada é iniciada. É importante ressaltar que ambos os usuários devem estar usando o *SSuite FaceTime Video Phone* para que a chamada funcione corretamente.

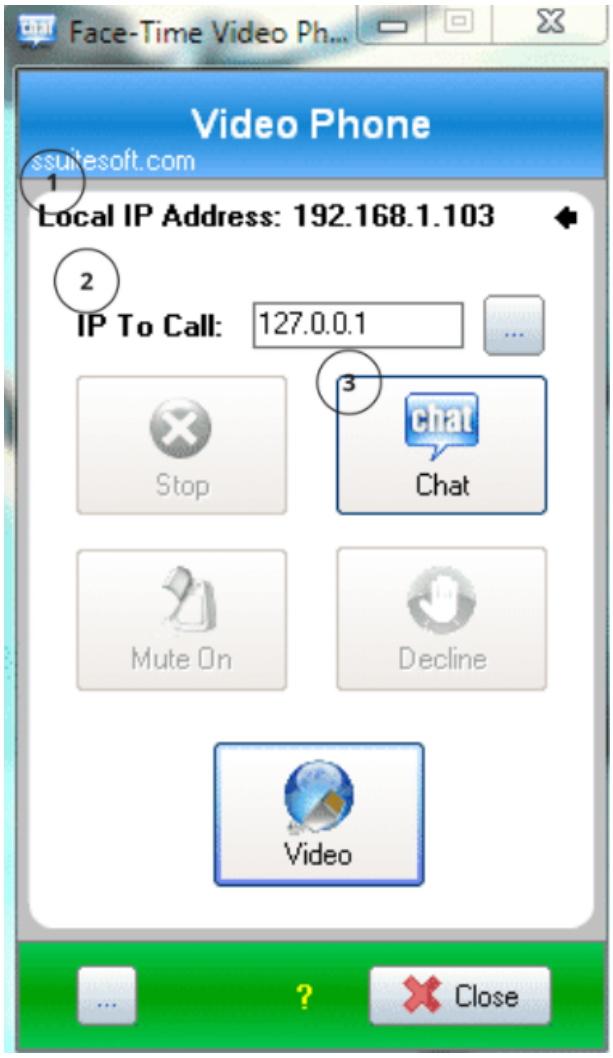


Figura 9 – Ssuite Face Time Video Phone

3.5.3 MorphVOX Pro

O MorphVOX Pro[5], como apresentado na Figura 10, é um software de alteração de voz desenvolvido pela Screaming Bee que oferece diversas funcionalidades, permitindo que o usuário adicione efeitos, altere o timbre e o tom da voz, deixando-a mais aguda ou grave. Além disso, o software já inclui algumas vozes pré-instaladas, como as de criança, robô, homem e mulher.

Figura 10 – MorphVOX Pro

No contexto em questão, o MorphVOX Pro será utilizado para tornar a voz do usuário mais amigável para a criança. As configurações utilizadas foram selecionar a voz de criança e alterar o valor do *Pitch Shift* para 0.4, enquanto o *Strengh* do *Timbre* foi configurado para 10, deixando a voz mais aguda. Essas configurações específicas foram escolhidas para criar um ambiente mais agradável e amigável durante a interação com a criança.

[5] Fonte:<https://screamingbee.com/morphvox-voice-changer>

3.5.4 Interface gráfica (GUIPsyin) para manipulação do NAO

O desenvolvimento da interface contou com a utilização de dois frameworks distintos: o Qt Designer versão 4.8.7 e o NAOqi versão 2.1.14. Cada um deles desempenha funções específicas e complementares no funcionamento da interface. O Qt Designer, por sua vez, é responsável por criar e gerenciar a interface gráfica da aplicação, proporcionando uma experiência visual amigável e intuitiva para o usuário. Já o NAOqi assume a responsabilidade pelos movimentos e comunicação do robô NAO, permitindo que ele possa receber comandos e interagir com o ambiente ao seu redor de forma eficiente e precisa. Assim, ao combinar esses dois frameworks, foi possível criar uma interface integrada e funcional que oferece uma interação fluida e intuitiva com o robô NAO.

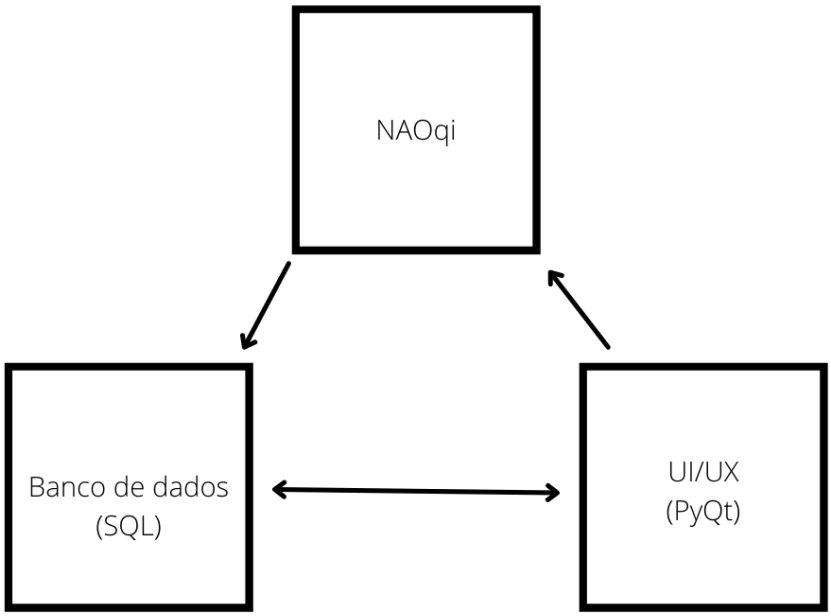


Figura 11 – Esquema de comunicação da interface GUIPsyin

Fonte: Elaborada pelo autor.

Na Figura 11, representa-se a comunicação entre os componentes que compõem a interface. Como mencionado anteriormente, o PyQt foi utilizado para o desenvolvimento da interface gráfica. Através dela, o usuário envia comandos para o NAOqi, responsável pela comunicação e pelos movimentos do robô NAO. Após o envio do comando, a resposta é encaminhada para o banco de dados, que registra se o comando foi enviado com sucesso ou se algum erro ocorreu. Em seguida, o banco de dados envia essa informação de volta para a interface gráfica, que informa o usuário sobre o sucesso ou falha da operação.

Figura 12 – Configurações da Interface Gráfica de Interação Psicológica Infantil

Fonte: Elaborada pelo autor.

A interface é dividida em cinco partes, conforme apresentado na Figura 12. À esquerda da figura (identificada como 1), estão as configurações definidas pelo usuário. É possível estabelecer a conexão da interface com o robô NAO inserindo o endereço IP fornecido pelo próprio robô no campo "IP do robô". Para obter esse endereço, basta pressionar o botão localizado no peito do robô. Após estabelecer a conexão, basta pressionar o botão "Conectar"da interface. Com a conexão ativa, é possível acessar a imagem da câmera "0"do robô, localizada em sua testa, através do botão "Câmera NAO". Além disso, o sistema de detecção de emoções é ativado, mostrando a emoção prevista no rosto da pessoa que está sendo observada pelo NAO. Por fim,

para encerrar a sessão, basta pressionar o botão "Desconectar", que realiza a desconexão entre a interface e o robô NAO.

Na parte (2) da Figura 12, temos a inicialização da sessão de terapia. Nela, o usuário tem acesso ao modo de vida autônomo do robô NAO, ou seja, é possível ativar o detector de presença humana. Ao pressionar o botão "Iniciar Vida", o robô NAO ativa todos os seus leds e entra em estado autônomo, ficando em alerta para procurar por um rosto humano e começar a interagir. Os botões "Girar para Esquerda"e "Girar para a Direita"fazem o NAO girar o corpo para os respectivos lados. Quando não for mais necessário continuar com a interação, o botão "Encerrar Vida"coloca o NAO em estado de dormir, desligando os leds e fazendo-o sentar. Em (3) desta mesma figura, encontramos o botão "Desligar/Emergência", que é uma medida de segurança para situações em que o robô NAO precisa ser desligado imediatamente. Ao pressionar esse botão, todos os motores do NAO são desligados e ele é colocado para sentar.

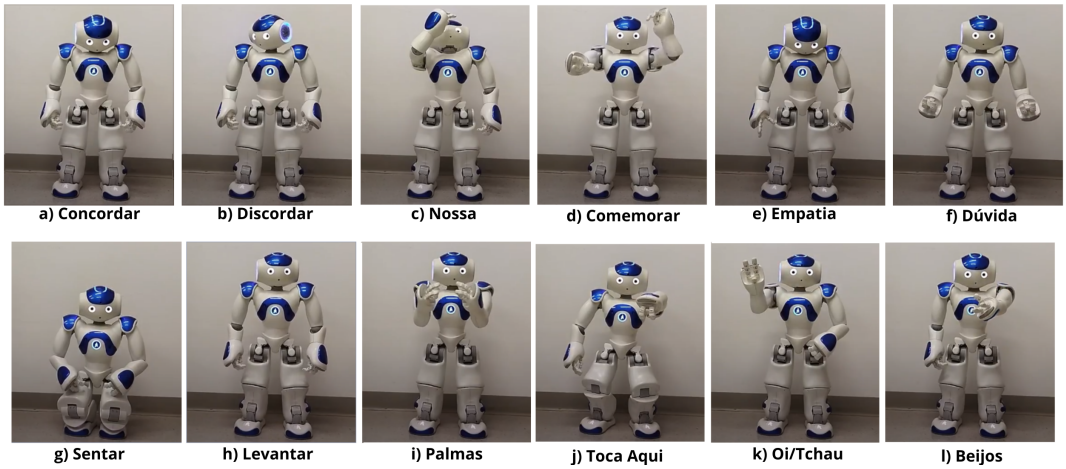


Figura 13 – Movimentos realizados pelo NAO

Fonte: Elaborada pelo autor.

Do lado direito da mesma figura, encontra-se a seção de Movimentos (ver (4)), onde estão presentes 11 botões que controlam os movimentos do NAO, incluindo "Concordar", "Discordar", "Nossa", "Comemorar", "Empatia", "Dúvida", "Sentar", "Palmas", "Toca aqui", "Oi/Tchau"e "Beijos", como demonstrado na Figura 13. Por fim, na seção denotada por (5), temos a área de Avisos, onde são registrados os movimentos que o usuário enviou para o NAO, as mensagens de alerta, erros e sucesso, por exemplo, se a conexão foi estabelecida ou se o movimento foi enviado com sucesso.

3.6 Desenvolvimento do detector de emoções

Nesta seção, é apresentado o módulo de detecção de emoção integrado na interface. Esse módulo foi desenvolvido utilizando Redes Neurais Artificias (RNA), que são sistemas baseados na forma como o cérebro funciona. Essas redes são compostas por um grande número de nós

computacionais interconectados, denominados neurônios, que trabalham de forma distribuída para aprender com o dado de entrada e otimizar em uma saída final.

Mais especificamente, foi utilizado um tipo de RNA chamado Rede Neural Convolucional (RNC). Esse tipo de rede neural é altamente eficiente na análise de imagens, sendo capaz de extrair informações relevantes de forma automática e precisa. A detecção de emoção utiliza uma RNC treinada em um grande conjunto de dados de faces humanas previamente rotuladas com emoções correspondentes.

Dessa forma, quando o módulo é ativado na interface, a imagem capturada pela câmera do robô é enviada para a RNC, que analisa a imagem e identifica a emoção mais provável na face da pessoa que está interagindo com o robô. Essa emoção é então exibida na interface, sobreposta à imagem da pessoa, permitindo que o usuário do robô visualize a emoção identificada pelo sistema.

3.6.1 Redes Neurais Convolucionais

Uma Rede Neural Convolucional, ou RNC, é uma rede de aprendizado profundo capaz de receber uma imagem de entrada e atribuir importância aos aspectos da imagem, diferenciando diferentes regiões da imagem. Dessa forma, o pré-processamento em uma RNC é menor em comparação a outros algoritmos de classificação, já que nos métodos convencionais são utilizados filtros para identificar cada região da imagem. Entretanto, as RNCs têm a capacidade de aprender esses filtros de forma autônoma (O'SHEA; NASH, 2015).

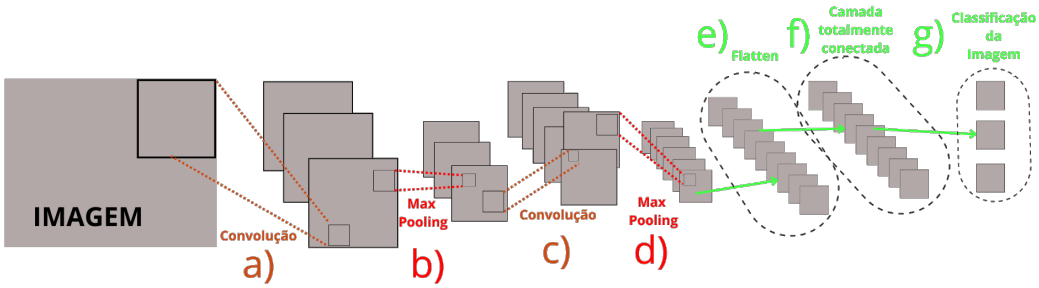


Figura 14 – Representação das camadas de uma Rede Neural Convolucional. Em a) Convolução, em b) Max Pooling, em c) Convolução, em d) Max Pooling, em e) Flatten, em f) Camada totalmente conectada e g) Classificação da imagem

Fonte: Elaborada pelo autor.

De modo geral, uma RNC é composta por camadas de Convolução, Max Pooling, Flatten e uma Camada totalmente conectada. Um exemplo de como uma RNC funciona é mostrado na Figura 14. Na camada de convolução, são extraídas as características da imagem de entrada, ou seja, são aprendidos os componentes presentes na imagem, como narizes, olhos, boca e orelhas. Para isso, essa camada age como um filtro na imagem. A camada de Max Pooling extrai os dados importantes da camada de convolução, reduzindo seu tamanho e deixando somente os dados relevantes. A camada Flatten transforma a saída produzida em um vetor para que a camada totalmente conectada possa receber os dados e realizar a classificação da imagem (ALBAWI; MOHAMMED; AL-ZAWI, 2017).

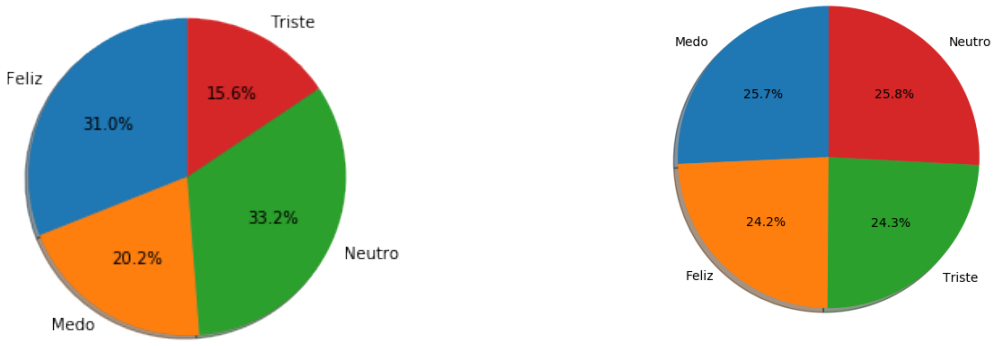
Neste trabalho, foram consideradas duas arquiteturas diferentes, as quais foram testadas com a base de dados para mensurar qual é a mais adequada aos dados fornecidos.

3.6.1.1 Treinamento das ConvNets

Como mencionado na seção anterior, para que uma rede neural profunda possa aprender uma tarefa específica, é necessário que ela processe imagens para extrair suas características principais. A fim de refinar esse processo de extração, é necessário ter um grande volume de imagens.

Para treinar a RNC, foi utilizado o conjunto de dados CAFE (The Child Affective Facial Expression) de LoBue e Thrasher (2014). Essa base de dados contém dados de 100 crianças com idades entre 2 e 8 anos, com 7 expressões faciais diferentes: medo (140 imagens), raiva (205 imagens), nojo (191 imagens), felicidade (215 imagens), neutralidade (230 imagens), tristeza (108 imagens) e surpresa (103 imagens), totalizando 1.192 imagens.

Vários testes foram realizados para otimizar o tempo de processamento e a acurácia da rede neural, levando em consideração a tarefa proposta. Por isso, foram selecionadas apenas 4 expressões faciais (classes de imagens): medo, tristeza, neutralidade e felicidade, totalizando 693 imagens. As imagens foram distribuídas em 4 classes para serem treinadas pela RNC. No entanto, as classes estavam desbalanceadas, com as classes de medo e tristeza contendo menos imagens do que as outras expressões faciais (Figura 15). Para resolver esse problema, foi aplicada a técnica de *data augmentation*, que consiste em aplicar técnicas de processamento de imagem para criar imagens sintéticas, como adição de ruído, mudanças no brilho, contraste e zoom na imagem. Cada imagem do conjunto de dados foi processada e novas imagens foram criadas. Depois, foram removidas imagens que não continham informações adicionais ou que continham erros no processamento, como algumas imagens que tiveram o zoom aplicado apenas no nariz da criança ou sofreram distorções excessivas. A quantidade de imagens da classe tristeza aumentou para 1239, a classe medo para 1303, a classe felicidade para 1203 e a classe neutralidade para 1291, totalizando 5036 imagens. Em seguida, todo o conjunto de dados foi redimensionado para o tamanho de 160 x 120 pixels. Para o treinamento, o conjunto de imagens foi separado em 3 conjuntos de dados: treinamento, validação e teste.



(a) (b)

Figura 15 – Em a) composição da distribuição das classes antes do balanceamento e em b) composição da distribuição das classes depois do balanceamento

3.6.1.2 Arquiteturas das RNCs utilizadas

Com a base de dados pronta, foram treinadas duas arquiteturas diferentes para a realização da tarefa. A primeira, descrita no trabalho de E. Pranav, propõe uma RNC (PRANAV *et al.*, 2020), que pode ser observada na Tabela 2. Para o treinamento desta rede, foi utilizado o filtro de Gabor, responsável pelo pré-processamento nas imagens visando aumentar o desempenho. No entanto, como a quantidade de imagens não era grande o suficiente, não foi necessária a sua utilização. Além disso, na arquitetura descrita, a entrada dos dados requer uma imagem de dimensões 256x256 pixels. A entrada foi alterada para dimensões de 160 x 120 pixels, pois a captura da imagem feita pelo robô NAO utiliza essas dimensões.

Tabela 2 – Arquitetura da rede RNC proposta por Pranav *et al.* (2020) com as modificações feitas.

Layer (type)	Output Shape	Param
Conv2D	(None, 118, 158, 6)	168
MaxPooling2D	(None, 59, 79, 6)	0
Conv2D	(None, 57, 77, 16)	880
MaxPooling2D	(None, 28, 38, 16)	0
Conv2D	(None, 26, 36, 120)	17400
Dropout	(None, 26, 36, 120)	0
Flatten	(None, 112320)	0
Dense	(None, 84)	9434964
Dropout	(None, 84)	0
Dense	(None, 4)	340

A segunda arquitetura apresentada na Tabela 3 é baseada no trabalho de M. M. Taghi Zadeh, que propõe uma Rede Neural Convolucional Densa (RNCD) (ZADEH; IMANI; MAJIDI, 2019). Algumas alterações foram feitas para um melhor desempenho da rede, como a inclusão da operação de Dropout. Além disso, a entrada dos dados requer uma imagem de dimensões

32x32 pixels, mas foi adaptada para suportar uma imagem de 160x120 pixels. Como pode ser observado, o autor adota uma estratégia diferente da primeira arquitetura, mantendo o mesmo número de filtros nas camadas de convolução, porém com um número maior de neurônios na camada totalmente conectada.

Tabela 3 – Arquitetura da rede RNCD proposta por Zadeh, Imani e Majidi (2019) com as modificações feitas.

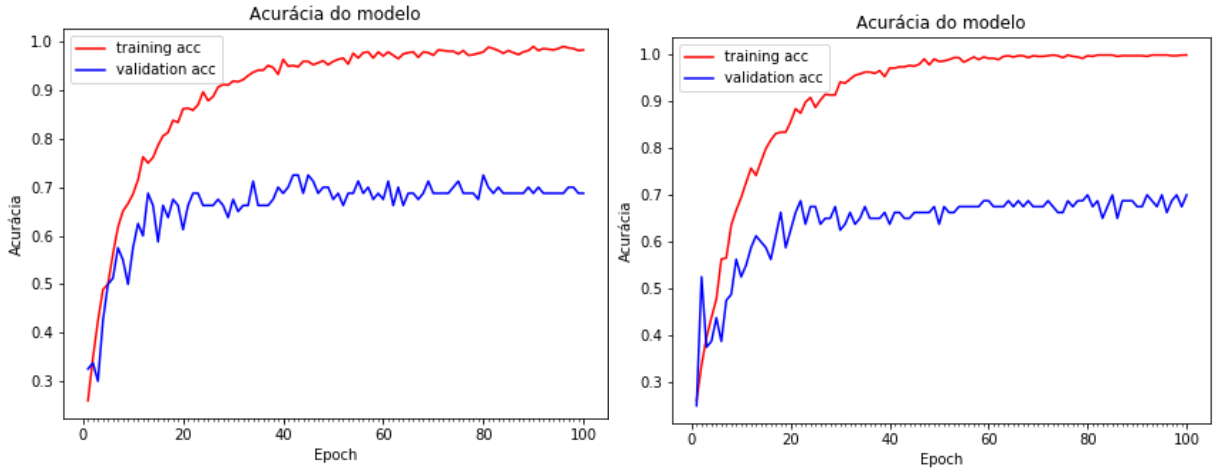
Layer (type)	Output Shape	Param
Conv2D	(None, 118, 158, 32)	896
MaxPooling2D	(None, 59, 79, 32)	0
Conv2D	(None, 57, 77, 32)	9248
MaxPooling2D	(None, 28, 38, 32)	0
Dropout	(None, 28, 28, 32)	0
Flatten	(None, 34048)	0
Dense	(None, 256)	8716544
Dense	(None, 4)	1028

3.6.1.3 Avaliação do sistema de emoção

Como mencionado em 3.6.1.2, os dados foram separados em três conjuntos: treinamento, utilizado para a extração das características da imagem; conjunto de validação, usado para ajustar a topologia e os hiperparâmetros, se necessário; e conjunto de teste, empregado para avaliar o desempenho do modelo.

Para implementar os modelos das duas redes descritas anteriormente e mencionadas aqui como RNC e RNCD, respectivamente, foram utilizadas funções disponíveis na biblioteca scikit-learn para separar esses conjuntos. Em seguida, foi utilizado o pacote OpenCV para processar as imagens de acordo com o que foi mencionado. Cada rede foi treinada por 100 épocas.

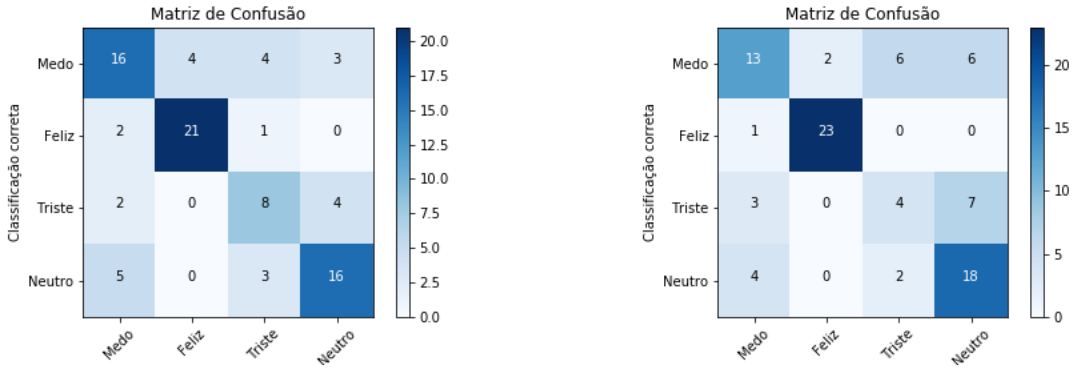
Na Figura 16, pode-se observar a acurácia da rede durante o treinamento em cada época realizada. Observa-se que ambas as redes foram treinadas com sucesso, atingindo uma acurácia de pouco mais de 80% no conjunto de validação. A rede RNC obteve um resultado de 81%, enquanto a rede RNCD obteve 80%.



(a) (b)

Figura 16 – Histórico do treinamento da RNC em a) e da RNCD em b)

Por fim, as redes RNC e RNCD foram avaliadas no conjunto de teste utilizando métricas baseadas na matriz de confusão, como mostrado nas Figuras 17(a) e 17(b), respectivamente. A RNC obteve uma acurácia geral de 81%, enquanto a RNCD obteve 80%.



(a) (b)

Figura 17 – Resultado do conjunto de teste da RNC em a) e da RNCD em b)

As métricas calculadas a partir da matriz de confusão foram a precisão, que mostra dentre os resultados classificados como positivos, a porcentagem de positivos realmente eram corretos; o *Recall* que informa a taxa de previstos como positivos dentre os que são verdadeiramente positivos e por fim, a métrica F1-Score, que combina a precisão e o *Recall*, indicando a qualidade geral do modelo, onde 1 é o resultado para um modelo perfeito.

Tabela 4 – Tabela com os resultadas da RNC e RNCD obtidos a partir da matriz de confusão

	RNC			RNCD		
Emoção	Precisão	Recall	F1-Score	Precisão	Recall	F1-Score
Medo	0,76	0,76	0,76	0,77	0,76	0,76
Feliz	0,89	0,94	0,91	0,90	0,91	0,90
Triste	0,75	0,70	0,72	0,71	0,69	0,70
Neutro	0,84	0,84	0,84	0,80	0,83	0,82
Média	0,81	0,81	0,81	0,80	0,80	0,80

A Tabela 4 mostra os valores obtidos das métricas a partir da matriz de confusão. Como pode ser observado, a arquitetura da RNC obteve uma média dos valores de Precisão, Revocação e do F1-Score superiores à arquitetura da RNCD. No caso da RNC, pode-se observar que ela obteve resultados aproximados para precisão e *recall*, com uma taxa média aceitável de 67% e 63%, respectivamente. A emoção feliz foi a que obteve o melhor resultado em ambas as arquiteturas, com uma taxa de *recall* de 94% para a RNC e 91% para a RNCD, e uma precisão de 89% e 90% para RNC e RNCD, respectivamente.

É importante ressaltar que ambas as redes foram treinadas apenas com faces de crianças, tornando a rede especialista em identificar emoções em crianças. Os trabalhos publicados levam em consideração um modelo mais genérico e, portanto, conseguem adquirir mais informações acerca das emoções.

3.7 Considerações Finais

Neste capítulo, foram apresentados todos os componentes que constituem o sistema proposto, incluindo a arquitetura desenvolvida, os softwares e componentes de hardware necessários, bem como as funcionalidades da interface GUIPsyin. É importante destacar que cada um desses elementos é essencial para o funcionamento adequado do sistema e para a realização dos experimentos que serão descritos no próximo capítulo.

No capítulo seguinte, serão apresentados os mecanismos de avaliação do sistema adotados, bem como os resultados dos experimentos realizados. Com base nesses resultados, será possível avaliar a eficácia do sistema proposto e discutir possíveis melhorias e aprimoramentos para futuras versões.

CAPÍTULO

4

MÉTODO DE AVALIAÇÃO DO SISTEMA

Neste capítulo, apresenta-se a metodologia adotada para avaliar o sistema proposto, bem como as métricas utilizadas. A avaliação do sistema ocorrerá em duas frentes: a primeira relacionada à usabilidade e facilidade de uso do sistema e a segunda relacionada a sua aplicação no ambiente forense.

4.1 Método de avaliação do sistema de software

A metodologia *Goal Question Metric* (GQM) é uma abordagem sistemática para estabelecer objetivos, formular questões para alcançar esses objetivos e definir métricas para responder às questões formuladas. Essa metodologia foi originalmente proposta por Victor Basili em 1984, como uma maneira de avaliar a qualidade de processos de software (BASILI; CALDIERA; ROMBACH, 1994).

A metodologia GQM é dividida em três etapas principais: definição de objetivos, formulação de questões e definição de métricas. Na primeira etapa, o objetivo geral do projeto ou do processo é definido, juntamente com os objetivos específicos que devem ser alcançados para atingir o objetivo geral. Na segunda etapa, as questões são formuladas para avaliar se os objetivos foram alcançados. Essas questões devem ser formuladas de forma clara e precisa, para que possam ser respondidas de maneira mensurável. Na última etapa, as métricas são definidas para responder às questões formuladas. Essas métricas devem ser mensuráveis e relevantes para as questões formuladas.

A metodologia GQM é útil em uma variedade de contextos, desde o desenvolvimento de software até a avaliação de sistemas, processos e produtos. A abordagem sistemática permite que as empresas e organizações estabeleçam objetivos claros, meçam seu progresso em relação a esses objetivos e tomem decisões informadas com base nos resultados da avaliação.

4.2 Desenvolvimento do método de avaliação

Como apresentado na seção 4.1, a metologia GQM foi utilizada para a avaliação do sistema.

4.2.1 Etapa 1 - Objetivos

Como a primeira etapa do GQM é estabelecer uma meta a ser alcançado, definiu-se como objetivo principal verificar se o sistema é capaz de mediar a interação entre uma criança e um psicólogo por meio da teleoperação de um robô. Como objetivos específicos, temos:

1. Verificar se o sistema desenvolvido é caracterizado como um sistema de teleoperação gestual.
2. Verificar se a sessão durou mais tempo, uma vez que a primeira pergunta que a criança fez para o robô ocorreu em um curto intervalo de tempo.
3. Verificar se a criança continuou interagindo com o robô NAO mesmo quando interagindo com outros brinquedos.

Os dois últimos objetivos foram levantados após a implementação do sistema no ambiente forense, no qual as interações geradas foram gravadas e posteriormente analisadas.

4.2.2 Etapa 2 - Questões

Na segunda etapa do processo de avaliação, a fim de validar o objetivo específico 1, foram elaboradas uma série de questões que constam na tabela 5. Essas questões foram criadas com o intuito de avaliar o sistema desenvolvido.

Tabela 5 – Questões utilizadas para avaliar o sistema

	Pergunta
Q1:	O sistema foi de fácil utilização?
Q2:	O sistema se mostrou confiável?
Q3:	O sistema pode ser implementado no seu dia-a-dia?
Q4:	Houve problemas na comunicação por voz?
Q5:	Houve problemas na captura da imagem?
Q6:	O robô NAO conseguiu se expressar como você queria?
Q7:	O sistema apresentou algum risco para a criança?

Tais questões foram utilizadas na etapa de validação que foi fundamental para garantir a eficácia e a confiabilidade do sistema, uma vez que um sistema de teleoperação gestual inadequado pode gerar falhas na comunicação entre a criança e o psicólogo e comprometer o processo terapêutico.

Então para avaliar os objetivos específicos 2 e 3, foram levantadas as seguintes questões da Tabela 6:

Tabela 6 – Questões utilizadas para avaliar a interação do robô com a criança

	Pergunta
Q1:	Quantas perguntas a criança iniciou durante a interação com o robô NAO?
Q2:	Quanto tempo a criança levou para fazer a primeira pergunta durante a interação com o robô NAO?
Q3:	Qual foi a duração total da sessão, ou seja, o tempo em que a criança interagiu com o robô NAO ou ficou próxima a ele?
Q4:	Qual foi o tempo em que a criança interagiu com o robô NAO utilizando outros meios, como brinquedos ou lousa?

Com essas questões é possível avaliar se o sistema é capaz de mediar a interação entre a criança e o psicólogo de forma eficiente, bem como verificar se a interação com o robô NAO teve um impacto positivo na duração e qualidade da sessão.

A quantidade de perguntas iniciadas pela criança durante a interação com o robô NAO e o tempo que ela levou para fazer a primeira pergunta podem indicar a qualidade da comunicação estabelecida entre a criança e o robô.

A avaliação da duração total da sessão e do tempo em que a criança interagiu com o robô NAO utilizando outros meios pode indicar a efetividade do sistema em manter a atenção e o interesse da criança na interação, bem como a sua capacidade de atuar como um mediador da interação entre a criança e o psicólogo.

Por tanto a análise dessas serviram para avaliar se o sistema atinge os objetivos específicos e se é capaz de mediar de forma eficiente a interação entre a criança e o psicólogo.

4.2.3 Etapa 3 - Métricas

Na terceira etapa do processo com as questões levantadas anteriormente as seguinte métricas foram utilizadas para avaliar as questões.

Com o objetivo de avaliar se o software desenvolvido estava pronto para ser utilizado no ambiente proposto. Foram adotadas algumas características descritas na norma ISO/IEC 25010. Essa norma estabelece um modelo de qualidade de software que define as seguintes características a serem avaliadas para garantir a qualidade do produto:

- Funcionalidade, que é a capacidade do software satisfazer o usuário em seu desempenho e funcionalidades.
- Confiabilidade, que é a capacidade do software de manter um desempenho esperado sob condições específicas de uso.

- Usabilidade, que é a facilidade de uso do software e à satisfação do usuário com sua interface e usabilidade.
- Eficiência, que é a capacidade do software de realizar suas funções com desempenho adequado em relação aos recursos utilizados.

As características de funcionalidade, confiabilidade, usabilidade e eficiência foram avaliadas por meio das questões elaboradas na Tabela 5. A escala Likert foi utilizada para estabelecer uma métrica de avaliação, na qual o valor 1 corresponde a "discordo completamente"ou "nada satisfatório", enquanto o valor 5 corresponde a "concordo totalmente"ou "totalmente satisfatório". As respostas foram avaliadas com base no *Mean Opinion Score* (MOS). A Tabela 7 apresenta a relação entre as questões e as características estabelecidas, indicadas por um "X".

Tabela 7 – Relação entre as perguntas da Tabela 5 e as características da ISO/IEC 25010 estabelecidas

Questões	Funcionalidade	Confiabilidade	Usabilidade	Eficiência
Q1	X		X	
Q2	X	X		
Q3		X		
Q4	X	X		X
Q5	X	X		X
Q6	X	X	X	X
Q7		X		

Por fim para avaliar as questões elaboradas na Tabela 6, foram utilizadas as seguintes métricas:

- Q1 - Foram contabilizadas as perguntas que a criança iniciou com o robô.
- Q2 - Foi medido em minutos o tempo que a criança levou para fazer a primeira pergunta para o robô.
- Q3 - Foi medido em minutos o tempo total que a criança participou da sessão.
- Q4 - Foi medido em minutos o tempo que a criança interagiu com o robô e com outros recursos.

4.3 Considerações Finais

Neste capítulo, foi apresentada a metodologia de avaliação do sistema, composta por três etapas: formulação dos objetivos, elaboração de questões para verificação do cumprimento dos objetivos e métricas para avaliação das perguntas realizadas.

Dessa forma, no próximo capítulo serão apresentados os experimentos realizados, incluindo uma discussão dos resultados obtidos.

CAPÍTULO

5

RESULTADOS E EXPERIMENTOS REALIZADOS

Neste capítulo, serão apresentados os resultados das avaliações realizadas no sistema, que foram conduzidas por meio das questões elaboradas pelo método GQM apresentado no capítulo anterior. Além disso, serão descritos os experimentos que demonstraram a evolução do sistema em quatro etapas. O primeiro experimento envolveu a participação de 10 crianças da cidade de Botucatu-SP, com a presença de uma psicóloga. O segundo experimento foi realizado no laboratório LAR-ICMC-USP com o objetivo de testar o funcionamento do sistema integrado, que envolve a coordenação dos componentes de hardware e software e a interface GUIPsyin. O terceiro experimento teve como propósito avaliar a funcionalidade, usabilidade, confiabilidade e eficiência do sistema. Por fim, serão apresentados os resultados obtidos durante a utilização do sistema no contexto da psicologia forense.

5.1 Testes Iniciais

A primeira versão do sistema integrado foi testada em Botucatu, no primeiro semestre de 2021, para verificar se os gestos e o sistema de operação poderiam ser funcionais. Na primeira versão não foi possível testar o sistema completo pois não foi utilizado um sistema de transmissão por voz, mas sim um sistema de texto para voz. Nesse sistema o psicólogo digitava as frases que o robô iria falar e o robô as repetia. Alguns problemas foram identificados com isso.

- O Psicólogo não conseguia escrever rápido o suficiente para manter uma conversa espontânea.
- Algumas palavras eram faladas erradas pelo robô NAO, pois com a digitação rápida o usuário cometia erros de ortografia e consequentemente o NAO as pronunciava de maneira errônea.

- A voz do robô NAO não tinha uma velocidade boa o suficiente para a compreensão da criança.

Além disso, foi testada a aceitação do vídeo de apresentação do NAO. Esse vídeo foi desenvolvido apenas como um piloto para melhorias futuras[1]. Nenhuma das crianças se recusou a conhecer o NAO.

Os pontos positivos constatados foram: o sistema permaneceu funcional durante sessões de 1 hora conseguindo manter a atenção da criança. O robô não teve nenhuma falha na execução dos movimentos.

A segunda versão do sistema foi testada no ICMC, com voz sendo transmitida de uma sala para a outra. Nesse teste foi feito um ajuste da sensibilidade do microfone, para que problemas como microfonia não surgissem durante sua aplicação. Na Figura 18, é mostrado a disposição da integração dos softwares.

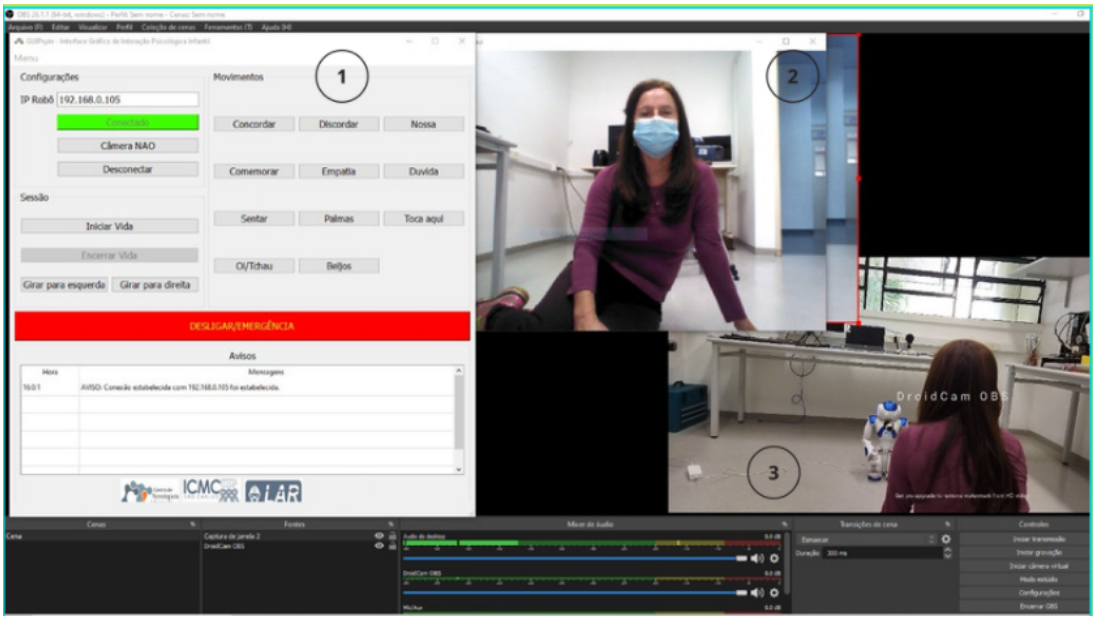


Figura 18 – Sistema de interação psicológica infantil: GUIPsyin em (1), em (2) captura do vídeo da câmera do NAO e em (3) a captura do video através do celular

Fonte: Elaborada pelo autor.

5.2 Caracterização de um sistema de teleoperação gestual

O sistema proposto pode ser considerado teleoperado com a capacidade de transmissão gestual com base nos seguintes princípios sugeridos por Stahl, Anastasiou e Latour (2018).

São três princípios citados para o desenvolvimento de um sistema teleoperado capaz de transmitir gestos e linguagem corporal. O primeiro princípio recomenda que o sistema

[1] Video de apresentação<https://www.youtube.com/watch?v=0r1eNbWP8oI>

deve prover um meio de comunicação intencional por voz, ou seja, através de ligações de voz, mensagens instantâneas, e-mails ou chamadas de vídeo. O computador que está na sala de Entrevista por meio do *Ssuite FaceTime Video Phone* faz a captação do áudio da sala através do microfone conectado e ao mesmo tempo transmite a voz do mágico (psicólogo) modificada pelo *MorphVox Pro* para a caixa *bluetooth* que está localizada no peito do robô NAO, além disso o microfone que está na mesma sala também capta o som da sala de Entrevista e envia para o operador na sala de Controle, além disso o robô NAO transmite o vídeo para o computador da sala de Controle, além do celular que capta a imagem do ambiente, atendendo o primeiro princípio.

O segundo princípio diz que o sistema deve prover meios de interação gestual apropriada para a comunicação. Por meio da interface GUIPsyin o robô NAO é capaz de gesticular de acordo com a vontade do mágico (psicólogo) pelos movimentos disponíveis. Entre os movimentos possíveis encontram-se: o movimento: "Toca aqui", que convida a criança a cumprimentar o NAO, o "Oi/Tchau", que é usado para dar as saudações. Por outro lado, o "Concordar"e o "Discordar"passam a ideia de aprovação e desaprovação, enquanto a "Empatia", que faz o NAO abaixar a cabeça com um movimento leve, passando a ideia de tristeza, o "Comemorar"que faz uma dança alegre. Além disso, o "Nossa"que faz o NAO levantar a mão em direção a testa, passando a ideia de surpresa, a "Dúvida"que leva o NAO na postura de "Superman"gerando uma dúvida, a "Palmas"para reforçar um comportamento positivo e por fim, o "Beijos"que faz o NAO enviar beijos para a criança criando uma despedida agradável.

O terceiro princípio recomenda que o sistema deve prover um meio não intencional de comunicação individual, ou seja, movimentos que são capazes de identificar um humano no ambiente e ações atreladas à conversação. Em relação a isto, a interface GUIPsyin fornece essas opções através da sessão, quando o ciclo "Iniciar Vida"é ativado e o robô NAO entra em estado de alerta. Este estado faz o NAO procurar um rosto humano por meio do sistema de detecção de face que ele possui, e quando encontrado ele faz o contato visual, esse contato faz o NAO focar no humano localizado.

Portanto, pode-se dizer que o sistema tem os requisitos necessários para ser caracterizado como um sistema de teleoperação gestual, pois atende aos três princípios de forma completa.

5.3 Testes de aceitação do sistema

Conforme mencionado na seção 4.2, a avaliação do sistema considerou o método GQM em conjunto com a norma ISO/IEC 25010, permitindo avaliar a funcionalidade, confiabilidade, usabilidade e eficiência do sistema.

Essa validação teve como objetivo verificar se o sistema estava suficientemente maduro para ser utilizado em um ambiente forense. Para isso, foram convidados seis psicólogos atuantes em uma clínica para passarem por três fases:

1. Na primeira fase, foi explicado o propósito do projeto, os componentes necessários e fornecido um treinamento de cinco minutos sobre como operar a interface GUIPsyin.

2. Na segunda fase, foi realizada uma simulação de como ocorreria a implementação do sistema durante uma sessão de terapia.

3. Na terceira fase, foi aplicado o questionário, apresentado na Tabela 5.

Os resultados foram obtidos a partir do questionário e medidos usando a escala Likert. Esses resultados ajudaram a avaliar a percepção dos psicólogos em relação ao sistema, auxiliando na determinação da maturidade do sistema para uso no contexto forense.

Tabela 8 – Resultados obtidos a partir das questões realizadas

	S1	S2	S3	S4	S5	S6	Média
Q1	5,0	5,0	5,0	5,0	5,0	3,0	4,67
Q2	5,0	5,0	5,0	5,0	5,0	5,0	5,00
Q3	5,0	5,0	5,0	5,0	5,0	5,0	5,00
Q4	5,0	5,0	3,0	4,0	4,0	4,0	4,17
Q5	4,0	5,0	5,0	4,0	5,0	5,0	4,67
Q6	5,0	5,0	5,0	5,0	4,0	5,0	4,83
Q7	5,0	5,0	5,0	5,0	5,0	5,0	5,00

Na Tabela 8, são apresentados os resultados obtidos a partir da implementação do questionário. É possível observar que todos os participantes indicaram que o sistema poderia ser implementado no dia a dia, conforme a Q3. Além disso, é importante ressaltar que as questões Q2 e Q7 receberam nota máxima de todos os participantes, indicando que o sistema foi considerado "Muito confiável"e não apresentou "Nenhum risco"para as crianças, respectivamente.

Como mencionado no capítulo anterior, para avaliar se o sistema estava maduro o suficiente para ser implantado no ambiente forense, foram avaliados a funcionalidade, confiabilidade, usabilidade e eficiência do sistema. A partir das médias calculadas de cada questão, foi gerada a Tabela 9.

Tabela 9 – Resultados obtidos para avaliar a funcionalidade, confiabilidade, usabilidade e eficiência do sistema

	Funcionalidade	Confiabilidade	Usabilidade	Eficiência
Q1	4,67		4,67	
Q2	5,00	5,00		
Q3		5,00		
Q4	4,17	4,17		4,17
Q5	4,67	4,67		4,67
Q6	4,83	4,83	4,83	4,83
Q7		5,00		
Média	4,67	4,78	4,75	4,56

Como pode-se observar, o sistema obteve uma média acima de 4,5 em todas as características, com destaque em sua usabilidade e confiabilidade. Demonstrando sua capacidade de ser utilizado no ambiente forense.

A seguir, são apresentados os resultados obtidos na fase de aplicação do sistema, na qual o sistema foi implementado no contexto da psicologia forense.

5.4 Aplicação do sistema na Psicologia Forense

No contexto da Psicologia Forense, o experimento foi aprovado de acordo com o parecer do Comitê de Ética no. 5.233.368 e pelo Processo n°: 2021/88468 (CPA) – SPI 1 para ser aplicado na 1ª Vara Criminal da Infância e da Juventude do Fórum Comarca de Botucatu-SP. O sistema proposto foi aplicado na interação com 11 crianças com idades entre 4 a 8 anos. Uma psicóloga ficou responsável de utilizar o sistema durante a interação. Cada interação foi gravada para ser análisada posteriormente. Como proposto, o sistema foi montado em dois ambientes: em uma sala de Controle e em uma sala de Entrevista como mostram as Figuras 19 e 20, respectivamente.

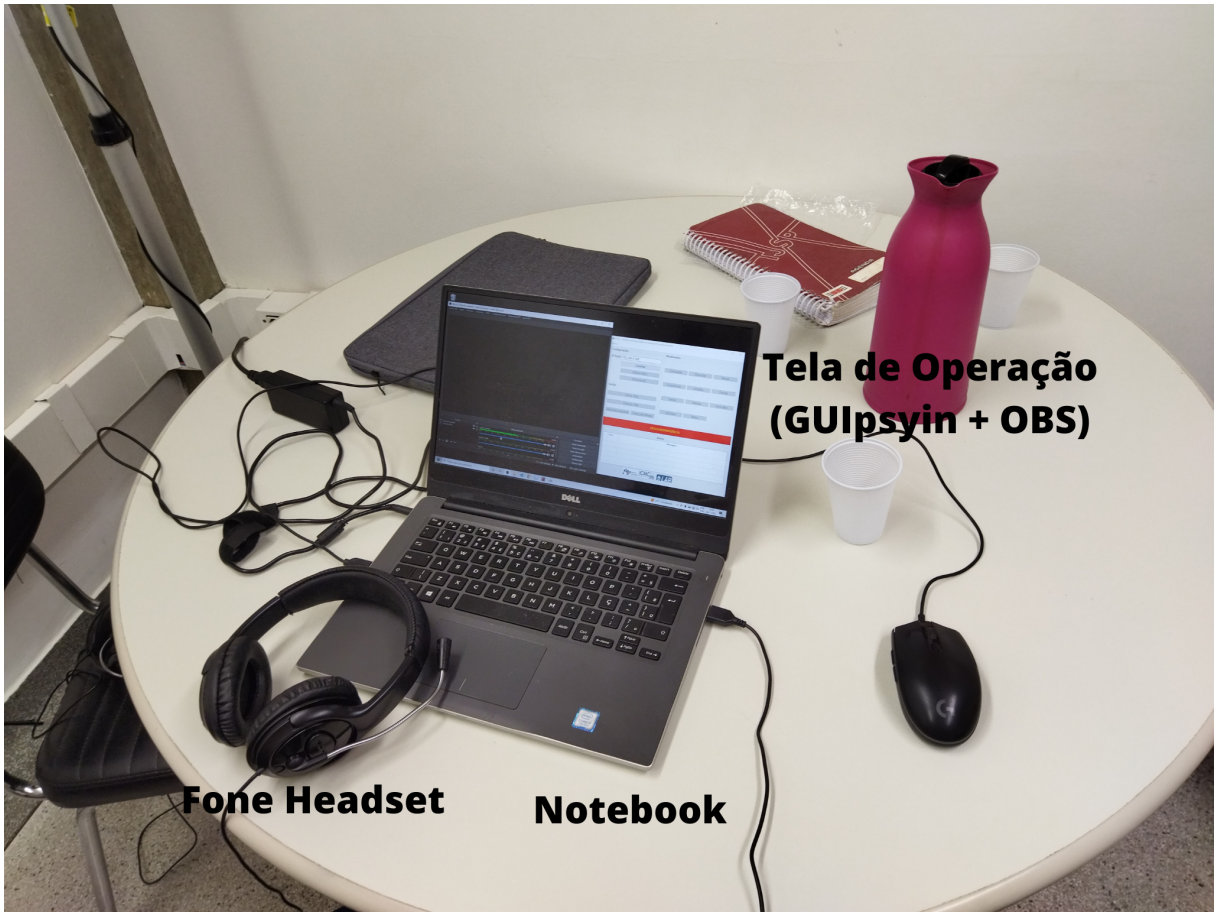
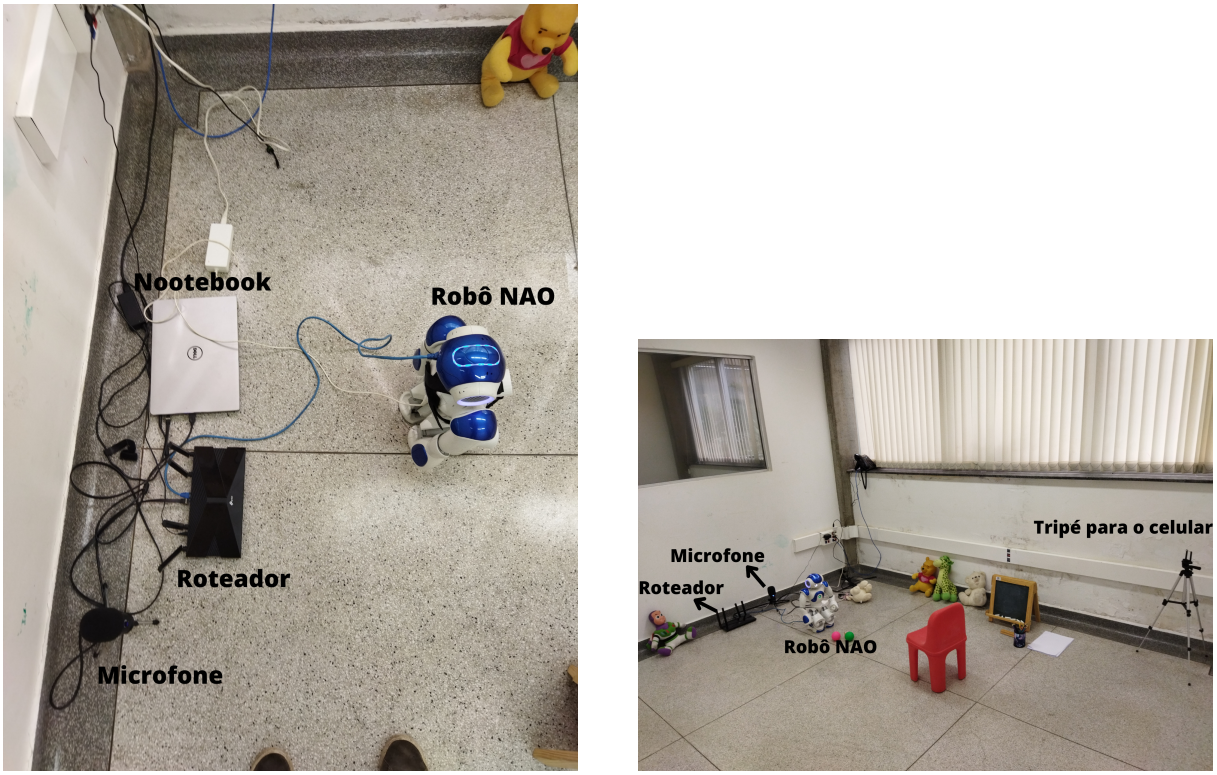


Figura 19 – Disposição dos equipamentos na sala de Controle.

Fonte: Elaborada pelo autor.



(a) (b)

Figura 20 – Em a) temos a Disposição dos equipamentos ligados ao robô NAO na sala de entrevista e em b) a visão geral da sala.

Fonte: Elaborada pelo autor.

O sistema foi avalidado como um todo, observando-se como o sistema de áudio e vídeo se comportavam e as dificuldades da utilização da interface GUIPsyin pelo psicólogo.

Foram realizadas 11 sessões ao todo. O sistema de áudio funcionou sem nenhum tipo de erro em 8 delas. O sistema de vídeo do robô NAO funcionou em 8 sem nenhum erro. Em uma única sessão ocorreu apenas problemas no áudio e em outra apenas no vídeo da captura do robô NAO. Problemas no sistema de áudio e vídeo do NAO, conjuntamente, ocorreram em apenas duas sessões. Os problemas que ocorreram no sistema de vídeo e áudio foram, respectivamente, congelamento da imagem que o NAO transmite, grande quantidade de ruídos sendo captados pelo microfone, impossibilitando o psicólogo de escutar o que as crianças falavam. Tal travamento no sistema de vídeo foi causado pelo grande número de imagens processadas, pois as imagens transmitidas pelo robô eram enviadas para o computador da sala de Controle, no qual eram processadas e enviadas para o sistema de detecção de emoção. A solução foi substituir a conexão via wifi por uma conexão de rede a cabo, na qual o robô seria ligado diretamente no roteador. No entanto, tais congelamentos não foram prejudiciais para a interação do robô NAO com as crianças. Os problemas de ruídos foram resolvidos apenas ajustando-se a sensibilidade do microfone utilizado.

Outro aspecto analisado foi o desempenho do detector de emoções, que foi analisado de

forma subjetiva. Como mencionado na seção 3.6.1.2, a RNC foi treinada para detectar apenas 4 emoções: feliz, triste, medo e neutro com uma acurácia de 80.0%. No entanto, durante a sessão foi observado que a rede classificava cada criança apenas com as emoções triste e medo, mesmo em momentos nos quais ela se mostrava feliz e sorridente. Alguns fatores que contribuíram para essa classificação errada podem ter sido a luz do ambiente que variava de acordo com o horário da sessão e a qualidade da imagem obtida pelo robô NAO e processada para a rede, pois a camera do robô não é das mais atuais. Com isso, o psicólogo tirou pouco proveito do detector de emoções, sendo considerado como um dos tópicos de pesquisa que precisa ser aperfeiçoado.

Contudo, a interface GUIPsyin funcionou bem em todas as sessões, enviando os comandos de movimentos para o robô sem nenhum erro. Apenas em uma interação foi perdida a conexão entre a interface e o NAO, pois enquanto o robô movimentava a sua cabeça, o cabo RJ 45 acabou se desconectando dele. Este fato impossibilitou a comunicação entre o robô e o computador da sala de Controle.

Com base nas sessões de interação realizadas, os seguintes dados foram coletados durante a análise dos vídeos:

- (A) Total de perguntas que a criança teve a iniciativa de fazer para o NAO durante a sessão.
- (B) O tempo que a criança levou para fazer a primeira pergunta.
- (C) A duração da sessão, ou seja, o tempo que a criança ficou interagindo com o robô ou ficou próximo a ele em minutos.
- (D) O tempo que a criança interagiu com o NAO utilizando algum outro meio, como por exemplo, uma lousa para desenhar ou um brinquedo em minutos.

Tabela 10 – Resultados obtidos da análise dos vídeos obtidos nas interações

Criança	A	B	C	D
1	4	4:38	35:20	00:00
2	32	3:42	42:02	05:30
3	8	4:19	45:40	31:51
4	9	4:50	47:12	08:54
5	34	1:40	59:45	06:17
6	7	2:20	29:37	00:00
7	8	5:50	42:39	22:49
8	13	4:00	39:37	07:19
9	38	0:19	40:03	00:46
10	2	0:22	45:05	05:55
11	41	0:55	43:30	14:36

De acordo com os dados da Tabela 10 pode-se responder as perguntas elaboras na Tabela 6. Além disso foi realizada uma análise de correlção entre as variáveis. Para isso os valores

das colunas B, C e D da Tabela 10 foram transformados para segundos. Após isso utilizando o método de transformação MinMax os dados da tabela foram normalizados e análisados através da matriz de correlação de Pearson. Com isso pode-se observar a correlação entre total de perguntas e o tempo da primeira pergunta, a correlação entre o tempo de interação utilizando algum brinquedo e a duração da sessão e por fim a correlação entre o total de perguntas e a duração da sessão.

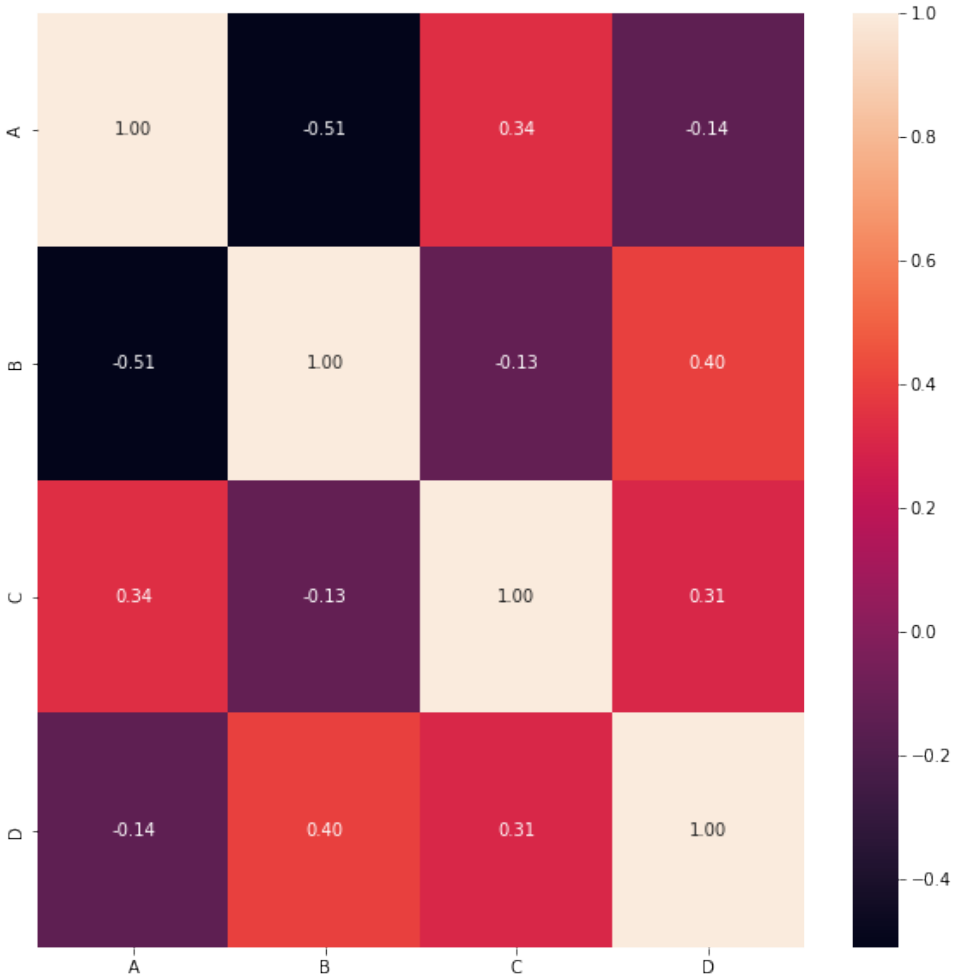


Figura 21 – Correlação de Pearson entre as variáveis A, B, C e D

Fonte: Elaborada pelo autor.

A Figura 21, mostra a correlação entre as váriaveis: total de perguntas feitas para o NAO (A), primeira pergunta (B), Duração da sessão (C) e o tempo que a criança interagiu com o NAO utilizando algum outro meio (D).

Considerando os valores apresentados, é possível observar que a variável A apresenta uma correlação negativa moderada com a variável B (-0.51), indicando que há uma relação inversa entre o tempo que a criança leva para fazer a primeira pergunta e a quantidade de perguntas feitas durante a sessão. A correlação entre A e C (0.34) é moderadamente positiva, sugerindo que há uma associação entre a duração da sessão e o número de perguntas feitas pelas

crianças ao robô. Por outro lado, a correlação entre as variáveis A e D (-0.14) é fraca, o que pode indicar uma relação menos significativa entre a interação da criança com outros meios e o número de perguntas feitas ao robô.

Ao analisar a correlação entre as variáveis B e C (-0.13), é observada uma correlação negativa fraca, o que pode sugerir uma relação inversa entre o tempo para a primeira pergunta e a duração da sessão. Além disso, a variável B exibe uma correlação moderada com a variável D (0.40), indicando uma possível relação entre o tempo para a primeira pergunta e a interação da criança com outros meios, bem como com o robô.

Por fim, ao avaliar a correlação entre as variáveis C e D (0.31), é identificada uma correlação moderada, indicando uma possível associação entre a duração da sessão e o tempo de interação da criança com outros meios.

5.5 Considerações Finais

Neste capítulo, foram descritos os experimentos realizados, com destaque para o experimento que foi realizado no Fórum de Botucatu, no contexto da psicologia forense. Foram usadas algumas medidas para avaliar a interação da criança com o sistema e também do psicólogo com o sistema. Os resultados foram promissores. No entanto, observou-se que o módulo de detecção de emoções não contribuiu para a análise, uma vez que não classificou bem as imagens das crianças obtidas pelo robô. No próximo capítulo, será apresentada a conclusão do trabalho e alguns trabalhos futuros serão apontados.

CAPÍTULO

6

CONCLUSÃO

Neste trabalho, foi proposto e desenvolvido um sistema integrado para auxiliar psicólogos em sessões de terapia. O sistema é um sistema de teleoperação que envolve, por um lado, um psicólogo e, por outro, um robô social (NAO) interagindo com uma criança. O psicólogo interage com a criança por meio do robô, enviando comandos de voz para o robô. A criança interage com o robô e o psicólogo recebe informações dessa interação, tanto das imagens quanto do áudio correspondente. O sistema foi proposto para ser utilizado em dois contextos: na Psicologia Forense e na Psicologia Clínica. Foi desenvolvida uma interface amigável por meio da qual um profissional da área pode realizar a interação sem ter que entender de computação e dos recursos computacionais envolvidos.

Recursos de hardware e software tiveram que ser utilizados para atingir os objetivos do sistema, cada um com um papel claramente definido no ambiente proposto. O sistema foi projetado para operar em duas salas: a Sala de Entrevista e a Sala de Controle. Na sala de Entrevista, são encontrados o robô NAO com uma caixa Bluetooth acoplada em seu peito, um notebook, um microfone, um roteador de Wi-Fi e a criança com a qual o NAO irá interagir. Na Sala de Controle, deve estar presente um Psicólogo que irá operar o robô através da GUIPsyin, com o auxílio de outros softwares, como o MorphVox para alteração da voz do Psicólogo, o OBS Studio para a gravação da sessão e o SStudio Face Phone para permitir a transmissão da fala do Psicólogo para a Sala de Entrevista. Ao final da sessão, o psicólogo poderá fazer uma avaliação da interação ocorrida.

A interface desenvolvida, denominada GUIPsyin, faz parte central do sistema, pois através dela, o usuário (Psicólogo) na Sala de Controle consegue operar o robô NAO na Sala de Entrevista. Ela é composta por 5 módulos: Configurações, onde é feita a conexão com o Robô NAO, Sessão, onde é iniciado o procedimento de engajamento do robô, Movimentos, onde os movimentos pré-programados ficam disponíveis para o operador, o botão de emergência para que faça o robô parar de operar em determinada circunstância e, por fim, a área de avisos que

informa o operador sobre o sucesso ou falha dos comandos enviados.

O sistema foi avaliado utilizando o método GQM (BASILI; CALDIERA; ROMBACH, 1994), no qual foram necessárias duas etapas para validar o sistema. Na primeira etapa, o sistema passou por uma avaliação na qual 06 psicólogos avaliaram os seguintes fatores: a funcionalidade, usabilidade, confiabilidade e eficiência do sistema, onde todos foram avaliados com uma nota superior a 4.

Na segunda etapa, o sistema foi testado no contexto da Psicologia Forense. Infelizmente, o sistema não pode ser utilizado no contexto da Psicologia Clínica e dessa forma essa pesquisa fará parte dos trabalhos futuros.

Diante disso, é importante observar que a utilização do sistema apresenta algumas limitações que devem ser consideradas. Uma delas diz respeito a quantidade de componentes de hardware e software necessária para o funcionamento do sistema, o que pode representar um desafio em termos de custo computacional e complexidade. Além disso, o ambiente no qual o sistema pode ser utilizado deve ser adequadamente isolado acusticamente, a fim de evitar que a criança ouça o operador durante a interação com o robô. Outra limitação é o custo associado à aquisição e manutenção do robô NAO, que pode ser elevado para algumas instituições. Essas questões devem ser cuidadosamente consideradas pelos profissionais que desejam utilizar o sistema em sua prática clínica ou forense.

No entanto, o trabalho realizado pode ser considerado uma inovação no Brasil, destacando o sucesso na interação entre um robô humanóide e uma criança. Durante as interações realizadas, as criança demonstraram interesse em verbalizar e interagir devido à presença do robô, o que evidencia o potencial dessa tecnologia como ferramenta de estímulo ao desenvolvimento infantil.

Assim as seguintes publicações foram geradas durante o período dessa Dissertação de Mestrado:

- Artigo Godoi *et al.* (2020) apresentado e publicado nos Anais do Simpósio de Robótica da América Latina 2020 (LARS), Simpósio Brasileiro de Robótica 2020 (SBR) e no Workshop de Robótica na Educação 2020 (WRE).
- Artigo Garcia *et al.* (2021) apresentado no 10º Congresso de Avaliação Psicológica.
- Resumo do Artigo Godoi *et al.* (2022) apresentado e publicado nos Anais do I Simpósio de Engenharia Biomédica do Vale do Paraíba, 2022, São José dos Campos. Onde recebeu o prêmio de trabalho destaque.

Como trabalhos futuros, pretende-se ampliar os testes realizados com as crianças no ambiente forense e iniciar os testes com crianças com Mutismo Seletivo ou Fobia Social no Instituto de Psiquiatria do Hospital das Clínicas da Faculdade de Medicina da USP. Outra possibilidade de trabalho a ser realizado é o aperfeiçoamento do sistema multimídia, implementando novos

métodos visando diminuir os recursos de hardware e software utilizados. É necessário também aprimorar o modelo da rede neural utilizada para a detecção de emoções das crianças, para que ela possa reconhecer satisfatoriamente as emoções básicas em diferentes ambientes.

REFERÊNCIAS

ALBAWI, S.; MOHAMMED, T. A.; AL-ZAWI, S. Understanding of a convolutional neural network. In: **2017 International Conference on Engineering and Technology (ICET)**. [s.n.], 2017. p. 1–6. Disponível em: <https://doi.org/10.1109/ICEngTechnol.2017.8308186>. Citado na página 46.

ALQUISIRIS-QUECHA, O.; MALDONADO-REYES, A. E.; MORALES, E. F.; SUCAR, L. E. Teleoperation and control of a humanoid robot nao through body gestures. In: **2018 Seventeenth Mexican International Conference on Artificial Intelligence (MICAI)**. [s.n.], 2018. p. 88–94. Disponível em: <https://doi.org/10.1109/MICAI46078.2018.00022>. Citado na página 26.

AZIZ, A.; MOGANAN, F.; ISMAIL, A.; LOKMAN, A. Autistic children's kansei responses towards humanoid-robot as teaching mediator. **Procedia Computer Science**, v. 76, p. 488–493, 12 2015. Disponível em: <https://doi.org/10.1016/j.procs.2015.12.322>. Citado na página 27.

BASILI, V. R.; CALDIERA, G.; ROMBACH, H. D. The goal question metric approach. 1994. Disponível em: <http://www.cs.umd.edu/~mvz/handouts/gqm.pdf>. Citado nas páginas 51 e 66.

CLABAUGH, C.; MATARIĆ, M. Escaping Oz: Autonomy in Socially Assistive Robotics. **Annual Review of Control, Robotics, and Autonomous Systems**, v. 2, n. 1, p. 33–61, 2019. ISSN 2573-5144. Disponível em: <https://doi.org/10.1146/annurev-control-060117-104911>. Citado nas páginas 21 e 22.

DIO, C. D.; MANZI, F.; PERETTI, G.; CANGELOSI, A.; HARRIS, P. L.; MASSARO, D.; MARCHETTI, A. Shall i trust you? from child-robot interaction to trusting relationships. **FRONTIERS IN PSYCHOLOGY**, FRONTIERS MEDIA SA, AVENUE DU TRIBUNAL FEDERAL 34, LAUSANNE, CH-1015, SWITZERLAND, v. 11, APR 3 2020. ISSN 1664-1078. Disponível em: <https://doi.org/10.3389/fpsyg.2020.00469>. Citado na página 30.

FERREIRA, L. M. S.; CHATELARD, D. S. O sujeito e a musicalidade da fala. **Analytica: Revista de Psicanálise**, scielopepsic, v. 8, p. 1 – 10, 06 2019. ISSN 2316-5197. Disponível em: <http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S2316-51972019000100004&nrm=iso>. Citado na página 36.

GARCIA, M. A. T.; ROMERO, R. A.; GODOI, D.; ; AZEVEDO, H.; RAMOS, J.; FILHO, G. B. a. Robô humanoide em avaliações forenses: suporte no atendimento psicológico infantil. p. 1, 2021. Disponível em: <https://www.congresso2021.ibapnet.org.br/site/anais#M>. Citado na página 66.

GODOI, D.; ROMERO, R. A.; AZEVEDO, H.; RAMOS, J.; FILHO, G. B.; GARCIA, M. A. T. Proteger: A social robotics system to support child psychological evaluation. In: **2020 Latin American Robotics Symposium (LARS), 2020 Brazilian Symposium on Robotics (SBR) and 2020 Workshop on Robotics in Education (WRE)**. [s.n.], 2020. p. 1–6. Disponível em: <https://doi.org/10.1109/LARS/SBR/WRE51543.2020.9306975>. Citado na página 66.

GODOI, D.; ROMERO, R. A.; AZEVEDO, H.; RAMOS, J.; GODOI, B. Guipsyin: Interface gráfica para interação psicológica infantil. 2022. Disponível em: <https://www.even3.com.br/sebvap/>. Citado na página 66.

GULLIVER, A.; GRIFFITHS, K. M.; CHRISTENSEN, H. Perceived barriers and facilitators to mental health help-seeking in young people: A systematic review. **BMC Psychiatry**, v. 10, 2010. ISSN 1471244X. Disponível em: <https://doi.org/10.1186/1471-244X-10-113>. Citado na página 21.

HUIJNEN, C. A. G. J.; LEXIS, M. A. S.; JANSENS, R.; WITTE, L. P. de. How to Implement Robots in Interventions for Children with Autism? A Co-creation Study Involving People with Autism, Parents and Professionals. **Journal of Autism and Developmental Disorders**, v. 47, n. 10, p. 3079–3096, oct 2017. ISSN 0162-3257. Disponível em: <https://doi.org/10.1007/s10803-017-3235-9>. Citado na página 27.

LOBUE, V.; THRASHER, C. The child affective facial expression (cafe) set: Validity and reliability from untrained adults. **Frontiers in Psychology**, Frontiers Media S.A., v. 5, 2014. ISSN 16641078. Disponível em: <https://doi.org/10.3389/FPSYG.2014.01532>. Citado na página 46.

LOGAN, D. E.; BREAZEAL, C.; GOODWIN, M. S.; JEONG, S.; O'CONNELL, B.; SMITH-FREEDMAN, D.; HEATHERS, J.; WEINSTOCK, P. Social Robots for Hospitalized Children. **Pediatrics**, v. 144, n. 1, 07 2019. ISSN 0031-4005. E20181511. Disponível em: <https://doi.org/10.1542/peds.2018-1511>. Citado na página 30.

M., L. N.; DAJLES, D.; SILES, F. Teleoperation of a humanoid robot using an optical motion capture system. In: **2018 IEEE International Work Conference on Bioinspired Intelligence (IWOBI)**. [s.n.], 2018. p. 1–8. Disponível em: <https://doi.org/10.1109/IWOBI.2018.8464136>. Citado na página 26.

MARGE, M.; BONIAL, C.; BYRNE, B.; CASSIDY, T.; EVANS, A. W.; HILL, S. G.; VOSS, C. **Applying the Wizard-of-Oz Technique to Multimodal Human-Robot Dialogue**. 2017. Disponível em: <https://doi.org/10.48550/arXiv.1703.03714>. Citado na página 27.

MARTIN, D. U.; MACINTYRE, M. I.; PERRY, C.; CLIFT, G.; PEDELL, S.; KAUFMAN, J. Young children's indiscriminate helping behavior toward a humanoid robot. **Frontiers in Psychology**, v. 11, 2020. ISSN 1664-1078. Disponível em: <https://www.frontiersin.org/articles/10.3389/fpsyg.2020.00239>. Citado na página 28.

MOHAN, M.; KUCHENBECKER, K. J. A design tool for therapeutic social-physical human-robot interactions. In: **2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)**. [s.n.], 2019. p. 727–729. Disponível em: <https://doi.org/10.1109/HRI.2019.8673202>. Citado na página 26.

MURPHY, R. R.; NOMURA, T.; BILLARD, A.; BURKE, J. L. Human–robot interaction. **IEEE Robotics Automation Magazine**, v. 17, n. 2, p. 85–89, 2010. Disponível em: <https://doi.org/10.1109/MRA.2010.936953>. Citado na página 21.

ORGANIZATION, W. H. **WHO | Mental health: a state of well-being.** 2019. Disponível em: <https://www.who.int/features/factfiles/mental_health/en/.> Citado na página 21.

O'SHEA, K.; NASH, R. An introduction to convolutional neural networks. **CoRR**, abs/1511.08458, 2015. Disponível em: <https://doi.org/10.48550/arXiv.1511.08458>. Citado na página 45.

PALESTRA, G.; CAZZATO, D.; ADAMO, F.; BORTONE, I.; DISTANTE, C. Assistive robot, rgb-d sensor and graphical user interface to encourage communication skills in asd population. **Journal of Medical Robotics Research**, v. 02, n. 02, p. 1740002, 2017. Disponível em: <https://doi.org/10.1142/S2424905X17400025>. Citado na página 28.

PINO, O.; PALESTRA, G.; TREVINO, R.; CAROLIS, B. D. The humanoid robot nao as trainer in a memory program for elderly people with mild cognitive impairment. **International Journal of Social Robotics**, Springer Netherlands, Feb 2019. Disponível em: <https://doi.org/10.1007/s12369-019-00533-y>. Citado na página 30.

PRANAV, E.; KAMAL, S.; CHANDRAN, C. S.; SUPRIYA, M. H. Facial emotion recognition using deep convolutional neural network. **2020 6th International Conference on Advanced Computing and Communication Systems, ICACCS 2020**, Institute of Electrical and Electronics Engineers Inc., p. 317–320, 3 2020. Disponível em: <https://doi.org/10.1109/ICACCS48705.2020.9074302>. Citado nas páginas 15 e 47.

PROVOOST, S.; LAU, H. M.; RUWAARD1, J.; RIPER, H.; CLINICAL, D. o.; PROVOOST, C. A. **Embodied Conversational Agents in Clinical Psychology: A Scoping Review**. JMIR Publications Inc., Toronto, Canada, 2017. Disponível em: <https://doi.org/10.2196/jmir.6553>. Citado na página 31.

RIEK, L. Wizard of Oz Studies in HRI: A Systematic Review and New Reporting Guidelines. **Journal of Human-Robot Interaction**, p. 119–136, 8 2012. ISSN 21630364. Disponível em: <https://doi.org/10.5898/JHRI.1.1.Riek>. Citado na página 33.

RIEK, L. D. Wizard of oz studies in hri: A systematic review and new reporting guidelines. **J. Hum.-Robot Interact.**, Journal of Human-Robot Interaction Steering Committee, v. 1, n. 1, p. 119–136, jul. 2012. Disponível em: <https://doi.org/10.5898/JHRI.1.1.Riek>. Citado na página 27.

SOLER, M. V.; AGüERA-ORTIZ, L.; RODRíGUEZ, J. O.; REBOLLEDO, C. M.; MUñOZ, A. P.; PéREZ, I. R.; RUIZ, E. O.; SáNCHEZ, A. B.; CANO, V. H.; CHILLóN, L. C.; RUIZ, S. F.; ALVAREZ, J. L.; SALAS, B. L.; PLAZA, J. C.; RICO, F. M.; MARTíN, P. M. Social robots in advanced dementia. **Frontiers in Aging Neuroscience**, v. 7, p. 133, 2015. ISSN 1663-4365. Disponível em: <https://doi.org/10.3389/fnagi.2015.00133>. Citado na página 30.

STAHL, C.; ANASTASIOU, D.; LATOUR, T. Social Telepresence Robots. In: **Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference**. New York, NY, USA: ACM, 2018. p. 409–414. ISBN 9781450363907. Disponível em: <https://doi.org/10.1145/3197768.3203180>. Citado nas páginas 25 e 56.

STOWER, R.; CALVO-BARAJAS, N.; CASTELLANO, G.; KAPPAS, A. **A Meta-analysis on Children's Trust in Social Robots**. Springer Netherlands, 2021. Disponível em: <https://doi.org/10.1007/s12369-020-00736-8>. Citado na página 31.

TROVATO, G.; RAMOS, J. G.; AZEVEDO, H.; MORONI, A.; MAGOSSI, S.; SIMMONS, R.; ISHII, H.; TAKANISHI, A. A receptionist robot for brazilian people: study on interaction involving illiterates. **Paladyn, Journal of Behavioral Robotics**, v. 8, n. 1, p. 1–17, 4 2017. ISSN 2081-4836. Disponível em: <https://doi.org/10.1515/pjbr-2017-0001>. Citado na página 36.

WIJNEN, F. M.; DAVISON, D. P.; REIDSMA, D.; MEIJ, J. V. D.; CHARISI, V.; EVERS, V. Now we're talking: Learning by explaining your reasoning to a social robot. **J. Hum.-Robot Interact.**, Association for Computing Machinery, New York, NY, USA, v. 9, n. 1, dec 2019. Disponível em: <https://doi.org/10.1145/3345508>. Citado na página 29.

ZADEH, M. M. T.; IMANI, M.; MAJIDI, B. Fast facial emotion recognition using convolutional neural networks and gabor filters. **2019 IEEE 5th Conference on Knowledge Based Engineering and Innovation, KBEI 2019**, Institute of Electrical and Electronics Engineers Inc., p. 577–581, 2 2019. Disponível em: <https://doi.org/10.1109/KBEI.2019.8734943>. Citado nas páginas 15, 47 e 48.

ZHANG, Y.; SONG, W.; TAN, Z.; WANG, Y.; LAM, C. M.; HOI, S. P.; XIONG, Q.; CHEN, J.; YI, L. Theory of robot mind: False belief attribution to social robots in children with and without autism. **Frontiers in Psychology**, v. 10, 2019. ISSN 1664-1078. Disponível em: <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01732>. Citado na página 28.

GLOSSÁRIO

Framework: é uma abstração que une códigos comuns entre vários projetos de *software* provendo uma funcionalidade genérica.*Frameworks* são projetados com a intenção de facilitar o desenvolvimento de *software*, habilitando designers e programadores a gastarem mais tempo determinando as exigências do *software* do que com detalhes de baixo nível do sistema.

ICMC USP
SÃO CARLOS