

Papel original

O desempenho de classificadores de emoções para crianças com autismo relatado pelos pais: estudo quantitativo de viabilidade

Haik Kalantarian^{1,2}, BSc, MS, DPhil; Khaled Jedoui³, Licenciatura; Kaitlyn Dunlap^{1,2}, BA, MRes; Jessey Schwartz^{1,2}, BA; Peter Washington^{1,2}, BA, MSc; Arman Husic^{1,2}, Licenciatura; Qandeel Tariq^{1,2}, BSc, MSci; Michael Ning^{1,2}, Licenciatura; Aaron Kline^{1,2}, Licenciatura; Dennis Paul Wall^{1,2,4}, bacharelado, mestrado, doutorado

¹Departamento de Pediatria, Stanford University, Stanford, CA, Estados Unidos

²Departamento de Ciência de Dados Biomédicos, Universidade de Stanford, Stanford, CA, Estados Unidos

³Departamento de Matemática, Stanford University, Stanford, CA, Estados Unidos

⁴Departamento de Psiquiatria e Ciências Comportamentais, Universidade de Stanford, Stanford, CA, Estados Unidos

Autor correspondente:

Dennis Paul Wall, BSc, MSc, DPhil

Departamento de Pediatria

Universidade de Stanford

Shopping Serra 450

Stanford, CA, 94305

Estados Unidos

Telefone: 1 (650) 7232300 E-

mail: dpwall@stanford.edu

Abstrato

Fundo: O transtorno do espectro autista (TEA) é um transtorno do desenvolvimento caracterizado por déficits na comunicação e interação social e comportamentos e interesses restritos e repetitivos. A incidência de TEA tem aumentado nos últimos anos; estima-se agora que aproximadamente 1 em cada 40 crianças nos Estados Unidos seja afetada. Devido em parte ao aumento da prevalência, o acesso ao tratamento tornou-se restrito. A esperança está em soluções móveis que fornecem terapia por meio de abordagens de inteligência artificial (IA), incluindo modelos de IA de detecção facial e de emoção desenvolvidos por provedores de nuvem convencionais, disponíveis diretamente para os consumidores. No entanto, essas soluções podem não ser suficientemente treinadas para uso em populações pediátricas.

Objetivo: Classificadores de emoções disponíveis para o público em geral através da Microsoft, Amazon, Google e Sighthound são adequados para a população pediátrica e podem ser usados para desenvolver terapias móveis visando aspectos de comunicação e interação social, talvez acelerando a inovação em este espaço. Este estudo teve como objetivo testar esses classificadores diretamente com dados de imagens de crianças com TEA relatadas pelos pais recrutadas por meio de crowdsourcing.

Métodos: Usamos um jogo para celular chamado *Adivinha?* que desafia uma criança a encenar uma série de instruções exibidas na tela do smartphone colocado na testa de seu provedor de cuidados. O jogo pretende ser uma maneira divertida e envolvente para a criança e os pais interagirem socialmente, por exemplo, o pai tentando adivinhar qual emoção a criança está expressando (por exemplo, surpreso, assustado ou enojado). Durante uma sessão de jogo de 90 segundos, até 50 comandos são mostrados enquanto a criança age, e o vídeo registra as ações e expressões da criança. Devido em parte à natureza divertida do jogo, é uma maneira viável de envolver remotamente populações pediátricas, incluindo a população autista por meio de crowdsourcing. Recrutamos 21 crianças com TEA para jogar o jogo e reunimos 2602 quadros emotivos após suas sessões de jogo.

Resultados: Todos os classificadores tiveram um desempenho ruim para todas as emoções avaliadas, exceto feliz. Nenhum dos classificadores rotulou corretamente mais de 60,18% (1566/2602) dos quadros avaliados. Além disso, nenhum dos classificadores identificou corretamente mais de 11% (6/51) dos quadros de raiva e 14% (10/69) dos quadros de nojo.

Conclusões: Os resultados sugerem que os classificadores de emoções comerciais podem ser insuficientemente treinados para uso em abordagens digitais para tratamento de autismo e rastreamento de tratamento. Métodos seguros e de preservação de privacidade para aumentar os dados de treinamento rotulados são necessários para aumentar o desempenho dos modelos antes que eles possam ser usados em abordagens habilitadas por IA para terapia social do tipo comum em tratamentos de autismo.

PALAVRAS-CHAVE

celular; emoção; autismo; dados digitais; aplicativo móvel; mSaúde; afetar; aprendizado de máquina; inteligência artificial; saúde digital

Introdução

Fundo

O transtorno do espectro autista (TEA) é um transtorno do neurodesenvolvimento caracterizado por comportamentos e interesses estereotipados e repetitivos, bem como déficits na interação social e na comunicação.^{1]} Além disso, as crianças autistas lutam com o afeto facial e podem se expressar de maneiras que não se assemelham muito às de seus pares.^{2-4]} A incidência de TEA tem aumentado nos últimos anos; estima-se agora que aproximadamente 1 em cada 40 crianças nos Estados Unidos é afetada por esta condição.^{5]} Embora o autismo não tenha cura, há fortes evidências que sugerem que a intervenção precoce pode melhorar as habilidades de fala e comunicação.^{6]}

As abordagens comuns para a terapia do autismo incluem análise comportamental aplicada (ABA) e o modelo de início precoce de Denver (ESDM). Na terapia ABA, a intervenção é personalizada por um analista comportamental treinado para se adequar especificamente às habilidades e déficits do aluno.^{7]} A base deste programa é uma série de atividades estruturadas que enfatizam o desenvolvimento de habilidades transferíveis para o mundo real. Da mesma forma, intervenções comportamentais de desenvolvimento naturalistas, como ESDM, apoiam o desenvolvimento de habilidades sociais centrais por meio de interações com um terapeuta comportamental licenciado, enfatizando atividades conjuntas e trocas interpessoais.^{8]} Ambos os tipos de tratamento demonstraram ser seguros e eficazes, com seu maior potencial de impacto ocorrendo durante a intervenção precoce em idades mais jovens.^{9-11]}

Apesar do progresso significativo na compreensão dessa condição nos últimos anos, permanecem desequilíbrios na cobertura e barreiras ao diagnóstico e tratamento. Nos países em desenvolvimento, estudos notaram a falta de profissionais de saúde treinados, tratamentos inconsistentes e um caminho pouco claro do diagnóstico à intervenção.^{12-14]} Nos Estados Unidos, pesquisas mostraram que crianças em áreas rurais recebem diagnósticos aproximadamente 5 meses mais tarde do que crianças que vivem em cidades.^{15]} Além disso, observou-se que crianças de famílias próximas à linha de pobreza recebem diagnósticos quase um ano mais tarde do que aquelas de famílias de renda mais alta. Abordagens baseadas em dados estimam que mais de 80% dos condados dos EUA não contêm recursos de diagnóstico de autismo^[16]. Mesmo meses de atraso no acesso à terapia podem limitar a eficácia das intervenções comportamentais subsequentes.^{15]} Soluções alternativas que podem melhorar

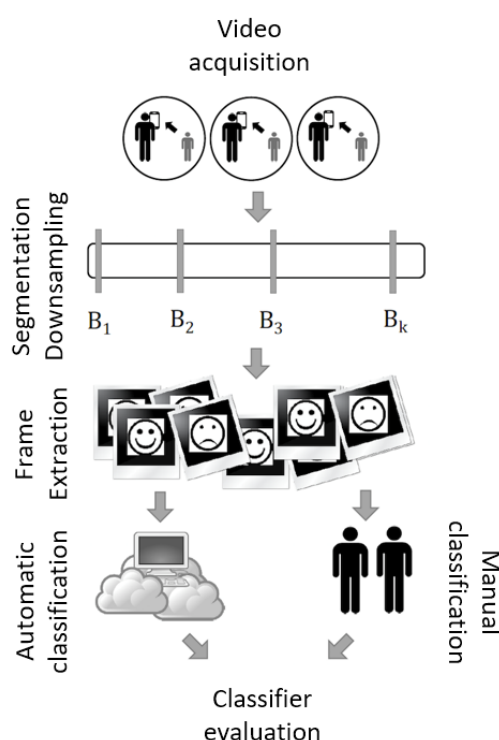
alguns desses desafios podem ser derivados de ferramentas digitais e móveis. Por exemplo, desenvolvemos um sistema vestível usando o Google Glass que aproveita algoritmos de classificação de emoções para reconhecer a emoção facial do parceiro de conversa de uma criança para feedback em tempo real e suporte social e mostrou a eficácia do tratamento em um ensaio clínico randomizado^[17-25].

Vários classificadores de emoções baseados em nuvem podem ajudar no valor e no alcance de ferramentas e soluções móveis. Estes incluem quatro sistemas disponíveis comercialmente: Interface de programação de aplicativos (API) do Microsoft Azure Emotion^[26], Amazon Rekognition^[27], Google Cloud Vision^[28], e Sighthound^[29]. Considerando que a maioria das implementações dessas APIs de reconhecimento de emoções são proprietárias, esses algoritmos são normalmente treinados usando grandes conjuntos de dados de emoções faciais, como o banco de dados Cohn-Kanade.^{30]} e banco de dados de emoções naturais induzidas por Belfast^[31], que possuem poucos exemplos de crianças. Devido a esse viés nos exemplos rotulados, é possível que esses modelos não generalizem bem para a população pediátrica, incluindo crianças com atrasos no desenvolvimento, como o autismo, que é avaliado neste estudo. Este estudo põe a disparidade à prova. Para isso, usamos nosso jogo para celular *Adivinha*^[32-35]. Este jogo (nativo para Android^[36] e iOS^[37]) promove o engajamento entre a criança e seu parceiro social, como um pai, por meio de jogos semelhantes a charadas enquanto constrói um banco de dados de imagens faciais enriquecidos para uma gama de emoções exibidas pela criança durante as sessões de jogo.

As principais contribuições deste estudo são as seguintes:

1. Apresentamos um jogo de charadas para celular, *Guess What?*, para obter vídeos emotivos de seus jogadores. Essa estrutura tem utilidade tanto como mecanismo para a avaliação de classificadores de emoções existentes quanto para o desenvolvimento de novos sistemas que generalizam adequadamente para a população de interesse.
2. Apresentamos um estudo no qual 2.602 quadros emotivos são derivados de 21 crianças com diagnóstico de autismo relatado pelos pais usando dados do jogo para celular *Guess What* coletados em uma variedade de ambientes heterogêneos.
3. Os dados foram usados para avaliar a precisão e o desempenho de vários classificadores de última geração usando o fluxo de trabalho mostrado em figura 1, para desenvolver uma compreensão da viabilidade de usar essas APIs em futuras abordagens de terapia móvel.

Figura 1. Um jogo de charadas para celular jogado entre cuidador e criança é usado para crowdsourcing vídeo emotivo, subamostrado e categorizado por avaliadores manuais e classificadores automáticos. Os quadros desses vídeos formam a base do nosso conjunto de dados para avaliar vários classificadores de emoções.



Trabalho relatado

Até onde sabemos, este é o primeiro trabalho até o momento que compara APIs públicas de reconhecimento de emoções em crianças com atrasos no desenvolvimento. No entanto, vários aplicativos interessantes foram propostos nos últimos anos, que empregam ferramentas baseadas em visão ou soluções de computação afetiva como auxílio para crianças com autismo. O surgimento dessas abordagens motiva uma investigação cuidadosa da viabilidade de algoritmos comerciais de classificação de emoções para a população pediátrica.

Motivados pelo fato de que crianças com autismo podem experimentar sobrecarga cognitiva ou emocional, o que pode comprometer suas habilidades de comunicação e experiência de aprendizagem, Picard et al.[38] forneceu uma visão geral dos avanços tecnológicos para detectar a ativação do sistema nervoso autônomo em tempo real, incluindo sensores de atividade eletrodérmica vestíveis. Uma visão mais geral do papel da computação afetiva no autismo é fornecida por Kalioby et al.[39], com os exemplos motivadores do uso da tecnologia para ajudar os indivíduos a navegar melhor no cenário socioemocional de suas vidas diárias. Entre os dispositivos enumerados estão aqueles desenvolvidos no laboratório de mídia do Massachusetts Institute of Technology, como *óculos de expressão que discriminam entre várias emoções*, *luvas sensíveis à condutância da pele para detecção de estresse* e um *mouse sensível à pressão para inferir o estado afetivo de como os indivíduos interagem com o dispositivo*. Os dispositivos fabricados pela indústria incluem a *braceleteira SenseWear Pro2*, que inclui uma variedade de sensores vestíveis que podem ser reaproveitados para detecção de estresse e produtividade, *luvas inteligentes que podem detectar frequência respiratória e pressão arterial e monitores de frequência cardíaca sem fio que podem ser analisados no contexto de estressores ambientais* [40].

Pesquisas anteriores realizadas por nós demonstraram a eficácia das abordagens de fenotipagem de vídeo móvel para crianças com TEA em geral [41-47] e por meio do uso de classificadores de emoções integrados à plataforma Google Glass para fornecer suporte comportamental em tempo real a crianças com TEA [17-25]. Além disso, outros estudos confirmaram a usabilidade, aceitação e impacto positivo geral nas famílias de sistemas baseados no Google Glass que usam a tecnologia de reconhecimento de emoções para auxiliar a comunicação e interação socioemocional para crianças autistas.[48,49]. Além desses esforços, uma variedade de outros dispositivos de vidro inteligente foi proposta. Por exemplo, o SenseGlass [50] está entre os primeiros trabalhos que propõem alavancar a plataforma Google Glass para capturar e processar informações afetivas em tempo real usando uma variedade de sensores. Os autores propuseram aplicativos, incluindo o desenvolvimento de interfaces de usuário baseadas em afeto e capacitando os usuários para mudanças comportamentais por meio de intervenções de gerenciamento de emoções.

O reconhecimento de efeito baseado em vidro que antecede a plataforma Google Glass também foi proposto. Scheirer et al [51] usaram sensores piezoelétricos para detectar expressões como confusão e interesse, que foram detectadas com precisão de 74%. Um trabalho mais recente propõe um dispositivo chamado *Óculos de empatia* [52] em que os usuários podem ver, ouvir e sentir da perspectiva de outro indivíduo. O sistema consiste em hardware vestível para transmitir o olhar e a expressão facial dos usuários e uma interface remota onde o feedback visual é fornecido e os dados são visualizados.

A pesquisa para intervenções baseadas em vidro inteligente é ainda apoiada por outros sistemas tecnológicos que foram desenvolvidos e examinados no contexto de atrasos de desenvolvimento, incluindo o uso de realidade aumentada para treinamento de discriminação de objetos [53], robótica assistiva para terapia [54-56] e tecnologias assistivas móveis para redes sociais em tempo real.

aprendizado de habilidades [57]. Além disso, o uso de visão computacional e sistemas gamificados para detectar e ensinar emoções continua a progredir. Uma abordagem computacional para **detectar expressões faciais otimizada para plataformas móveis** foi proposta [58], que **demonstrou uma precisão de 95% a partir de um conjunto de 6 classes de expressões**. Leo e outros [59] propuseram uma abordagem computacional para **avaliar a capacidade de crianças com TEA de produzir expressões faciais usando visão computacional, validada por três avaliadores especialistas**. Suas descobertas demonstraram a viabilidade de um sistema de visão computacional humano no circuito para analisar dados faciais de crianças com TEA. Semelhante a este estudo, que utiliza *Adivinha?*, **um jogo para celular no estilo de charadas para coletar dados emocionais de rostos**, Park e colegas **propuseram seis métodos de design de jogos** para o desenvolvimento de estruturas orientadas a jogos **no ensino de emoções** para crianças com TEA, dos quais incluem: **observação, compreensão, imitação e generalização**, e apoia o uso de jogos para produzir dados de valor para abordagens de visão computacional para crianças com autismo [60].

Embora nem todos os estudos de pesquisa mencionados acima empreguem modelos de reconhecimento de emoções diretamente, eles são indicativos de uma transição geral das práticas tradicionais de saúde para soluções móveis e digitais modernas que alavancam avanços recentes em visão computacional, realidade aumentada, robótica e inteligência artificial [61]. Assim, **a tendência motiva nossa investigação da eficácia de modelos de visão de última geração em populações com atraso no desenvolvimento**.

Métodos

Visão geral

Nesta seção, descrevemos **a arquitetura de Adivinha?** seguido por uma **descrição dos métodos empregados para obter dados de teste e processar os quadros neles para avaliar o desempenho de vários classificadores principais de emoção**. Embora dezenas de APIs estejam disponíveis, limitamos nossa análise a alguns dos sistemas mais **populares dos principais** provedores de serviços em nuvem como uma representação justa do estado da arte em APIs de reconhecimento de emoções disponíveis publicamente. Os sistemas avaliados em

este trabalho foram **Microsoft Azure Emotion API (Azure) [26], Amazon AWS Rekognition (AWS) [27], API do Google Cloud Vision (Google) [28], e Sighthound (SH) [29]**.

Arquitetura do sistema

A avaliação do estado da arte em APIs de classificação de emoções públicas em crianças com TEA **requer um conjunto de dados derivado de sujeitos do grupo populacional relevante com uma quantidade razoável de consistência em seu formato e estrutura**. Além disso, como os dados são limitados, **é fundamental que o vídeo contenha uma alta densidade de quadros emotivos para simplificar o processo de anotação manual ao estabelecer uma verdade geral**. Por isso, desenvolvemos e lançamos um jogo educacional para celular na **Google Play Store [34] e iOS App Store [35]**, *Adivinha?*, do qual derivamos o vídeo emotivo.

Neste jogo, **os pais seguram o telefone** de forma que **a câmera frontal e a tela estejam voltadas para a criança**. Quando a sessão de jogo começa, a criança recebe um aviso que o cuidador deve adivinhar com base nos gestos e expressões faciais da criança. **Depois que um palpite correto é reconhecido, o pai inclina o telefone para frente, indicando que um ponto deve ser concedido**. Neste momento, outro prompt é mostrado. **Se quem estiver segurando o telefone não puder adivinhar, ele inclinará o telefone para trás para pular o quadro e prosseguir automaticamente para o próximo**. Esse processo se repete até que a sessão de jogo de 90 segundos tenha decorrido. Meta-informações são geradas para cada sessão de jogo que indica os horários em que vários prompts são mostrados e quando ocorrem as suposições corretas.

Embora uma série de prompts variados estejam disponíveis, os dois que são mais relevantes para o reconhecimento de afeto facial e reconhecimento de emoção são **emojis e rostos**, conforme mostrado em **Figuras 2e3**, respectivamente. Após a conclusão da sessão do jogo, os cuidadores podem optar por **compartilhar seus arquivos e metadados associados** a um bucket seguro do Amazon S3 aprovado pelo conselho de revisão da instituição e totalmente compatível com os padrões de segurança de aplicativos de alto risco da Universidade de Stanford. Uma discussão mais detalhada da mecânica e aplicações de *Adivinha?* é descrito em [29-32].

Figura 2. Os prompts da categoria emoji são caricaturas, mas muitos ainda estão associados às emoções universais clássicas de Ekman.

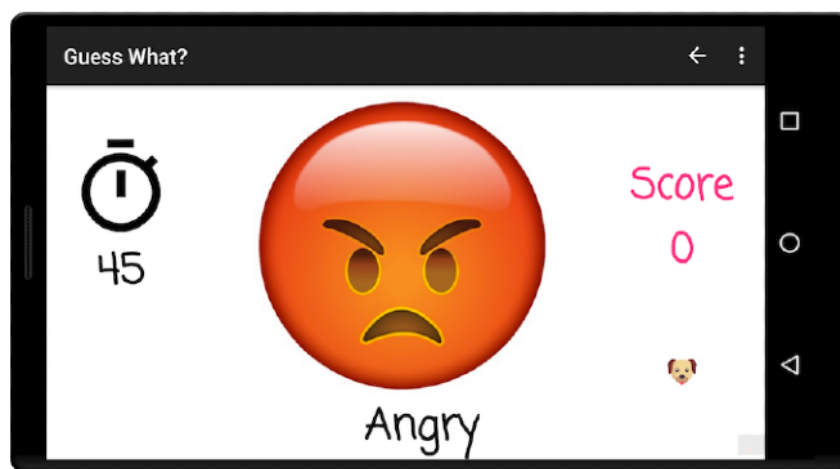
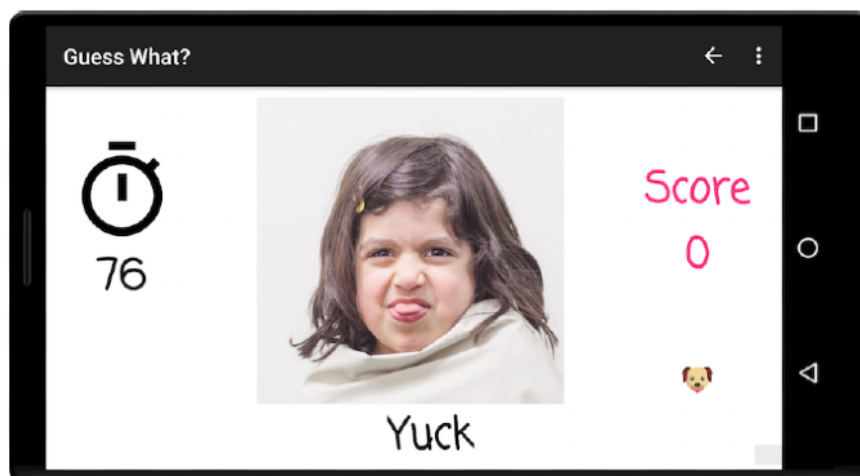


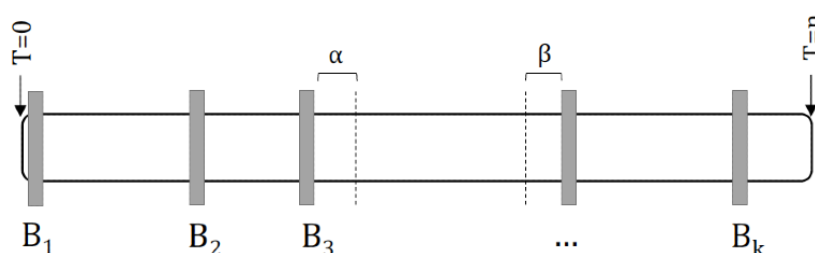
Figura 3. Os prompts da categoria de rostos são derivados de fotos reais de crianças sobre um fundo sólido.



A estrutura de um vídeo é mostrada em Figura 4. Cada vídeo enviado rende quadros de vídeo, delineados por pontos de fronteira, B_1-B_k , onde cada ponto de fronteira representa o momento em que um novo prompt é mostrado ao usuário. Para obter frames associados a uma determinada emoção, deve-se primeiro identificar o ponto limite associado a essa emoção através da metainformação do jogo, eu . Tendo identificado este ponto de fronteira, os quadros entre B_{eu} e B_{eu+1} pode ser associado a este prompt.

No entanto, dois fatores adicionais permanecem. Normalmente, leva algum tempo, α , para que a criança reaja após a exibição do prompt. Além disso, muitas vezes há um período de tempo, β , entre o reconhecimento da criança do palpite dos pais e a inclinação do telefone pelos pais, durante o qual a criança pode adotar uma expressão facial neutra. Portanto, os quadros de interesse são aqueles que se situam entre $B_{eu} + \alpha$ e $B_{eu+1} - \beta$.

Figura 4. A estrutura de um único vídeo é caracterizada por seus pontos de fronteira, que identificam os momentos em que vários prompts foram mostrados à criança.



O sistema proposto está centrado em dois objetivos principais. Primeiro, esse mecanismo facilita a aquisição de vídeos emotivos estruturados de crianças de uma maneira que desafia sua capacidade de expressar emoção facial. Enquanto outras formas de captura de vídeo podem ser empregadas, um sistema gamificado encoraja o uso repetido e tem o potencial de conter uma densidade muito maior de quadros emotivos do que um vídeo caseiro típico estruturado em atividades que não sejam jogos. Como a anotação manual é empregada como base para avaliar a classificação de emoções, uma alta concentração de quadros emotivos em um curto período de tempo é essencial para a simplificação e redução da carga associada a esse processo. Um segundo objetivo é potencialmente facilitar a agregação de vídeos emotivos rotulados de crianças usando um mecanismo de crowdsourcing.

Aquisição de dados

Um total de 46 vídeos de 21 sujeitos foram analisados neste estudo. Esses dados foram coletados ao longo de 1 ano. Dez vídeos foram

coletados em ambiente de laboratório de seis sujeitos com TEA que jogaram vários jogos em uma única sessão administrada por um membro da equipe de pesquisa. Outros 36 vídeos foram adquiridos por meio de crowdsourcing de 15 participantes remotos. O diagnóstico de qualquer forma de transtorno do desenvolvimento foi fornecido pelo cuidador por meio de autorrelato durante o processo de registro, juntamente com informações demográficas (sexo, idade, etnia). As informações coletadas incluíam diagnósticos de transtorno autista (autismo), TEA, síndrome de Asperger, transtorno invasivo do desenvolvimento (não especificado de outra forma), transtorno desintegrativo da infância, sem diagnóstico, sem diagnóstico, mas suspeito e transtorno de comunicação social (pragmático). Além disso, um campo de texto livre estava disponível para os pais especificarem condições adicionais.^{1]} Os cuidadores de todas as crianças que participaram do estudo selecionaram o transtorno do espectro do autismo opção.

O formato de um Adivinha? sessão de jogo geralmente impõe uma estrutura no vídeo derivado: o dispositivo é mantido no modo paisagem, o rosto da criança está contido no quadro,

e a distância entre a criança e a câmera é tipicamente entre 2 e 10 pés. No entanto, esses vídeos foram cuidadosamente selecionados por membros da equipe de pesquisa para garantir a confiabilidade e a qualidade dos dados neles contidos; vídeos que não incluíam crianças, eram corrompidos, filmados em más condições de iluminação ou não incluíam informações demográficas plausíveis foram excluídos da análise. A média de idade das crianças participantes foi de 7,3 (1,76) anos. Devido ao pequeno tamanho da amostra e à incidência não uniforme de autismo entre os gêneros.^{62]}, 18 dos 21 participantes eram do sexo masculino. Embora os participantes tenham explorado uma variedade de mecânicas de jogo, todos os vídeos analisados foram derivados das duas categorias mais úteis para o estudo do afeto facial: rostos e emojis. Após cada sessão de jogo, os vídeos eram carregados automaticamente em um bucket do Amazon S3 por meio de um processo em segundo plano do Android.

Processamento de dados

A maioria das APIs de classificação de emoções cobra dos usuários por uma solicitação http, tornando o processamento de cada quadro em um vídeo proibitivo em termos de tempo e custo. Para simplificar nossa avaliação, subamostramos cada vídeo a uma taxa de dois quadros por segundo. Esses quadros formaram a base de nossos experimentos. Para obter a verdade, dois avaliadores atribuíram manualmente um rótulo de emoção a cada quadro com base nas sete emoções universais de Ekman.^{63]}, com a adição de um rótulo neutro. Alguns quadros

foram descartados quando não havia rosto, ou a qualidade era muito ruim para fazer uma avaliação. O desempenho de um classificador em um quadro foi avaliado apenas sob as condições em que o quadro era válido (de qualidade suficiente), e os dois avaliadores manuais concordaram com o rótulo de emoção associado ao quadro. Os quadros foram considerados de qualidade insuficiente se: (1) o quadro estava muito desfocado para discernir, (2) a criança não estava no quadro, (3) a imagem ou o vídeo estava corrompido ou (4) havia vários indivíduos dentro do quadro. quadro.

De um total de 5.418 frames revisados, 718 foram descartados por falta de concordância entre os avaliadores manuais. Um adicional de 2.123 quadros foram descartados porque pelo menos um avaliador atribuiu o não aplicável (N/A) etiqueta, indicando que o quadro era de qualidade insuficiente. Isso ocorreu devido a uma variedade de fatores, mas geralmente causado por artefatos de movimento ou pela saída da criança do quadro devido à inclinação do telefone em reconhecimento de um palpite correto. O número total de frames analisados foi de 2602 divididos entre as categorias apresentadas em tabela 1.

Como mostrado, a maioria dos quadros foram *neutro*, com preponderância de *feliz* quadros na categoria não neutra. Devido ao número limitado de *assustada* e *confuso* frames, essa emoção foi omitida de nossa análise. Também fundimos o *desprezo* e *raiva* categorias devido à sua semelhança de afeto e agilizarmos a análise.

Tabela 1. A distribuição de quadros por categoria (N=2602).

Emoção	Quadros, n
Neutro	1393
Emotivo	1209
Feliz	864
Triste	60
Surpreso	165
Com nojo	69
Bravo	51

Como nem todos os classificadores de emoções representaram suas saídas em um formato consistente, algumas simplificações adicionais foram feitas em nossa análise. Primeiro, foi necessário fazer pequenas correções no formato dos dados de saída. Por exemplo, *feliz* *felicidade* foram considerados idênticos. No caso da AWS, o *confuso* class foi ignorado, pois muitos outros classificadores não a suportavam. Além disso, *calma* foi renomeado *neutro*. Como AWS, Azure e Sighthound retornaram probabilidades em vez de um único rótulo, um quadro no qual nenhuma classe de emoção foi associada a uma probabilidade superior a 70% foi considerado uma falha. Para o Google Vision, a confiança da classificação foi associada a um rótulo categórico em vez de uma porcentagem. Nesse caso, os frames não receberam uma classificação de emoção como *provável* ou *muito provável* foram considerados fracassos. Também vale a pena notar que esta plataforma, ao contrário de todas as outras, não contém *desgosto* ou *neutro* Aulas. As emoções finais avaliadas neste estudo foram *feliz*, *triste*, *surpresa*, *raiva*, *desgosto*, e *neutro*, com os dois últimos omitidos para o Google Cloud Vision.

Como o uso em tempo real é um aspecto importante das terapias e auxílios móveis, avaliamos o desempenho de cada classificador calculando o número de segundos necessários para processar cada vídeo de 90 segundos subamostrado para um quadro por segundo. Essa avaliação foi realizada em uma rede Wi-Fi testada com velocidade média de download de 51 Mbps e velocidade média de upload de 62,5 Mbps. Para cada classificador, esse experimento foi repetido 10 vezes para obter o tempo médio necessário para processar o vídeo subamostrado.

Resultados

Visão geral

Nesta seção, apresentamos os resultados de nossa avaliação de *Adivinha*. Bem como o desempenho dos classificadores avaliados: Microsoft Azure Emotion API (Azure) [26], AWS [27], API do Google Cloud Vision (Google) [28], e SH [29]. As abreviações para emoções descritas nesta seção podem ser encontradas em Caixa de texto 1.

Caixa de texto 1. Abreviaturas para emoções.

HP: Feliz
FC: Confuso
N/A: Não aplicável
SC: Assustado
SP: Surpreso
DG: Nojo
AG: Com raiva

Precisão do classificador

Comparação com Ground Truth (Classificadores)

mesa 2 mostra o desempenho de cada classificador calculado pela porcentagem de quadros corretamente identificados em comparação com a verdade do terreno para as categorias neutro, emotivo, estudo. UMA neutro

O quadro é aquele em que o rosto é reconhecido e o rótulo neutro recebe alta confiança. Qualquer outro quadro dentro das categorias de feliz, triste, surpreso, com nojo, e Bravo, são considerados quadros emotivos. Uma divisão mais detalhada do desempenho por emoção é mostrada em Tabela 3. Observe que, como antes, a API do Google não suporta o neutro e desgosto categorias.

Mesa 2. Porcentagem de frames identificados corretamente pelo classificador: Azure (Azure Cognitive Services), AWS (Amazon Web Services), SH (Sighthound) e Google (Google Cloud Vision). Esses resultados incluem apenas quadros em que havia um rosto, e os dois avaliadores manuais concordaram com a classe. A API do Google Vision não é compatível com o rótulo neutro;

Classificador	Tipo de moldura		
	Emotivo (n=1209), n (%)	Neutro (n=1393), n (%)	Todos (n=2602), n (%)
Azure	798 (66,00)	744 (53,40)	1542 (59,26)
AWSuma	829 (68,56)	679 (48,74)	1508 (57,95)
Google	785 (64,92)	N / Db	N / D
Galgo	664 (54,92)	902 (64,75)	1566 (60,18)

umaAWS: Amazon AWS Rekognition.

bN/A: não aplicável.

Tabela 3. Porcentagem de frames identificados corretamente por tipo de emoção por cada classificador: Azure (Azure Cognitive Services), AWS (Amazon Web Services), SH (Sighthound) e Google (Google Cloud Vision). Esses resultados incluem apenas quadros em que havia um rosto, e os dois avaliadores manuais concordaram com a classe.

Observação: a API do Google Vision não é compatível com os rótulos neutros ou nojentos.

Classificador	Tipo de moldura					
	Neutro (n=1394), n (%)	Feliz (n=864), n (%)	Triste (n=60), n (%)	Surpreso (n=165), n (%)	Desgostoso (n=69), n (%)	Irritado (n=51), n (%)
AWS	679 (48,74)	709 (82,0)	19 (31)	94 (56,9)	4 (5)	3 (5)
Galgo	902 (64,75)	545 (63,0)	13 (21)	90 (54,5)	10 (14)	6 (11)
Azure	744 (53,41)	695 (80,4)	20 (33)	80 (48,4)	0 (0)	3 (5)
Google	N / Duma	676 (78,2)	10 (16)	93 (56,3)	N / D	6 (11)

umaN/A: não aplicável.

Confiabilidade entre avaliadores (Classificadores)

A estatística kappa de Cohen é uma medida de confiabilidade entre observadores que leva em consideração a porcentagem de concordância devido ao acaso; uma consideração importante quando as classes possíveis são poucas em número. Figura 5 mostra a concordância entre cada par de classificadores com base em sua pontuação Cohen kappa calculada com base em cada quadro avaliado, em que uma pontuação de 1 indica perfeito

acordo. Os resultados refletem baixa concordância entre a maioria das combinações de classificadores. Isso é particularmente verdadeiro para a falta de acordo entre Google e Sighthound, com uma pontuação kappa de Cohen de 0,2. Isso provavelmente ocorre devido às diferenças em como os classificadores são ajustados para precisão e recuperação; Sighthound identificou corretamente mais neutro frames do que os outros, mas o desempenho foi menor para o rótulo emotivo mais predominante: feliz,

Figura 5. O Cohen's Kappa Score é uma medida de concordância entre dois avaliadores e foi calculado para todos os quatro classificadores avaliados: Azure (Azure Cognitive Services), AWS (Amazon Web Services), SH (Sighthound) e Google (Google Cloud Vision). Os resultados indicam concordância fraca entre todos os pares de classificadores.

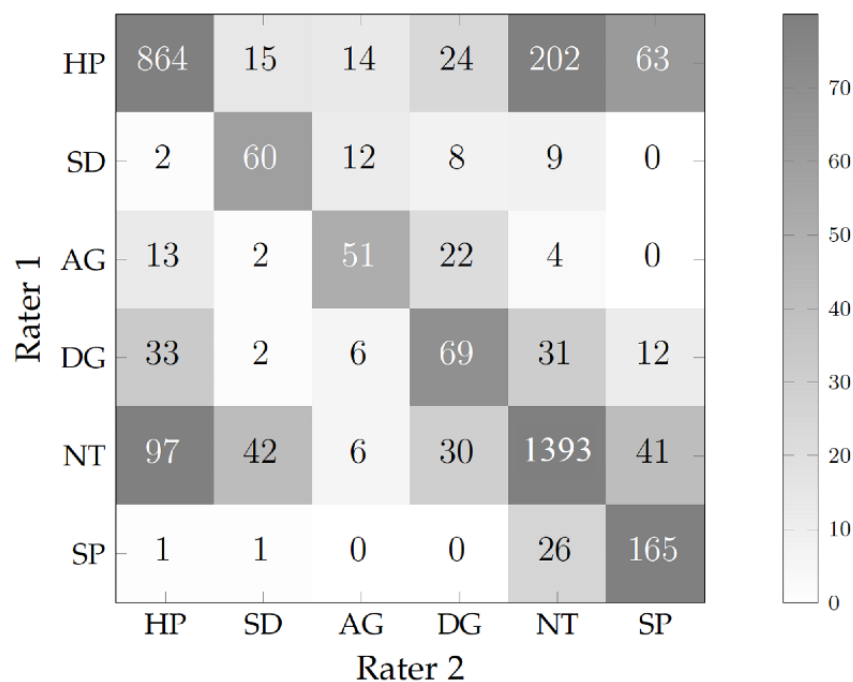


Confiabilidade entre avaliadores (avaliadores humanos)

O coeficiente kappa de Cohen para concordância entre os dois avaliadores manuais foi de 0,74, que foi maior do que qualquer combinação de classificadores automáticos avaliados neste estudo. Embora isso indique concordância substancial, vale a pena explorar as características dos quadros em que houve discordância

entre os dois avaliadores. A matriz de confusão completa pode ser vista em [Figura 6](#), que mostra a distribuição de todos os quadros avaliados pelos avaliadores. Os resultados indicam que a maioria das discrepâncias foi entre feliz e neutro. Essas discrepâncias provavelmente eram diferenças sutis em como os avaliadores percebiam um rosto feliz devido à subjetividade inerente a esse processo. A falta de concordância também pode ser observada entre as categorias de desgosto-raiva.

Figura 6. A distribuição de quadros entre os dois avaliadores humanos para cada emoção: HP (Feliz), SD (Triste), AG (Raiva), DG (Nojo), NT (Neutro) e SC (Assustado).



Velocidade do Classificador

As soluções vestíveis e móveis para autismo geralmente exigem desempenho de classificação eficiente para fornecer informações em tempo real.

feedback aos usuários. Em alguns casos, isso pode ser um feedback ambiental, como no [Autism Glass \[17-25\]](#), que usa a câmera voltada para fora do Google Glass para ler as emoções de quem está ao redor da criança e fornecer filas sociais em tempo real.

No caso de *Adivinha*, a câmera frontal do telefone é usada para ler a expressão da criança, que pode ser analisada para determinar se a expressão facial corresponde ao prompt exibido naquele momento.

Para determinar se o desempenho de classificação em tempo real é viável com implementações computacionais descarregadas de classificadores de emoção, medimos a quantidade de tempo necessária para processar um vídeo de 90 segundos gravado a 30 quadros por segundo e

subamostrados para um quadro por segundo, resultando em um total de 90 quadros. Para cada classificador, esse experimento foi repetido 10 vezes para obter o número médio de segundos necessários para processar o vídeo subamostrado. Tabela 4 mostra a velocidade dos classificadores baseados em API usados neste estudo. Os valores mostrados nesta tabela representam a quantidade de tempo necessária para enviar cada quadro ao serviço da Web por meio de uma solicitação de postagem http e receber uma resposta http com o rótulo de emoção. Esses quadros foram processados sequencialmente, sem sobreposição entre as requisições http.

Tabela 4. Velocidade dos classificadores avaliados.

Classificador	Tempo (segundos)
Azure	28,6
AWS	90,6
Google	55,9
Galgo	41.1

As descobertas indicaram que o classificador mais rápido foi o Azure, processando todos os 90 quadros em um total de 28,6 segundos. Usando o Azure com uma conexão de internet rápida, pode ser possível obter desempenho de classificação de emoções em tempo semi-real, um tempo de 28,6 segundos corresponde a 3,14 quadros por segundo, o que está dentro dos limites do que poderia ser considerado tempo real. O classificador mais lento foi o AWS, que processou esses 90 quadros em 90,6 segundos. Isso corresponde a uma taxa de quadros de 0,99 quadros por segundo. Em resumo, o desempenho em tempo real ou semi-real é possível com classificadores de emoção baseados na Web em conexões de internet Wi-Fi rápidas. Para conexões de celular ou aplicativos que exigem taxas de quadros superiores a três quadros por segundo, essas abordagens podem ser insuficientes.

Discussão

Desempenho do classificador

Os resultados indicam que o Google e a AWS produziram a maior porcentagem de quadros emotivos classificados corretamente, enquanto o Sighthound produziu a maior porcentagem de quadros neutros corretamente identificados. A API do Google não forneceu um neutro rótulo e, portanto, não pôde ser avaliada. O melhor sistema em termos de precisão de classificação geral foi o Sighthound por uma pequena margem, com 60,18% (1566/2602) dos quadros corretamente identificados. Outros resultados indicam que nenhum dos classificadores teve um bom desempenho para qualquer categoria além de feliz, que foi a emoção mais representada no conjunto de dados, como mostrado em tabela 1. Além disso, parece haver um viés sistemático de alta evocação e baixa precisão para a feliz categoria: aqueles classificadores que identificaram a maioria dos feliz quadros tiveram um desempenho pior para aqueles no neutro categoria.

Em resumo, os dados sugerem que, embora um quadro com um sorriso seja identificado corretamente na maioria dos casos, a capacidade dos classificadores avaliados em identificar outras expressões para crianças com TEA é duvidosa e apresenta um obstáculo no design de dispositivos móveis baseados em emoções e medidas de resultados vestíveis, ferramentas de triagem e terapias.

Análise de quadros

Figura 7 mostra seis quadros de um participante do estudo, reproduzidos com permissão dos pais da criança. A parte superior de cada quadro lista a anotação padrão-ouro na qual ambos os avaliadores concordaram em um rótulo adequado para o quadro. A parte inferior de cada quadro enumera os rótulos atribuídos de cada classificador em ordem: Amazon Rekognition, Sighthound, Azure Cognitive Services e Google Cloud Vision AI. Deve-se notar que, como antes, esses rótulos são normalizados para comparação porque cada classificador gera dados em um formato específico. Por exemplo, *N/D* de um classificador pode ser comparado com um campo em branco em outro, enquanto alguns, como o Google Cloud, declaram explicitamente *Não tenho certeza*; para nossos propósitos, todos esses três cenários foram rotulados como *N/D* durante a análise.

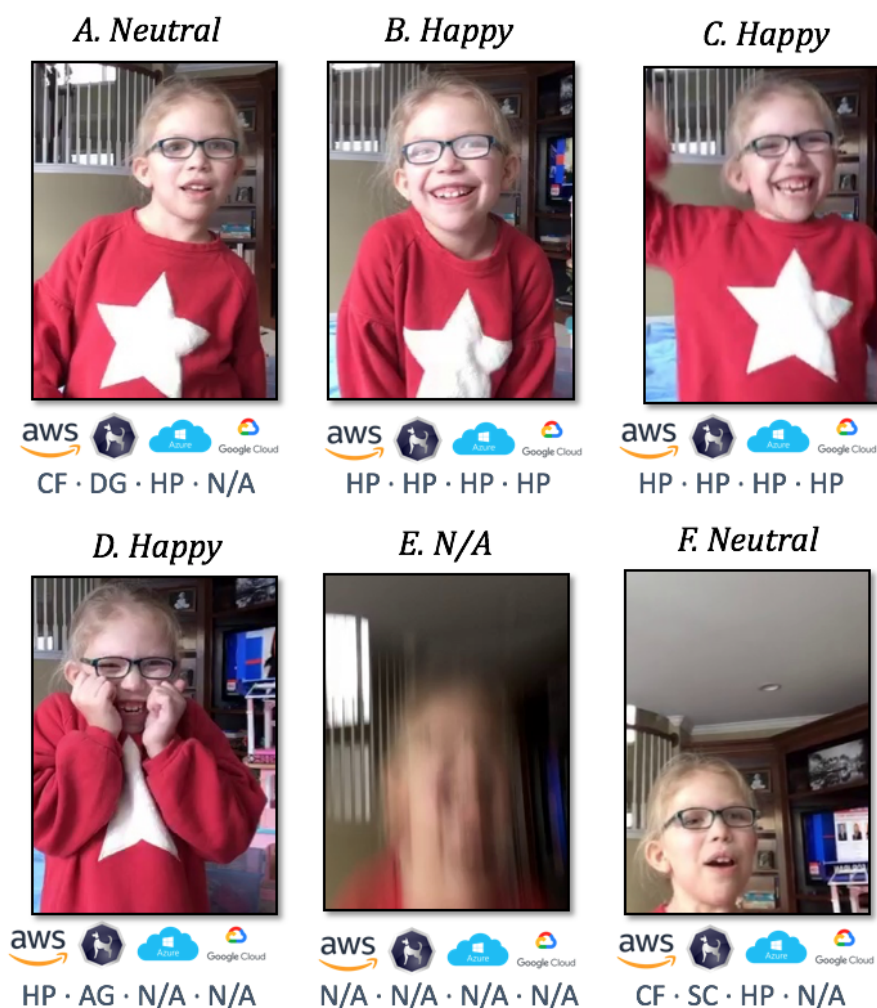
O quadro A mostra um quadro que foi rotulado como neutro pelos avaliadores, embora cada classificador tenha fornecido um rótulo diferente: *confuso*, *com nojo*, *feliz*, e *N/D*. Este é um exemplo de um falso-positivo, detectando uma emoção em um quadro neutro. Um exemplo semelhante é mostrado no quadro F; a maioria dos classificadores não conseguiu identificar o rótulo neutro. Esses falsos positivos são particularmente problemáticos, pois o neutro rótulo é o mais prevalente, como mostrado em tabela 1.

Em contraste, os quadros B e C são exemplos em que os rótulos atribuídos por cada classificador correspondem aos rótulos atribuídos pelos avaliadores manuais; todos os classificadores identificaram corretamente o feliz etiqueta. Como mostrado em tabela 1, feliz foi a emoção não neutra mais comum por uma margem considerável, e a maioria dos classificadores teve um bom desempenho nesta categoria; AWS, Azure e Sighthound, todos identificados corretamente entre 78,2% (676/864) e 82,0% (709/864) desses quadros, embora à custa do aumento de falsos positivos, como os mostrados nos quadros A e F. Um exemplo de um feliz O quadro que foi identificado como tal pelos avaliadores humanos, mas incorretamente pela maioria dos classificadores, é o quadro D. É possível que as mãos da criança cobrindo parte de seu rosto contribuam para esse erro, pois o quadro é bastante semelhante ao quadro B. Por fim, o quadro E é um exemplo de um quadro que foi processado pelos classificadores, mas não incluído em nossos resultados experimentais porque os avaliadores humanos sinalizaram o quadro como insuficiente

devido a artefatos de movimento. Neste caso, todos os quatro classificadores corretamente

determinou que o quadro não pôde ser processado.

Figura 7. Uma comparação do desempenho de cada classificador em um conjunto de quadros destaca cenários que podem levar a discrepâncias nas saídas do classificador para várias emoções: HP (feliz), CF (confuso), DG (desgosto), N/A (não aplicável), AG (Irritado), SC (Assustado). Os rótulos do manual da verdade do terreno são mostrados na parte superior, com rótulos derivados de cada classificador na parte inferior.



Limitações

Existem várias limitações associadas a este estudo, que serão abordadas em trabalhos futuros. **Primeiro, analisamos apenas** um subconjunto de classificadores de emoções existentes, enfatizando **os de provedores dos principais serviços em nuvem**. **Esforços futuros estenderão essa avaliação para incluir aqueles que são menos prolíficos e exigem licenças pagas**. Uma **segunda limitação é o uso de diagnósticos relatados pelos pais, que nem sempre podem ser factuais**. Uma **terceira limitação é que, embora tenhamos descartado algumas comorbidades, não descartou todas as comorbidades**, incluindo foi de 63,60% (769/1209) e 55,63% (775/1393), respectivamente, Transtorno de Déficit de Atenção/Hiperatividade, que foi demonstrado para impactar o processamento emocional e a função em crianças [64]. Uma quarta limitação decorre da falta de crianças neurotípicas. Por fim, o conjunto de dados que usamos incluiu uma distribuição desigual de quadros entre as categorias de emoção. No futuro, **investigaremos maneiras de reunir números iguais de quadros e se essa distribuição pode estar relacionada a déficits sociais associados ao autismo ou ao aumento da prevalência de felizes e neutros devido à natureza inerente da jogabilidade**. Embora nossos resultados suportem a conclusão de que os classificadores comerciais de emoção testados aqui

ainda não estão em um nível necessário para uso com crianças autistas, ainda não está claro como esses modelos funcionarão com uma amostra maior, mais diversificada e estratificada.

Conclusões

Neste estudo de viabilidade, avaliamos o desempenho de quatro classificadores de reconhecimento de emoções em crianças com TEA: Google Cloud Vision, Amazon Rekognition, Microsoft Azure Emotion API e Sighthound. A porcentagem média de quadros emotivos e neutros corretamente identificados para todos os classificadores combinados variando muito entre os classificadores com base em como sua sensibilidade e especificidade foram ajustadas. **Os resultados também demonstraram que enquanto a maioria dos classificadores foram capazes de identificar consistentemente felizes quadros, o desempenho para triste, desgosto, e raiva foi ruim: nenhum classificador identificou mais de um terço dos quadros de qualquer uma dessas categorias**. **Concluímos que o desempenho dos classificadores avaliados ainda não está no nível para uso em soluções de terapia móvel e/ou vestível para crianças autistas, sendo necessário o desenvolvimento de conjuntos de dados de treinamento maiores dessas populações para desenvolver modelos mais específicos de domínio**.

Agradecimentos

Este trabalho foi apoiado em parte por fundos para DW do Instituto Nacional de Saúde (1R01EB025025-01 e 1R21HD091500-01), a Fundação Hartwell, Fundação Bill e Melinda Gates, Fundação Coulter, e subsídios de programa do Programa de Inteligência Artificial Centrado no Humano de Stanford, Precision Health and Integrated Diagnostics Center, Beckman Center, Bio-X Center, Predictives and Diagnostics Accelerator Spectrum, Wu Tsai Neurosciences Institute Neuroscience: Translate Program, Stanford Spark e Weston Havens Foundation. Também reconhecemos o generoso apoio de Bobby DeKesyer e Peter Sullivan. O Dr. Haik Kalantarian gostaria de agradecer o apoio do Thrasher Research Fund e do programa Stanford NLM Clinical Data Science (T-15LM007033-35).

Conflitos de interesse

DW é o fundador da Cognoa. Esta empresa está desenvolvendo soluções digitais para a saúde comportamental pediátrica, incluindo condições de neurodesenvolvimento, como autismo, que são detectadas e tratadas usando técnicas, incluindo classificação de emoções. AK trabalha como consultor em tempo parcial para Cognoa. Todos os outros autores declaram não haver interesses conflitantes.

Referências

1. Associação Americana de Psicologia. Manual Diagnóstico e Estatístico de Transtornos Mentais. Quinta edição. Arlington, VA: Pub Psiquiátrico Americano; 2013.
2. Lozier LM, Vanmeter JW, Marsh AA. Prejuízos no reconhecimento de afeto facial associados a transtornos do espectro do autismo: uma meta-análise. *Dev Psychopathol* 2014 Nov;26(4 Pt 1):933-945. [doi:[10.1017/S0954579414000479](https://doi.org/10.1017/S0954579414000479)] [Linha médica:[24915526](https://pubmed.ncbi.nlm.nih.gov/24915526/)]
3. Loth E, Garrido L, Ahmad J, Watson E, Duff A, Duchaine B. Reconhecimento de expressão facial como marcador candidato para transtorno do espectro do autismo: quão frequentes e graves são os déficits? *Mol Autismo* 2018;9:7 [Texto completo GRATUITO] [doi: [10.1186/s13229-018-0187-7](https://doi.org/10.1186/s13229-018-0187-7)] [Linha médica:[29423133](https://pubmed.ncbi.nlm.nih.gov/29423133/)]
4. Fridenson-Hayo S, Berggren S, Lassalle A, Tal S, Pigat D, Bölte S, et al. Reconhecimento básico e complexo de emoções em crianças com autismo: achados transculturais. *Mol Autismo* 2016;7:52 [Texto completo GRATUITO] [doi:[10.1186/s13229-016-0113-9](https://doi.org/10.1186/s13229-016-0113-9)] [Linha médica: [28018573](https://pubmed.ncbi.nlm.nih.gov/28018573/)]
5. Kogan MD, Vladutiu CJ, Schieve LA, Ghandour RM, Blumberg SJ, Zablotsky B, et al. A prevalência de transtorno do espectro do autismo relatado pelos pais entre crianças dos EUA. *Pediatrics* 2018 Dez;142(6) [Texto completo GRATUITO] [doi:[10.1542/peds.2017-4161](https://doi.org/10.1542/peds.2017-4161)] [Linha médica: [30478241](https://pubmed.ncbi.nlm.nih.gov/30478241/)]
6. Rogers SJ. Breve relato: intervenção precoce no autismo. *J Autism Dev Disord* 1996 Abr;26(2):243-246. [doi: [10.1007/bf02172020](https://doi.org/10.1007/bf02172020)] [Linha médica:[8744493](https://pubmed.ncbi.nlm.nih.gov/8744493/)]
7. Cooper JO, Heron TE, Heward WL. Análise de comportamento aplicado. Nova Jersey: Pearson; 2007.
8. Dawson G, Rogers S, Munson J, Smith M, Winter J, Greenson J, et al. Ensaio randomizado e controlado de uma intervenção para crianças com autismo: o Early Start Denver Model. *Pediatrics* 2010 Jan;125(1):e17-e23 [Texto completo GRATUITO] [doi: [10.1542/peds.2009-0958](https://doi.org/10.1542/peds.2009-0958)] [Linha médica:[19948568](https://pubmed.ncbi.nlm.nih.gov/19948568/)]
9. Dawson G. Intervenção comportamental precoce, plasticidade cerebral e prevenção do transtorno do espectro autista. *Dev Psychopathol* 2008;20(3):775-803. [doi:[10.1017/S0954579408000370](https://doi.org/10.1017/S0954579408000370)] [Linha médica:[18606031](https://pubmed.ncbi.nlm.nih.gov/18606031/)]
10. Dawson G, Jones EJ, Merkle K, Venema K, Lowy R, Faja S, et al. A intervenção comportamental precoce está associada à atividade cerebral normalizada em crianças pequenas com autismo. *J Am Acad Psiquiatria Infantil Adolescência* 2012 Nov;51(11):1150-1159 [Texto completo GRATUITO] [doi: [10.1016/j.jaac.2012.08.018](https://doi.org/10.1016/j.jaac.2012.08.018)] [Linha médica:[23101741](https://pubmed.ncbi.nlm.nih.gov/23101741/)]
11. Dawson G, Bernier R. Um quarto de século de progresso na detecção precoce e tratamento do transtorno do espectro do autismo. *Dev Psychopathol* 2013 Nov;25(4 Pt 2):1455-1472. [doi: [10.1017/S0954579413000710](https://doi.org/10.1017/S0954579413000710)] [Linha médica:[24342850](https://pubmed.ncbi.nlm.nih.gov/24342850/)]
12. Scherzer AL, Chhagan M, Kauchali S, Susser E. Perspectiva global sobre diagnóstico precoce e intervenção para crianças com atrasos no desenvolvimento e deficiências. *Dev Med Child Neurol* 2012 Dez;54(12):1079-1084 [Texto completo GRATUITO] [doi: [10.1111/j.1469-8749.2012.04348.x](https://doi.org/10.1111/j.1469-8749.2012.04348.x)] [Linha médica:[22803576](https://pubmed.ncbi.nlm.nih.gov/22803576/)]
13. van Cong T, Weiss B, Toan KN, Le Thu TT, Trang NT, Hoa NT, et al. Serviços de identificação precoce e intervenção para crianças com autismo no Vietnã. *Rep. de Psicologia da Saúde* 2015;3(3):191-200 [Texto completo GRATUITO] [doi: [10.5114/hpr.2015.53125](https://doi.org/10.5114/hpr.2015.53125)] [Linha médica: [27088123](https://pubmed.ncbi.nlm.nih.gov/27088123/)]
14. Samms-Vaughan ME. O status da identificação precoce e intervenção precoce em transtornos do espectro do autismo em países de baixa e média renda. *Int J Speech Lang Pathol* 2014 Fev;16(1):30-35. [doi: [10.3109/17549507.2013.866271](https://doi.org/10.3109/17549507.2013.866271)] [Linha médica: [24397842](https://pubmed.ncbi.nlm.nih.gov/24397842/)]
15. Mandell DS, Novak MM, Zubritsky CD. Fatores associados à idade do diagnóstico em crianças com transtornos do espectro do autismo. *Pediatrics* 2005 Dez;116(6):1480-1486 [Texto completo GRATUITO] [doi: [10.1542/peds.2005-0185](https://doi.org/10.1542/peds.2005-0185)] [Linha médica: [16322174](https://pubmed.ncbi.nlm.nih.gov/16322174/)]
16. Ning M, Daniels J, Schwartz J, Dunlap K, Washington P, Kalantarian H, et al. Identificação e quantificação de lacunas no acesso a recursos de autismo nos Estados Unidos: um estudo infodemiológico. *J Med Internet Res* 2019 10 de julho;21(7):e13094 [Texto completo GRATUITO] [doi: [10.2196/13094](https://doi.org/10.2196/13094)] [Linha médica: [31293243](https://pubmed.ncbi.nlm.nih.gov/31293243/)]
17. Washington P, Wall D, Voss C, Kline A, Haber N, Daniels J, et al. SuperpowerGlass: Uma ajuda vestível para a terapia em casa de crianças com autismo. In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. 2017 Apresentado em: ACM'17; 26 a 30 de junho de 2017; Cambridge, MA. [doi: [10.1145/3130977](https://doi.org/10.1145/3130977)]

18. Washington P, Catalin V, Nick H, Serena T, Jena D, Carl F, et al. Uma ajuda de interação social vestível para crianças com autismo. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems. 2016 Apresentado em: CHI EA'16; 7 a 12 de maio de 2016; Califórnia, EUA pág. 2348-2354. [doi:[10.1145/2851581.2892282](https://doi.org/10.1145/2851581.2892282)]
19. Daniels J, Schwartz J, Haber N, Voss C, Kline A, Fazel A, et al. 5.13 Projeto e eficácia de um dispositivo vestível para aprendizagem social afetiva em crianças com autismo. J Am Acad Child Psy 2017 Out;56(10):S257. [doi:[10.1016/j.jaac.2017.09.296](https://doi.org/10.1016/j.jaac.2017.09.296)]
20. Daniels J, Schwartz JN, Voss C, Haber N, Fazel A, Kline A, et al. Estudo exploratório examinando a viabilidade em casa de uma ferramenta vestível para aprendizagem socioafetiva em crianças com autismo. NPJ Digit Med 2018;1:32 [Texto completo GRATUITO] [doi:[10.1038/s41746-018-0035-3](https://doi.org/10.1038/s41746-018-0035-3)] [Linha médica:[31304314](https://pubmed.ncbi.nlm.nih.gov/31304314/)]
21. Daniels J, Haber N, Voss C, Schwartz J, Tamura S, Fazel A, et al. Teste de viabilidade de uma ajuda comportamental vestível para aprendizagem social em crianças com autismo. Appl Clin Informa 2018 Jan;9(1):129-140 [Texto completo GRATUITO] [doi:[10.1055/s-0038-1626727](https://doi.org/10.1055/s-0038-1626727)] [Linha médica:[29466819](https://pubmed.ncbi.nlm.nih.gov/29466819/)]
22. Voss C, Schwartz J, Daniels J, Kline A, Haber N, Washington P, et al. Efeito da intervenção digital vestível para melhorar a socialização em crianças com transtorno do espectro do autismo: um ensaio clínico randomizado. JAMA Pediatr 2019 1º de maio;173(5):446-454 [Texto completo GRATUITO] [doi:[10.1001/jamapediatrics.2019.0285](https://doi.org/10.1001/jamapediatrics.2019.0285)] [Linha médica:[30907929](https://pubmed.ncbi.nlm.nih.gov/30907929/)]
23. Voss C, Washington P, Haber N, Kline A, Daniels J, Fazel A, et al. Superpower Glass: fornecendo dicas sociais discretas em tempo real em sistemas vestíveis. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2016 Apresentado em: UbiComp'16; 12 a 16 de setembro de 2016; Heidelberg, Alemanha p. 1218-1226. [doi:[10.1145/2968219.2968310](https://doi.org/10.1145/2968219.2968310)]
24. Voss C, Haber N, Wall DP. O potencial de wearables baseados em aprendizado de máquina para melhorar a socialização em adolescentes e adultos com resposta ao transtorno do espectro do autismo. JAMA Pediatr 2019 9 de setembro [Online ahead of print]. [doi:[10.1001/jamapediatrics.2019.2969](https://doi.org/10.1001/jamapediatrics.2019.2969)] [Linha médica:[31498377](https://pubmed.ncbi.nlm.nih.gov/31498377/)]
25. Kline A, Voss C, Washington P, Haber N, Schwartz H, Tariq Q, et al. Superpotência de vidro. GetMobile 2019 14 de novembro;23(2):35-38. [doi:[10.1145/3372300.3372308](https://doi.org/10.1145/3372300.3372308)]
26. Microsoft Azure. Face: um serviço de IA que analisa rostos em imagens URL:<https://azure.microsoft.com/en-us/services/cognitive-services/face/> [acessado em 17/03/2020]
27. Amazon Web Services (AWS) - Serviços de computação em nuvem. URL do Amazon Rekognition:<https://aws.amazon.com/rekognition/> [acessado em 17/06/2018]
28. Google Cloud: Serviços de computação em nuvem. URL da Visão AI:<https://cloud.google.com/vision/> [acessado em 17/06/2018]
29. Sighthound. 17 de junho de 2018. API de detecção e URL da API de reconhecimento:<https://www.sighthound.com/products/cloud> [acessado em 17/06/2018]
30. Cohn J. O Robotics Institute Carnegie Mellon University. 1999. URL do Banco de Dados de Expressão Facial Codificado em AU Cohn-Kanade: <https://www.ri.cmu.edu/project/cohn-kanade-au-coded-facial-expression-database/> [acessado em 24/03/2020]
31. Sneddon I, McRorie M, McKeown G, Hanratty J. O banco de dados de emoções naturais induzidas por Belfast. IEEE Trans Affective Comput 2012 Jan;3(1):32-41. [doi:[10.1109/t-affc.2011.26](https://doi.org/10.1109/t-affc.2011.26)]
32. Kalantarian H, Washington P, Schwartz J, Daniels J, Haber N, Wall DP. Adivinha? J Healthc Informa Res 2018 Out;3(1):43-66. [doi:[10.1007/s41666-018-0034-9](https://doi.org/10.1007/s41666-018-0034-9)]
33. Kalantarian H, Washington P, Schwartz J, Daniels J, Haber N, Wall D. Um Sistema Móvel Gamificado para Crowdsourcing de Vídeo para Pesquisa do Autismo. In: Proceedings of the 2018 IEEE International Conference on Healthcare Informatics. 2018 Apresentado em: ICHI'18; 4 a 7 de junho de 2018; Nova York, NY, EUA p. 350-352. [doi:[10.1109/ichi.2018.00052](https://doi.org/10.1109/ichi.2018.00052)]
34. Kalantarian H, Jedoui K, Washington P, Tariq Q, Dunlap K, Schwartz J, et al. Rotular imagens com emoção facial e o potencial para a saúde pediátrica. Artif Intell Med 2019 julho;98:77-86 [Texto completo GRATUITO] [doi:[10.1016/j.artmed.2019.06.004](https://doi.org/10.1016/j.artmed.2019.06.004)] [Linha médica:[31521254](https://pubmed.ncbi.nlm.nih.gov/31521254/)]
35. Kalantarian H, Jedoui K, Washington P, Wall DP. Um jogo para celular para rotulagem automática de emoções de imagens. IEEE Trans Games 2018;1 epub antes da impressão. [doi:[10.1109/tg.2018.2877325](https://doi.org/10.1109/tg.2018.2877325)]
36. Google Play. Adivinha? URL:<https://play.google.com/store/apps/details?id=walllab.guesswhat> [acessado em 01/06/2018]
37. Loja da Apple. Adivinha? (Laboratório de parede) URL:<https://apps.apple.com/us/app/guess-what-wall-lab/id1426891832> [acessado em 18/03/2020]
38. Picard RW. Tecnologia afetiva futura para autismo e comunicação emocional. Philos Trans R Soc Lond B Biol Sci 2009 Dez 12;364(1535):3575-3584 [Texto completo GRATUITO] [doi:[10.1098/rstb.2009.0143](https://doi.org/10.1098/rstb.2009.0143)] [Linha médica:[19884152](https://pubmed.ncbi.nlm.nih.gov/19884152/)]
39. el Kaliouby R, Picard R, Baron-Cohen S. Computação afetiva e autismo. Ann NY Acad Sci 2006 Dez;1093:228-248. [doi:[10.1196/ana.1382.016](https://doi.org/10.1196/ana.1382.016)] [Linha médica:[17312261](https://pubmed.ncbi.nlm.nih.gov/17312261/)]
40. André D, Pelletier R, Farrington J, Er S, Talbott WA, Phillips PP, et al. Estudioso Semântico. 2006. O Desenvolvimento da Braçadeira SenseWear, Um Dispositivo Revolucionário de Avaliação de Energia para Avaliar Atividade Física e Estilo de Vida URL:<https://www.semanticscholar.org/paper/The-Development-of-the-SenseWear-%C2%AE-armband-%2C-a-to-Andre-Pelletier/e9e115cb6f381a706687982906d45ed28d40bbac> [acessado em 24/03/2020]
41. Tariq Q, Daniels J, Schwartz JN, Washington P, Kalantarian H, Wall DP. Detecção móvel de autismo por meio de aprendizado de máquina em vídeo caseiro: um estudo de desenvolvimento e validação prospectiva. PLoS Med 2018 Nov;15(11):e1002705 [Texto completo GRATUITO] [doi:[10.1371/journal.pmed.1002705](https://doi.org/10.1371/journal.pmed.1002705)] [Linha médica:[30481180](https://pubmed.ncbi.nlm.nih.gov/30481180/)]

42. Tariq Q, Fleming SL, Schwartz JN, Dunlap K, Corbin C, Washington P, et al. Detectando atraso no desenvolvimento e autismo por meio de modelos de aprendizado de máquina usando vídeos caseiros de crianças de Bangladesh: estudo de desenvolvimento e validação. J Med Internet Res 24 de abril de 2019;21(4):e13822 [Texto completo GRATUITO] [doi:10.2196/13822] [Linha médica:31017583]
43. Washington P, Kalantarian H, Tariq Q, Schwartz J, Dunlap K, Chrisman B, et al. Validade da triagem online para autismo: estudo de crowdsourcing comparando tarefas de diagnóstico pagas e não pagas. J Med Internet Res 2019 23 de maio;21(5):e13668 [Texto completo GRATUITO] [doi:10.2196/13668] [Linha médica:31124463]
44. Washington P, Paskov KM, Kalantarian H, Stockham N, Voss C, Kline A, et al. Seleção de recursos e redução de dimensão de dados de autismo social. Pac Symp Biocomput 2020;25:707-718 [Texto completo GRATUITO] [Linha médica:31797640]
45. Abbas H, Garberson F, Glover E, Wall DP. Abordagem de aprendizado de máquina para detecção precoce do autismo, combinando questionário e triagem de vídeo doméstico. J Am Med Inform Assoc 1 de agosto de 2018;25(8):1000-1007. [doi:10.1093/jamia/ocy039] [Linha médica:29741630]
46. Levy S, Duda M, Haber N, Wall DP. Modelos esparsivos de aprendizado de máquina identificam subconjuntos estáveis de recursos preditivos para detecção comportamental de autismo. Mol Autismo 2017;8:65 [Texto completo GRATUITO] [doi:10.1186/s13229-017-0180-6] [Linha médica:29270283]
47. Fusaro VA, Daniels J, Duda M, DeLuca TF, D'Angelo O, Tamburello J, et al. O potencial de acelerar a detecção precoce do autismo por meio da análise de conteúdo de vídeos do YouTube. PLoS One 2014;9(4):e93533 [Texto completo GRATUITO] [doi:10.1371/journal.pone.0093533] [Linha médica:24740236]
48. Sahin NT, Keshav NU, Salisbury JP, Vahabzadeh A. Segunda versão do Google Glass como auxílio socioafetivo vestível: desejo escolar positivo, alta usabilidade e estrutura teórica em uma amostra de crianças com autismo. JMIR Hum Factors 2018 Jan 4;5(1):e1 [Texto completo GRATUITO] [doi:10.2196/fatores humanos.8785] [Linha médica:29301738]
49. Keshav NU, Salisbury JP, Vahabzadeh A, Sahin NT. Smartglasses de coaching de comunicação social: bem tolerados em uma amostra diversificada de crianças e adultos com autismo. JMIR Mhealth Uhealth 2017 21 de setembro;5(9):e140 [Texto completo GRATUITO] [doi:10.2196/mhealth.8534] [Linha médica:28935618]
50. Hernandez J, Picard W. SenseGlass: Usando o Google Glass para sentir emoções diárias. In: Anais da publicação adjunta do 27º simpósio anual da ACM sobre software e tecnologia de interface do usuário. 2014 Apresentado em: UIST'14; 5 a 8 de outubro de 2014; Havaí EUA p. 77-78. [doi:10.1145/2658779.2658784]
51. Scheirer J, Fernandez R, Picard R. Óculos de expressão: um dispositivo vestível para reconhecimento de expressão facial. In: Proceedings of the Extended Abstracts on Human Factors in Computing Systems. 1999 Apresentado em: CHI'99; 15-20 de maio de 1999; Pittsburgh, Pensilvânia p. 262-263. [doi:10.1145/632716.632878]
52. Katsutoshi M, Kunze K, Billingham M. Empathy Glasses. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems. 2016 Apresentado em: CHI EA'16; 7 a 12 de maio de 2016; San Jose, Califórnia p. 1257-1263. [doi:10.1145/2851581.2892370]
53. Escobedo L, Tentori M, Quintana E, Favela J, Garcia-Rosas D. Usando realidade aumentada para ajudar crianças com autismo a manter o foco. IEEE Pervasive Comput 2014 Jan;13(1):38-46. [doi:10.1109/mprv.2014.19]
54. Scassellati B, Admoni H, Matarić M. Robôs para uso na pesquisa do autismo. Annu Rev Biomed Eng 2012;14:275-294. [doi:10.1146/annurev-bioeng-071811-150036] [Linha médica:22577778]
55. Feil-Seifer D, Matarić M. Rumo à robótica socialmente assistida para aumentar as intervenções para crianças com transtornos do espectro do autismo. In: Khatib O, Kumar V, Pappas GJ, editores. Robótica Experimental: O XI Simpósio Internacional. Berlim, Heidelberg: Springer; 2009:201-210.
56. Robins B, Dautenhahn K, Boekhorst RT, Billard A. Assistentes robóticos em terapia e educação de crianças com autismo: um pequeno robô humanóide pode ajudar a estimular as habilidades de interação social? Univ Access Inf Soc 2005 Jul 8;4(2):105-120. [doi:10.1007/s10209-005-0116-3]
57. Escobedo L, Nguyen DH, Boyd LA, Hirano S, Rangel A, Garcia-Rosas D, et al. MOSOCO: Uma ferramenta de assistência móvel para apoiar crianças com autismo praticando habilidades sociais em situações da vida real. In: Anais da Conferência SIGCHI sobre Fatores Humanos em Sistemas Computacionais. 2012 Apresentado em: CHI'12; 5 a 10 de maio de 2012; Austin, EUA pág. 2589-2598. [doi:10.1145/2207676.2208649]
58. Palestra G, Pettinicchio A, Del CM, Carcagni P. Melhor desempenho no reconhecimento de expressões faciais usando 32 recursos geométricos. In: Proceedings of the International Conference on Image Analysis and Processing. 2015 Apresentado em: ICIAP'15; 7 a 11 de setembro de 2015; Génova, Itália p. 518-528. [doi:10.1007/978-3-319-23234-8_48]
59. Leo M, Carcagni P, Distant C, Spagnolo P, Mazzeo P, Rosato A, et al. Avaliação computacional da produção de expressões faciais em crianças com TEA. Sensores (Basileia) 16 de novembro de 2018;18(11) [Texto completo GRATUITO] [doi:10.3390/s18113993] [Linha médica:30453518]
60. Park JH, Abirached B, Zhang Y. Uma estrutura para projetar tecnologias assistivas para ensinar emoções a crianças com TEA. In: Proceedings of the Extended Abstracts on Human Factors in Computing Systems. 2012 Apresentado em: CHI EA'12; 5 a 10 de maio de 2012; Texas, Austin, EUA p. 2423-2428. [doi:10.1145/2212776.2223813]
61. Washington P, Park N, Srivastava P, Voss C, Kline A, Varma M, et al. Diagnóstico baseado em dados e o potencial da inteligência artificial móvel para fenotipagem terapêutica digital em psiquiatria computacional. Biol Psychiatry Cogn Neurosci Neuroimaging 2019 13 de dezembro [Online ahead of print]. [doi:10.1016/j.bpsc.2019.11.015] [Linha médica:32085921]

62. Christensen DL, Baio J, Braun KN, Bilder D, Charles J, Constantino JN, Centros de Controle e Prevenção de Doenças (CDC). Prevalência e características do transtorno do espectro do autismo em crianças de 8 anos - rede de monitoramento de autismo e deficiências de desenvolvimento, 11 sites, Estados Unidos, 2012. MMWR Surveill Summ 2016 Apr 1;65(3):1-23. [doi: [10.15585/mmwr.ss6503a1](https://doi.org/10.15585/mmwr.ss6503a1)] [Linha médica: [27031587](https://pubmed.ncbi.nlm.nih.gov/27031587/)]
63. Ekman P, Friesen WV, O'Sullivan M, Chan A, Diacoyanni-Tarlatzis I, Heider K, et al. Universais e diferenças culturais nos julgamentos de expressões faciais de emoção. J Pers Soc Psychol 1987 Oct;53(4):712-717. [doi: [10.1037//0022-3514.53.4.712](https://doi.org/10.1037//0022-3514.53.4.712)] [Linha médica: [3681648](https://pubmed.ncbi.nlm.nih.gov/3681648/)]
64. Sinzig J, Morsch D, Lehmkuhl G. Hiperatividade, impulsividade e desatenção têm impacto na capacidade de reconhecimento de afeto facial em crianças com autismo e TDAH? Eur Psiquiatria Infantil Adolescência 2008 Mar;17(2):63-72. [doi: [10.1007/s00787-007-0637-9](https://doi.org/10.1007/s00787-007-0637-9)] [Linha médica: [17896119](https://pubmed.ncbi.nlm.nih.gov/17896119/)]

Abreviaturas

ABA:Análise de comportamento

aplicado **IA:**inteligência artificial

API:interface de programação de aplicativos

TEA:transtorno do espectro do autismo **ESDM:**

modelo de Denver de início precoce **SH:**Galgo

Editado por G Eysenbach; apresentado em 18.12.18; revisado por M Leo, N Sahin; comentários ao autor 27.04.19; versão revisada recebida em 03.07.19; aceito em 23.02.20; publicado em 01.04.20

Por favor, cite como:

Kalantarian H, Jedoui K, Dunlap K, Schwartz J, Washington P, Husic A, Tariq Q, Ning M, Kline A, Wall DP O desempenho de classificadores de emoções para crianças com autismo relatado pelos pais: estudo quantitativo de viabilidade JMIR Ment Health 2020; 7(4):e13174

URL:<https://mental.jmir.org/2020/4/e13174>

doi:[10.2196/13174](https://doi.org/10.2196/13174)

PMID:[32234701](https://pubmed.ncbi.nlm.nih.gov/32234701/)

© Haik Kalantarian, Khaled Jedoui, Kaitlyn Dunlap, Jessey Schwartz, Peter Washington, Arman Husic, Qandeel Tariq, Michael Ning, Aaron Kline, Dennis Paul Wall. Originalmente publicado em JMIR Mental Health (<http://mental.jmir.org>), 01.04.2020. Este é um artigo de acesso aberto distribuído sob os termos da Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o original trabalho, publicado pela primeira vez no JMIR Mental Health, é devidamente citado. Devem ser incluídas as informações bibliográficas completas, um link para a publicação original em <http://mental.jmir.org/>, bem como esta informação de copyright e licença.