



**UNIVERSIDADE FEDERAL DA FRONTEIRA SUL
CAMPUS DE CHAPECÓ
CURSO DE CIÊNCIA DA COMPUTAÇÃO**

RICHARD HENRIQUE HERRERA SILVA

**ANÁLISE FACIAL PARA DETECÇÃO DE ESTRESSE E TÊDIO EM JOGOS
DIGITAIS**

**CHAPECÓ
2022**

RICHARD HENRIQUE HERRERA SILVA

**ANÁLISE FACIAL PARA DETECÇÃO DE ESTRESSE E TÉDIO EM JOGOS
DIGITAIS**

Trabalho de conclusão de curso apresentado como requisito para obtenção do grau de Bacharel em Ciência da Computação da Universidade Federal da Fronteira Sul.
Orientador: Prof. Dr. Fernando Bevilacqua

CHAPECÓ
2022

Silva, Richard Henrique Herrera

Análise facial para detecção de estresse e tédio em jogos digitais /
Richard Henrique Herrera Silva. – 2022.

31 f.: il.

Orientador: Prof. Dr. Fernando Bevilacqua.

Trabalho de conclusão de curso (graduação) – Universidade Federal
da Fronteira Sul, curso de Ciência da Computação, Chapecó, SC, 2022.

1. Interação humano-computador. 2. Jogos. 3. Computação afeti-
tiva. 4. Visão Computacional. 5. *MediaPipe*. I. Bevilacqua, Prof.
Dr. Fernando, orientador. II. Universidade Federal da Fronteira Sul.
III. Título.

© 2022

Todos os direitos autorais reservados a Richard Henrique Herrera Silva. A reprodução de partes
ou do todo deste trabalho só poderá ser feita mediante a citação da fonte.

E-mail: henriqueherrera9@gmail.com

RICHARD HENRIQUE HERRERA SILVA

**ANÁLISE FACIAL PARA DETECÇÃO DE ESTRESSE E TÉDIO EM JOGOS
DIGITAIS**

Trabalho de conclusão de curso apresentado como requisito para obtenção do grau de Bacharel em Ciência da Computação da Universidade Federal da Fronteira Sul.

Orientador: Prof. Dr. Fernando Bevilacqua

Este trabalho de conclusão de curso foi defendido e aprovado pela banca avaliadora em: 11/4/2022.

BANCA AVALIADORA

A handwritten signature in black ink, appearing to read 'Bevilacqua', enclosed within a large, loopy oval shape. The signature is centered on a light gray rectangular background.

Prof. Dr. Fernando Bevilacqua – UFFS

Prof. Dr. Denio Duarte – UFFS

Profa. Me. Andressa Sebben – UFFS

AGRADECIMENTOS

Agradeço a minha namorada por todo apoio que sempre me dedicou. Sou grato a minha mãe por sempre me incentivar e acreditar que eu seria capaz de superar os obstáculos que a vida me apresentou. A todos os meus amigos do curso de graduação que compartilharam dos inúmeros desafios que enfrentamos, sempre com o espírito colaborativo, e um agradecimento especial ao Comassetto e ao Jean que sempre me ajudaram com suas vasta experiências desde o início da graduação. Também quero agradecer à Universidade Federal da Fronteira Sul e o seu corpo docente que demonstrou estar comprometido com a qualidade e excelência do ensino.

RESUMO

A classificação de expressões faciais usando visão computacional, a fim de obter uma extração remota dos sinais fisiológicos do indivíduo, geralmente relaciona-se a partir de métodos que estimulam o usuário a ter um envolvimento emocional ou esboçar alguma reação fisiológica, como, por exemplo, usando vídeos e imagens. Geralmente, os modelos usados são construídos para identificar uma grande parte de comportamentos emocionais de um determinado grupo, o que faz com que as individualidades comportamentais de cada pessoa sejam desconsideradas. No decorrer deste trabalho apresentamos um método em que seja possível analisar e classificar os estados emocionais de cada indivíduo de forma remota, treinando uma rede neural que gera um modelo exclusivo para cada um dos usuários, por meio de jogos de calibração. Com isso será possível analisar e classificar o estado emocional de um jogador, mantendo as suas peculiaridades comportamentais.

Palavras-chave: Interação humano-computador. Jogos. Computação afetiva. Visão Computacional. *MediaPipe*.

ABSTRACT

The classification of facial expressions using computer vision, in order to obtain a remote extraction of the individual's physiological signals, is usually related to methods that stimulate the user to have an emotional involvement or to sketch some physiological reaction, such as using, for example, videos and images. Generally, the models used are built to identify a large part of the emotional behaviors of a given group, which causes the behavioral individualities of each person to be disregarded. In the course of this work, we present a method where it is possible to analyze and classify the emotional states of each individual remotely, training a neural network that generates an exclusive model for each of the users, through calibration games. With this, it will be possible to analyze and classify the emotional state of a player, maintaining their behavioral peculiarities.

Keywords: Human-computer interaction. Games. Affective computing. Computer Vision. *MediaPipe*.

LISTA DE ILUSTRAÇÕES

Figura 1 – Ilustração do <i>workflow</i> da inferência de emoções.	10
Figura 2 – O modelo Circumplex de Afeto (MCA) proposto por Russel.	11
Figura 3 – Representação das 6 categorias de emoções básicas, sendo elas respectivamente, felicidade, medo, surpresa, nojo, tristeza e raiva, apresentadas por Ekman P. (5).	16
Figura 4 – Padrões de movimento produzidos por cada uma das seis emoções investigadas por JN. (11)	17
Figura 5 – Detecção facial realizada pela ferramenta <i>MediaPipe</i>	19
Figura 6 – Expressões de emoções mescladas (surpresa e felicidade).	20
Figura 7 – Ilustração do <i>workflow</i> proposto para extração remota de sinais. Fonte: Adaptado de Bevilacqua; Engström; Backlund (3)	21
Figura 8 – Estrutura de um jogo de calibração. eixo-x mostra a progressão do tempo. O eixo-y mostra a dificuldade do jogo e o estado emocional que os usuários devem vivenciar, ou seja, estresse e tédio. Fonte: Adaptado de Bevilacqua; Engström; Backlund (3)	22
Figura 9 – Ilustração das etapas da metodologia da pesquisa. Fonte: Adaptado de Giannakakis et al. (6)	23
Figura 10 – Ilustração das etapas da construção do modelo proposto. Fonte: Adaptado de Giannakakis; Marias; Tsiknakis (7)	24
Figura 11 – Técnica de validação cruzada. Fonte: Elaborada pelo autor	24
Figura 12 – Ilustração das etapas da pesquisa. Fonte: Elaborada pelo autor	26
Figura 13 – Ilustração dos resultados obtidos. Fonte: elaborado pelo autor	28

SUMÁRIO

1	INTRODUÇÃO	9
1.1	APRESENTAÇÃO	11
1.2	PROBLEMÁTICA	12
2	OBJETIVOS	14
2.1	OBJETIVOS GERAIS	14
2.2	OBJETIVOS ESPECÍFICOS	14
3	JUSTIFICATIVA	15
4	REVISÃO BIBLIOGRÁFICA	16
4.1	RECONHECIMENTO DE EMOÇÕES HUMANAS E EXPRESSÕES FACIAIS	16
4.2	<i>MEDIAPIPE</i>	18
4.3	TRABALHOS RELACIONADOS	19
4.3.1	Estado da arte na análise automática de expressão facial	19
4.3.2	Detecção remota de estresse e tédio em jogos de calibração personalizada por usuário	20
4.3.3	Detecção de estresse e ansiedade usando pistas faciais de vídeos	22
4.3.4	Um sistema de reconhecimento de estresse usando parâmetros HRV e técnicas de aprendizado de máquina	23
5	METODOLOGIA	25
5.1	DESENVOLVIMENTO USANDO MEDIAPIPE	25
5.2	PREPARAÇÃO DOS MATERIAIS, PROCESSAMENTO DOS VÍDEOS E EXTRAÇÃO DOS DADOS	25
5.3	REDE NEURAL UTILIZADA	26
5.4	EXECUÇÃO E APURAÇÃO DOS RESULTADOS	27
6	RESULTADOS	28
6.1	CONFIGURAÇÕES DO EXPERIMENTO	28
6.2	OBSERVAÇÕES DOS TESTES	28
7	CONCLUSÃO	29
	REFERÊNCIAS	30

1 INTRODUÇÃO

Expressões faciais podem ser consideradas tanto respostas emocionais quanto ferramentas utilizadas para a comunicação social. Nesse sentido, elas são responsáveis por gerar informações a partir das quais pode-se inferir rapidamente a resposta emocional do outro, possibilitando a compreensão e a adequação do comportamento frente à resposta facial expressa (1). Além disso, expressões faciais são socialmente aprendidas, por isso também caracterizam-se por serem culturalmente variáveis. Sendo assim, não há uma forma fixa e geral da expressão e seu significado. Com expressões faciais também é possível classificar padrões que detectam estados emocionais de um usuário. Na computação afetiva, algo bastante usado atualmente são sensores físicos, que permitem extrair informações psicofisiológicas de um usuário, assim aumentando a fonte de informações extraídas a partir do comportamento físico do usuário. A combinação dessas informações leva ao aumento considerável do grau de acuracidade dos estudos realizados envolvendo detecção de emoções.

O uso de sensores físicos, porém, acaba obstruindo a coleta dos dados, afeta o comportamento natural do usuário. Na área de jogos as pesquisas realizadas buscam preservar ao máximo possível a imersão durante um jogo, priorizando emoções genuínas de jogadores. Por conta disso, sensores físicos acabam não sendo a melhor opção. Na literatura apresenta-se uma possível solução para este problema, que é o uso do sensoriamento remoto para a obtenção dos sinais psicofisiológicos dos usuários (2).

Com o grande avanço da tecnologia nos últimos tempos, a visão computacional foi uma das áreas que mais se beneficiou das crescentes evoluções da inteligência artificial. Hoje é possível fazer uma análise detalhada da atividade muscular da face de um indivíduo, permitindo a realização da análise facial de forma automatizada o suficiente para que seja possível detectar emoções de jogadores. Em termos de jogos, na maioria dos casos o foco principal é a detecção de estresse e tédio. Um método de análise de dados muito usado para essa vertente é o aprendizado de máquina, que viabiliza a automatização da construção de modelos analíticos. Isso baseado em um sistema que aprende com os dados e identifica padrões, possibilitando a tomada de decisões com o mínimo possível de interação humana.

Na maioria das pesquisas realizadas, os modelos gerados usando aprendizado de máquina são construídos de formas generalistas. Nesse contexto, um determinado grupo de pessoas é usado para a extração dos dados e depois é feito o treinamento do modelo. O problema que se tem é que, por conta de ser um modelo geral, as individualidades e peculiaridades de cada indivíduo desaparecem. Tentando evitar isso, existem métodos que constrói um modelo de análise facial único para cada um dos jogadores. Isso é importante para garantir que individualidades não sejam perdidas.

O principal foco deste é o aprimoramento do método de extração de micro expressões faciais com foco em detectar estados emocionais de estresse e tédio dos jogadores (2), sendo o objetivo fazer uso de métodos remotos que permitam realizar a classificação das ações e emoções

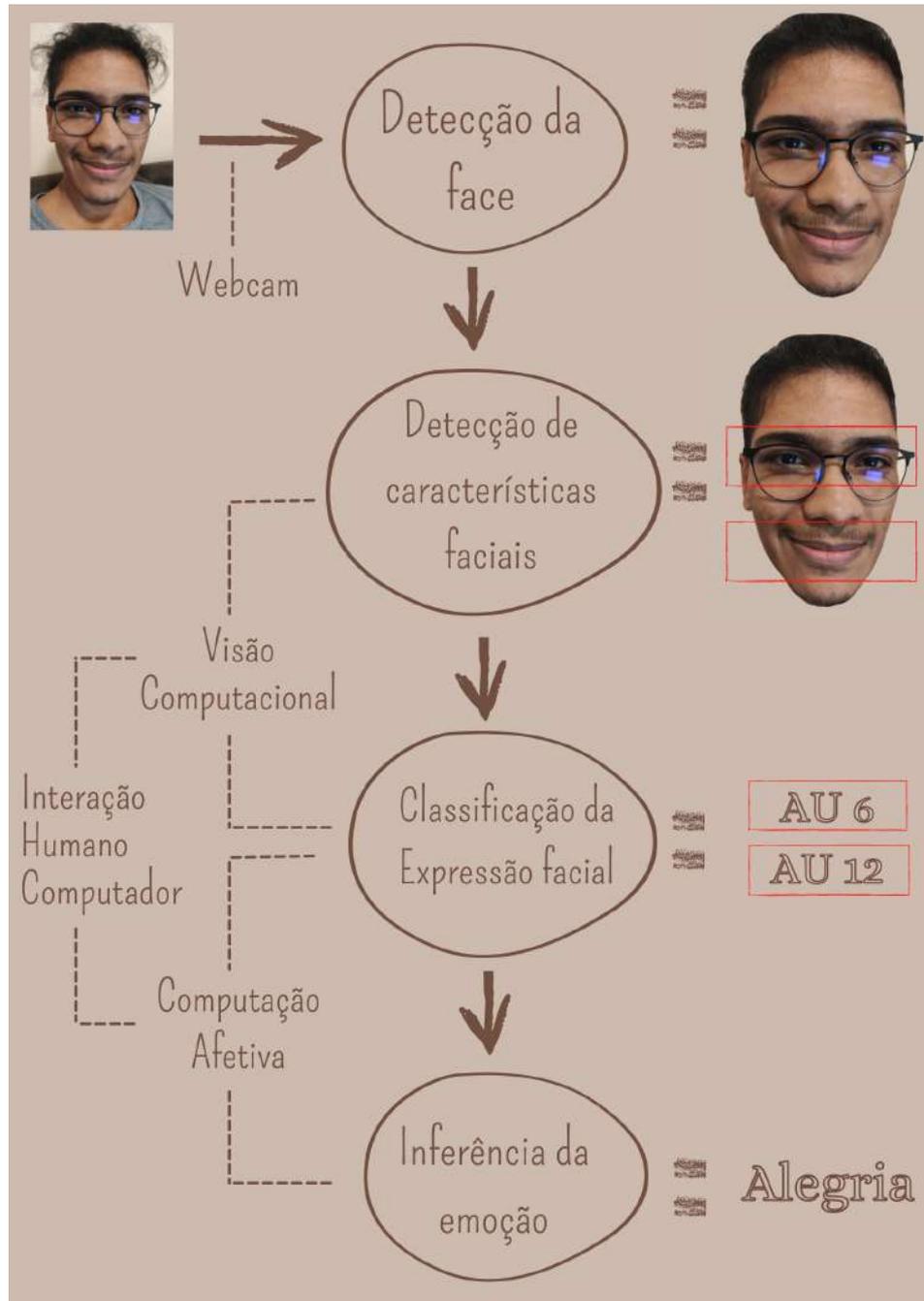


Figura 1 – Ilustração do *workflow* da inferência de emoções.

do usuário durante as sessões de jogos, conforme o exemplo na (Figura 1). A importância da detecção de estresse e tédio está ligada ao grau de imersão de um jogador, conforme é apresentado na teoria do fluxo (13). No mercado, a detecção de estresse e tédio pode ser usada com o intuito de analisar os estados emocionais do usuário durante o jogo, já que as emoções são de grande importância no ciclo de desafio e recompensa, o que torna um jogo interessante e imersivo. A análise das emoções permite ao desenvolvedor do jogo otimizar ao máximo possível o jogo desenvolvido, o que leva a maiores chances de ganhos financeiros. Essa tecnologia também possibilita ao desenvolvedor novos caminhos para desenvolver jogos que são capazes de detectar níveis de estresse e tédio sem que seja preciso ter um *hardware* ou sensores adicionais. Isso

permite o desenvolvimento de jogos que conseguem por si só autoajustar o nível de dificuldade durante a sua execução, permitindo melhorar ao máximo a experiência do usuário com o jogo.

1.1 APRESENTAÇÃO

A modelagem de emoções é um tema recorrente na literatura. Frequentemente utilizado como base para pesquisas relacionadas às emoções, o Modelo Circumplex de Afeto (MCA), mostrado na (Figura 2), propõe que todas as emoções emergem de interpretações cognitivas das sensações neurais centrais. Essas acontecem por conta de uma combinação linear de dois sistemas neurofisiológicos independentes: um relacionado à valência (um *continuum* de prazer-desprazer) e o outro à excitação (15). A partir deles, pode-se ter vários graus de combinações entre a valência e a excitação. O eixo vertical representa o nível de excitação (relaxado ou excitado) e o eixo horizontal indica o estado emocional (positivo ou negativo). Com base no modelo, entende-se que as emoções mudam conforme o seu nível de ativação. Aponta-se que os padrões de emoções são produtos da ativação dentro desses dois sistemas neurofisiológicos. A alegria, por exemplo, é classificada no MCA no quadrante superior direito, sendo uma emoção em que se tem uma grande carga de ativação entre os sistemas neurais associados, gerando uma valência positiva que se associa a uma intensificação moderada nos sistemas neurais relacionados à excitação.

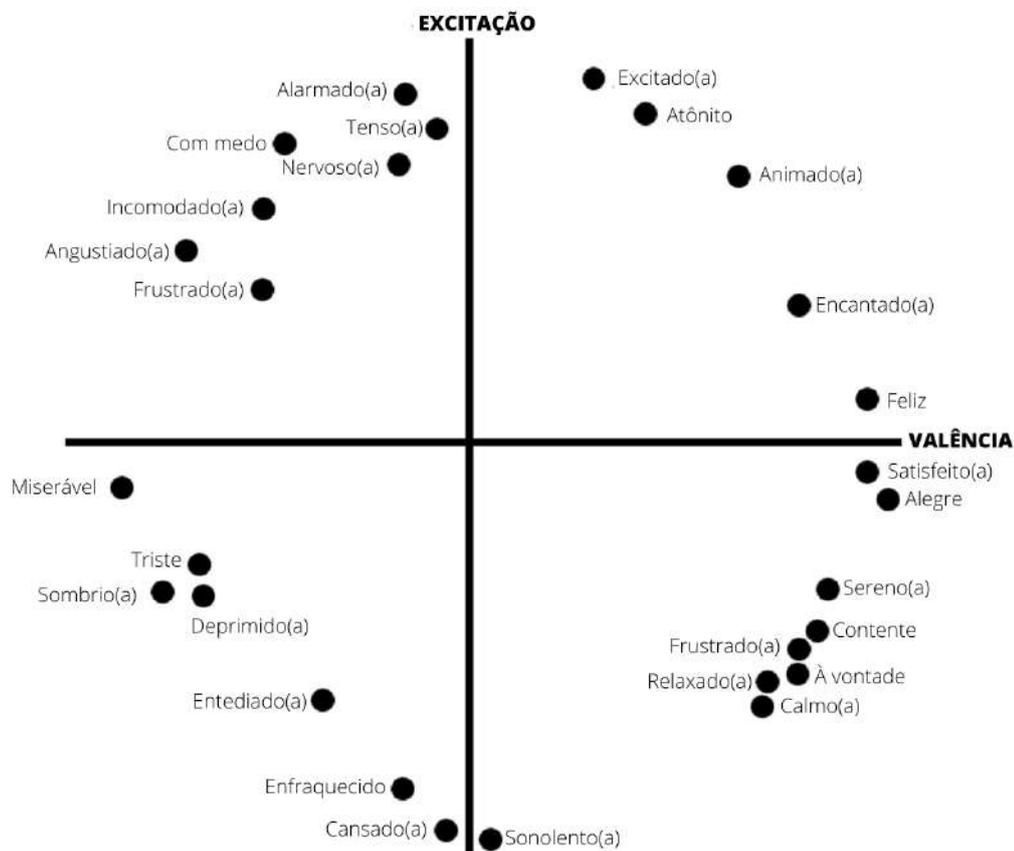


Figura 2 – O modelo Circumplex de Afeto (MCA) proposto por Russel.

Por outro lado, a emoção da tristeza encontra-se no quadrante inferior esquerdo, sugerindo que apesar de ela apresentar moderada excitação, sua valência apresenta-se como a menor carga de ativação dentre as emoções ilustradas (15). O tédio, uma das emoções foco do estudo deste projeto, pode ser definida como uma experiência negativa de ter o seu desejo de se engajar em uma atividade satisfatória não realizada. Sendo assim, também pode ser encontrada no quadrante inferior esquerdo, apresentando uma valência e excitação baixas, caracterizando-se por um estado de passivo negativo (16).

Parcialmente contrário a isso, SA. et al. (16) reconhece que uma pessoa entediada pode exibir tanto uma excitação baixa, quanto uma alta, dependendo do tipo de tédio apresentado. Nesse sentido, o tédio pode ser dividido em duas categorias: o tédio traço e o tédio de estado. O tédio traço aponta para a tendência de um indivíduo ficar entediado, normalmente associado à personalidade. O tédio de estado, entretanto, apresenta-se como uma emoção momentânea e breve, normalmente relacionado ao exterior da pessoa. Em vista disso, o tédio traço pode ser o motivo do interrompimento de uma atividade, como o jogo, impedindo a manutenção da concentração do jogador. Dessa forma, o tédio foco do presente projeto, o tédio traço, mostra-se como um sentimento aversivo à experiência de fluxo, a qual possibilita a imersão satisfatória do jogador.

Essas emoções podem ser peças fundamentais no desenvolvimento de jogos. Nesse sentido, diversas iniciativas exploram a percepção da emoção em um jogo a fim de preservar e estimular o envolvimento do jogador. Chanel et al. (4) apresenta o estado emocional do usuário utilizando um esquema de categorização composto por três classes: tédio, ansiedade e engajamento. Para manter o engajamento do jogador, eles propuseram um sistema de adequação ao seu estado emocional, isto é, a dificuldade do jogo seria automaticamente reajustada quando fosse detectado tédio ou ansiedade no jogador.

1.2 PROBLEMÁTICA

Atualmente grande parte dos estudos realizados que envolvem a classificação de emoções faz uso das medidas psicofisiológicas. Em sua grande maioria, essas precisam de sensores físicos colocados aos indivíduos para que se consiga extrair uma grande quantidade de informações psicofisiológicas, como as ações faciais, frequência respiratória, frequência cardíaca, entre outros. Esses sensores físicos fornecem aos pesquisadores uma grande quantidade de informações sobre o voluntário da pesquisa, com isso é possível fazer uma análise de diferentes informações extraídas e obter uma chance maior de conseguir resultados mais precisos. Entretanto, na maioria das vezes esses métodos intrusivos de estudo acabam afetando o comportamento natural do indivíduo, ainda mais na área dos jogos. Isso porque para uma classificação emocional de estresse e tédio precisa, é necessário que o jogador esteja o máximo possível imerso no jogo. Com o uso dos sensores físicos, os jogadores têm suas ações e imersão ao jogo afetadas, prejudicando o resultado das pesquisas realizadas. Com isso em mente, este trabalho explora o

desenvolvimento de um método no qual seja possível a realização da classificação remota dos estados emocionais de um indivíduo de forma não intrusiva e com uma precisão confiável (3).

2 OBJETIVOS

2.1 OBJETIVOS GERAIS

Aperfeiçoar o método de análise não invasivo de detecção de estresse e tédio proposto por Bevilacqua; Engström; Backlund (2), voltado para jogos e baseado em análise facial via Visão Computacional, implementado como um modelo capaz de classificar os estados emocionais de um jogador de forma individual.

2.2 OBJETIVOS ESPECÍFICOS

- Apresentar os modelos de Visão computacional voltados à classificação de expressões faciais e detecção de emoções;
- Elencar teorias relacionadas ao estado de emocional de um indivíduo para a criação de um modelo para detecção de emoções faciais;
- Explorar a utilização da biblioteca *MediaPipe* no contexto de análise facial e detecção de emoções na área de jogos digitais, mostrando suas principais características;
- Avaliar de forma experimental a acurácia das implementações realizadas.

3 JUSTIFICATIVA

A utilização de Visão computacional aplicada à detecção de emoções é uma área ampla. Ao analisar as lacunas ainda existentes no uso da Visão computacional para área de estudos envolvendo a classificação de emoções, nota-se a realização de estudos com ferramentas que são intrusivas. Nota-se, também, a viabilidade de elaboração de um método de pesquisa com ênfase no treinamento de modelos para classificação de expressões faciais que seja único para cada usuário. Nesse contexto, em particular, há pouca exploração voltada para área de jogos.

A pesquisa proposta neste projeto tem por foco, portanto, explorar um caminho alternativo para se realizar estudos envolvendo estados emocionais. Isso está fundamentado na carência significativa nesta vertente, pois a maioria dos estudos realizados envolvem modelos para classificação de um determinado grupo de pessoas ou fazem uso de sensores físicos. Conforme descrito anteriormente, sensores acabam por limitar o comportamento natural do usuário, enquanto treinamento em grupo remove individualidades.

Como um incentivo a mais que ratifica a importância do projeto, observou-se que este é pioneiro em abordar métodos de treinamento de modelos únicos para cada usuário, que exaltam as características e comportamentos singulares de cada pessoa. Isso tem o potencial de possibilitar uma classificação emocional confiável e também apresentar para área de jogos estudos de como fazer uso dessa ferramenta.

4 REVISÃO BIBLIOGRÁFICA

4.1 RECONHECIMENTO DE EMOÇÕES HUMANAS E EXPRESSÕES FACIAIS

Para a identificação do estado emocional é necessário a interpretação das expressões faciais. Comumente mencionado na literatura, há o Sistema de Codificação de Ação Facial (FACS), criado por Ekman e Friesen (5). Esse modelo analisa o padrão das diferentes categorias de expressões faciais, por meio da identificação da ativação muscular nas expressões faciais mediante Unidades de Ação (AUs). Essas AUs configuram-se como movimentos musculares visualmente discrimináveis na face. Como resultado, essas AUs constituem uma base de informações desses movimentos, atuando semelhantemente a um alfabeto. A partir disso, Ekman e Friesen identificaram seis categorias de emoções básicas: medo, felicidade, tristeza, nojo, raiva e surpresa, conforme a (Figura 3), as quais seguem um padrão similar independente das origens e culturas Ekman P. (5).

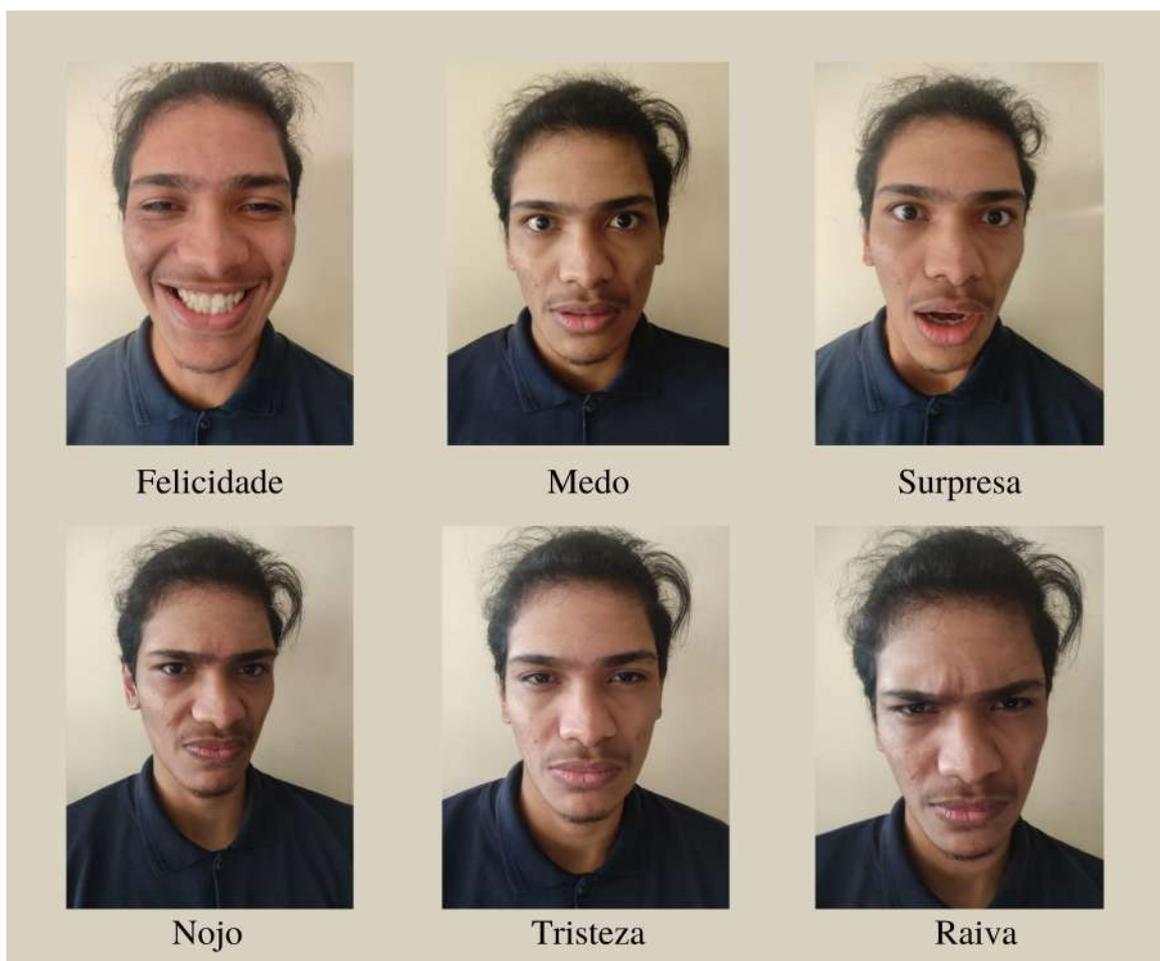


Figura 3 – Representação das 6 categorias de emoções básicas, sendo elas respectivamente, felicidade, medo, surpresa, nojo, tristeza e raiva, apresentadas por Ekman P. (5).

Em 1979, foi realizada uma pesquisa com 20 voluntários, sendo 10 mulheres e 10 homens, a fim de relacionar o movimento facial no reconhecimento de emoções. Os voluntários passaram

por treinamento, através de fotos, para o reconhecimento de algumas emoções: felicidade, tristeza, medo, surpresa, raiva e nojo. Posteriormente, alguns voluntários tiveram seu rosto coberto de maquiagem preta, com cerca de 50 manchas brancas, de 8 mm de diâmetro cada, distribuídas aleatoriamente em suas faces. Em seguida, realizou-se uma gravação de seus rostos pintados, demonstrando as seis emoções, conforme a (Figura 4). Após análise, constatou-se que telas de rosto inteiro foram reconhecidas com mais precisão do que nas exibições da parte inferior do rosto, que por sua vez foram mais reconhecidas do que as exibições da parte superior da face (11).

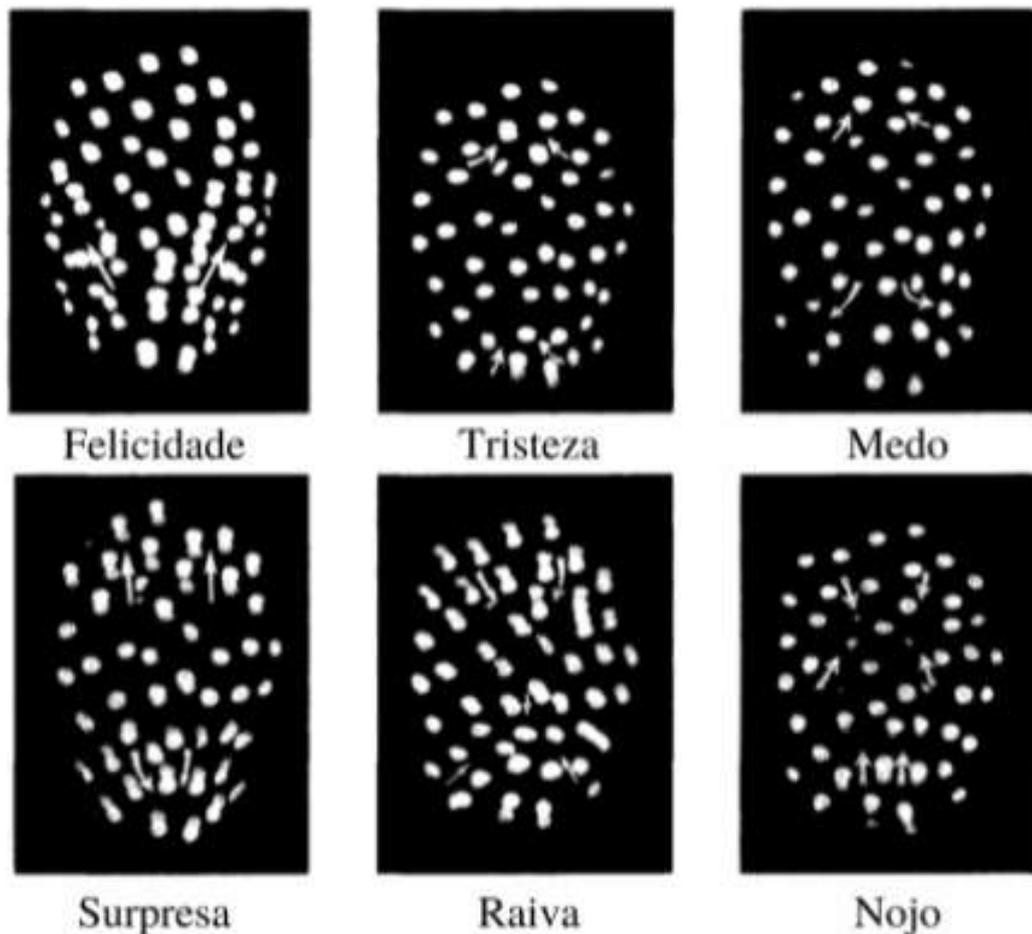


Figura 4 – Padrões de movimento produzidos por cada uma das seis emoções investigadas por JN. (11)

O padrão de movimento da felicidade é produzido a partir do sorriso, caracterizado pelo movimento ascendente nos cantos da boca e das bochechas. A tristeza consiste em um ligeiro e sutil deslocamento para cima na área do queixo, enquanto a área da testa reflete um movimento ascendente voltando-se para o interior. Já o medo, na região perioral, apresenta um padrão voltado para fora e descendente, enquanto que, na parte superior da face, o medo assemelha-se à tristeza, diferenciando-se por um movimento mais pronunciado e arqueado para cima. A expressão de surpresa envolve um deslocamento pronunciado para cima das sobrancelhas e

um deslocamento descendente na mandíbula. Na raiva, há uma movimentação descendente na parte superior da face, concomitantemente com uma compressão labial, na região perioral, caracterizada pelas movimentações que voltam-se para a região da boca. Por fim, a expressão de nojo é caracterizada pelo enrugamento do nariz, o que causa um aumento do movimento na região perinasal e nas bochechas, além disso, a expressão ainda pode envolver o movimento ascendente na área do queixo (11).

Para além desses, a expressão facial também pode combinar esses padrões e refletir o estresse. Dessa forma, o estresse emocional negativo pode ser observado a partir da combinação de três expressões faciais: raiva, tristeza e medo. Assim, ao identificar esse padrão de expressões faciais, pode-se prever o estresse (10). No entanto, o estresse expressa-se como um objeto de identificação complexa, tendo em vista a ampla variedade de estressores existentes, bem como seus sinais manifestados. Concomitante a isso, os humanos, como seres singulares, também expressam mecanismos de defesa a fim de evitar situações constrangedoras que podem ser desencadeadas por esse sentimento, dessa forma, dificultando sua identificação (8).

Com relação à área de pesquisa envolvendo jogos, uma das teorias mais importantes relacionada aos estados emocionais é a teoria do fluxo (13). Essa teoria é usada como base para várias pesquisas e entendimentos sobre as emoções e comportamentos humanos, englobando engajamento e imersão, senso de presença e aplicabilidade no *design* de jogos. A teoria do fluxo propõe que um indivíduo precisa estar se sentido desafiado e, simultaneamente a isso, precisa estar fazendo uso de suas competências máximas para executar uma atividade com um alto nível de performance. Para isso acontecer é necessário que a pessoa esteja experimentando um nível de imersão intenso. A fundamentação essencial do fluxo é qualificada pela imersão e foco completo pela atividade que está sendo exercida (13). Na área dos jogos tem-se que uma verticalidade entre desafio e habilidade em um jogo faz com que os jogadores atinjam um estado sublime de experiência e foco. Com isso, o jogador passa a desfrutar muito mais do que o jogo tem a oferecer, sem que fique estressado ou com tédio. Com base nisso, é correto afirmar que as construções de fluxo são essenciais durante o desenvolvimento de um jogo (2).

4.2 *MEDIAPIPE*

O *MediaPipe* é um *framework Open-Source* criado para o desenvolvimento de *pipelines* de aprendizado de máquina, inclusive análise facial. O *framework* é multiplataforma e fornece diversos recursos, sendo alguns deles, detecção facial, mapeamento da mão, detecção de gestos, detecção de objetos e detecção da íris do olho. Apesar de ter como foco os *smartphones*, também possui suporte para outras plataformas como navegadores e aplicações *desktop*.

A solução usada para a detecção facial é desenvolvida para ter um desempenho eficiente em GPU móvel. Isso possibilita um alto desempenho e com isso é possível realizar a detecção facial em tempo real. Existem também outros modelos com finalidades mais específicas, como o de detecção de micro expressões faciais. Um exemplo do detector facial sendo executado em

tempo real pode ser visto na (Figura 5).

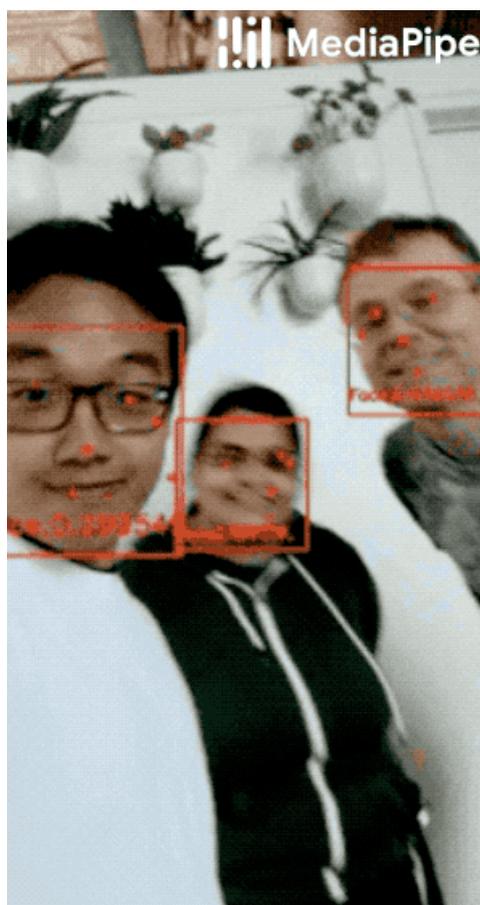


Figura 5 – Detecção facial realizada pela ferramenta *MediaPipe*.

Além de ser uma ferramenta gratuita mantida pelo *Google*, o *MediaPipe* destaca-se não apenas por permitir a detecção das características da face humana. Ela também apresenta uma diversidade de funcionalidades disponibilizadas para desenvolvimento. Essa ferramenta vem sendo usada em diversos projetos e por ser multiplataforma possibilita ao desenvolvedor trabalhar em linguagens diferentes.

4.3 TRABALHOS RELACIONADOS

4.3.1 Estado da arte na análise automática de expressão facial

Pantic; Rothkrantz (14) fazem uma análise sobre a visão computacional, comportamento humano e análise automática de expressão facial. Inicialmente é traçada uma perspectiva geral em relação à comunicação humana e comportamento humano, segmentado em duas categorias: canais de comunicação e modalidade. Os canais de comunicação seria a forma com que um humano se comunica com outro, por exemplo, canais auditivos e canais visuais. A modalidade seria a forma com que o humano faz uso dos seus sentidos para extrair informação do mundo exterior (14).

Em uma interação social entre duas ou mais pessoas, diversos canais são usados e diferentes modalidades são ativadas. Isso torna a comunicação bem mais fluida e clara para os interlocutores. Com base nisso, a fim de estudar a interação homem-máquina e torná-la mais eficaz, estudos com essa temática são contextualizados sobre a área conhecida como *Human-Computer Interaction* (HCI). Um dos assuntos de estudo da HCI é a face humana, tendo em vista o seu papel fundamental na comunicação interpessoal. A face humana auxilia no reconhecimento dos iguais, bem como na identificação de padrões de comportamentos, fases de vida e as emoções conforme a (Figura 6). Além disso, o ser humano tende a se comunicar de forma mais eficiente quando apresenta-se em contato face a face; assim, pode-se perceber as expressões faciais como uma peça fundamental na comunicação humana (14).



Figura 6 – Expressões de emoções mescladas (surpresa e felicidade).

Dada a importância da detecção de expressões faciais, houve vários avanços científicos que tornaram possível a criação de modelos que conseguem automatizar a identificação e classificação de sinais faciais emocionais e conversacionais. Automatizar a análise da expressão facial ajuda a aproximar a interação homem-máquina, e também possibilita o desenvolvimento de ferramentas em pesquisas envolvendo ciências comportamentais e medicina. Pantic; Rothkrantz (14) buscam trazer quais os três principais problemas que a automatização de detecção de expressões faciais enfrenta. Eles mostram quais as principais técnicas apresentadas para resolver esses problemas, bem como resumem suas características com base em requisitos impostos ao *design* de um analisador de expressão facial ideal. Isso traça os possíveis caminhos a se seguir para se aproximar cada vez mais de um analisador de expressão facial definido como o analisador ideal (14).

4.3.2 Detecção remota de estresse e tédio em jogos de calibração personalizada por usuário

O objetivo do artigo de Bevilacqua; Engström; Backlund (3), é apresentar a detecção do estado emocional usando Visão computacional e a detecção remota dos sinais demonstrados

pelos usuários de uma forma não intrusiva. A contribuição principal é o desenvolvimento de um método capaz de identificar e classificar estados emocionais de estresse e tédio dos jogadores durante um turno de jogo fazendo uso da extração remota de sinais via Visão computacional, conforme a (Figura 7). Os pesquisadores ressaltam que, na maioria das vezes, as pesquisas realizadas para se desenvolver modelos preditivos para classificações de emoções impõem situações limitadas ao usuário, o qual não possui um grande envolvimento emocional e também liberdade para interagir, como durante a manipulação de imagens e vídeos (3).



Figura 7 – Ilustração do *workflow* proposto para extração remota de sinais. Fonte: Adaptado de Bevilacqua; Engström; Backlund (3)

Relacionado a isso também tem-se outro problema, no qual os modelos preditivos desenvolvidos muitas vezes abrangem apenas um determinado grupo de pessoas, que acaba por limitar a acurácia da classificação de emoções e comportamentos a outros grupos. Isso faz com que os modelos preditivos não sejam capazes de classificar de forma eficiente. Isso leva à exclusão de um dos objetivos principais por trás das pesquisas envolvendo a detecção das emoções com base nas expressões faciais: as individualidades comportamentais de cada usuário e acurácia na classificação dos estados emocionais (3).

No texto, os autores apresentam um método de detecção emocional não intrusiva e personalizável de acordo com os sinais psicofisiológicos estimados de forma remota. O objetivo é potencializar a detecção comportamental e emocional de cada indivíduo (3). Preza-se pelo aprendizado da rede neural com base no perfil emocional de cada usuário durante a interação com os jogos de calibração desenvolvidos para o treinamento deste modelo. Além disso, o modelo é gerado com base nas informações extraídas do próprio usuário. O modelo gerado também é capaz de ser usado em vários ambientes diferentes e em situações que fogem do contexto normal, tornando-o assim um modelo multifatorial (3).

Os autores também apresentam dois experimentos para avaliar o estudo. Nos experimentos foram usados uma câmera de vídeo e um computador, não tendo nenhum sensor colocado diretamente ao corpo dos voluntários que participaram das pesquisas. No teste são apresentados aos voluntários dois tipos de jogos: os jogos de calibração na etapa em que é feito o treinamento do modelo com base nas expressões do jogador, e os jogos de validação na etapa cujo objetivo é detectar as emoções do usuário com base no modelo gerado (3) (Figura 8).

O foco dos experimentos foi realizar um teste de viabilidade e avaliação sistemática do método apresentado, assim testando a pressuposição geral que é a abordagem de detecção de emoção que faz uso dos sinais extraídos de forma remota. Após coletado o resultado dos experimentos, constatou-se que o modelo treinado para a detecção de expressões faciais via visão

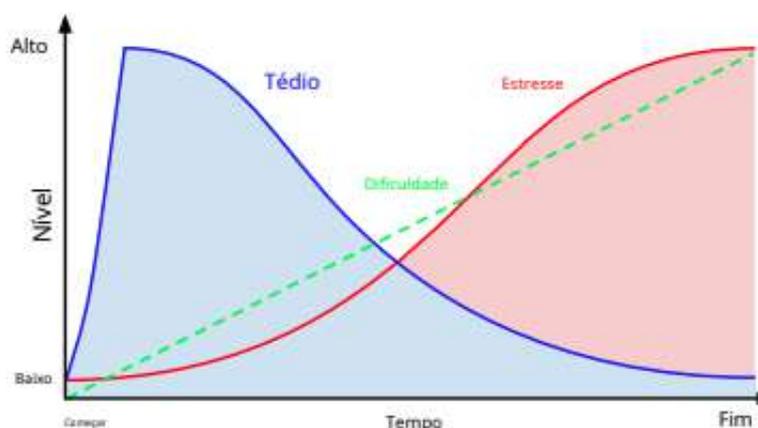


Figura 8 – Estrutura de um jogo de calibração. eixo-x mostra a progressão do tempo. O eixo-y mostra a dificuldade do jogo e o estado emocional que os usuários devem vivenciar, ou seja, estresse e tédio. Fonte: Adaptado de Bevilacqua; Engström; Backlund (3)

computacional obteve uma acurácia média de 61,6%. Em comparação a resultados produzidos por classificações realizado ao acaso, a abordagem apresentou uma margem superior de acertos (3).

A finalidade dos métodos propostos no artigo é a análise baseada em fatores únicos, por conta de sua forma de treinamento do modelo usando jogos de calibração e suas formas de testes usando os jogos de avaliação. Isso leva à preservação do comportamento singular de cada um dos indivíduos e a sua forma não intrusiva de adquirir dados em relação aos sinais expressos pelos usuários (3).

Por fim, com base nos resultados obtidos desses estudos, os autores concluem que esse método é viável. O método tem o potencial de mudar a forma pela qual é realizada uma pesquisa envolvendo a área da computação afetiva e também poderia eventualmente retirar sensores físicos, que são geralmente usados nessas pesquisas. Isso possibilitaria utilizar um recurso tecnológico não intrusivo para detectar o estado emocional de uma pessoa em um contexto que envolve jogos e comportamentos de usuário mais naturais (3).

4.3.3 Detecção de estresse e ansiedade usando pistas faciais de vídeos

Giannakakis et al. (6) apresenta uma metodologia para a captação e análise de emoções envolvendo estresse e ansiedade, com base em pistas faciais gravadas em vídeo. Na pesquisa foi implementado uma série de regras com o objetivo de causar uma diversidade abrangente de estados emocionais: neutro, relaxado, estressado e ansioso. Os estados emocionais foram elicitados por meio de uma variabilidade de estressores. Nas avaliações os pesquisadores tiveram como alvos principais as expressões faciais involuntárias e também as semi-voluntárias, resultando assim numa possível classificação de representações de estados emocionais de forma mais clara. Os aspectos comportamentais analisados englobam ações envolvendo olhos, movimentos da boca, ações voluntárias e involuntárias da cabeça e também a aferição da frequência

cardíaca obtida por fotopletiografia baseada em câmera Giannakakis et al. (6), as etapas da metodologia da pesquisa são ilustradas na (Figura 9).

A partir disso, os autores usaram de técnicas de seleção de características para se conseguir selecionar padrões mais robustos. Acompanhou-se a isso a criação de modelos de classificação diferenciando estresse, ansiedade e estados neutros, tendo como referência uma situação de relaxamento a cada fase experimental. Além do mais, também fez-se uso de classificação via auto relatos com o objetivo de analisar as conexões entre os parâmetros faciais e a quantidade identificada de estresse e ansiedade no participante Giannakakis et al. (6).

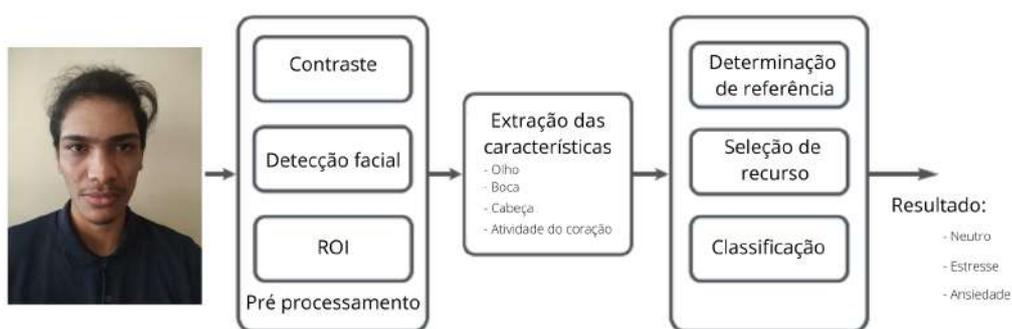


Figura 9 – Ilustração das etapas da metodologia da pesquisa. Fonte: Adaptado de Giannakakis et al. (6)

Em conclusão, com base nos resultados obtidos da pesquisa, pontua-se que as pistas faciais específicas, obtidas a partir de ações e movimentos envolvendo dos olhos, boca, da cabeça e, também, do monitoramento da atividade cardíaca com base na câmera geram uma boa acurácia. Pode se dizer, inclusive, que são recomendáveis para a classificação e a diferenciação de estresse e ansiedade (6).

4.3.4 Um sistema de reconhecimento de estresse usando parâmetros HRV e técnicas de aprendizado de máquina

Giannakakis; Marias; Tsiknakis (7) realizam uma pesquisa voltada a analisar a frequência cardíaca usando variabilidade de frequência cardíaca (HRV) de uma forma confiável. Tem-se como finalidade identificar o estresse. Para isso seguiu-se algumas regras estabelecidas e englobaram-se vários tipos de estressores equiparados a algumas situações rotineiras da vida cotidiana (7).

Um padrão, tido como linha de base personalizada, foi montado para cada um dos voluntários presente na pesquisa. O objetivo foi evitar qualquer diferença entre os voluntários e também para normalizar os dados extraídos. Os espólios da variabilidade de frequência cardíaca extraídos foram modificados conforme a transformação de pares, com o intuito de ressaltar a linha de base personalizada de uma das etapas da construção do modelo (Figura 10) que identifica o estresse (7).

Fase experimental	Afetivo Estado	Duração (min)
Exposição Social		
1.1 Neutro (referência)	N	1
1.2 Entrevista (autodescritiva)	S	1
Lembrança de evento estressante		
2.1 Neutro (referência)	N	1
2.2 Relembra o evento ansioso	S	1
2.3 Relembrar evento estressante	S	1
Carga cognitiva		
3.1 Imagens neutras / estressantes	S	2
3.2 Tarefa do Stroop Color Word	S	2
Vídeos estressantes		
4,1 Neutro (referência)	N	1
4,2 Vídeo relaxante	R	2
4,3 Vídeo de aventura	S	2
4,4 Vídeo de pressão psicológica	S	2

Figura 10 – Ilustração das etapas da construção do modelo proposto. Fonte: Adaptado de Giannakakis; Marias; Tsiknakis (7)

Os métodos mais robustos foram escolhidos aplicando o algoritmo de relevância máxima de redundância mínima (mRMR). Por fim, os métodos escolhidos serviram como entrada para um algoritmo de *Machine Learning* (ML). O resultado foi uma acurácia de classificação de 84,4% usando a técnica de validação cruzada, ilustrado na (Figura 11).

Esse processo serve para avaliar modelos de ML via treinamento de várias variações do mesmo modelo de aprendizado de máquina em subconjuntos de dados de entrada e análise deles no subconjunto complementar dos dados.

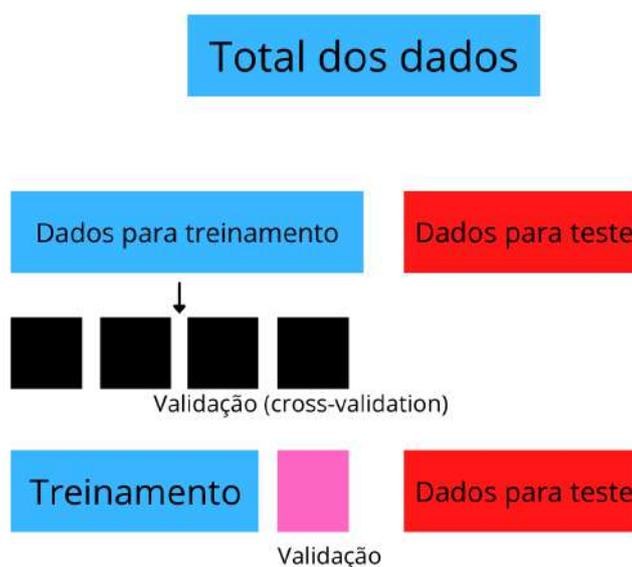


Figura 11 – Técnica de validação cruzada. Fonte: Elaborada pelo autor

5 METODOLOGIA

Inicialmente, elencou-se como prioridade o aprimoramento dos resultados obtidos a partir da pesquisa realizada por Bevilacqua; Engström; Backlund (2) e dos trabalhos desenvolvidos por Hilger; Ramisch; Bevilacqua (9). Para isso, os dados e os sistemas desenvolvidos serviram de fundamentação para este estudo e, a partir disso, foi realizada a revisão das características faciais a fim de catalogar as características relevantes a serem adicionadas, definindo assim um escopo de tarefas a serem implementadas dentro das funcionalidades que o *MediaPipe* oferece. Após essa etapa executaram-se testes de comparação, com o objetivo de avaliar se houve aprimoramento dos resultados obtidos quando comparado com o estudo base.

Em linhas gerais, a metodologia do projeto baseou-se na reimplementação de uma análise facial, a qual buscou utilizar uma tecnologia mais recente, possibilitando uma nova captação de dados com intuito de superar a acurácia anterior de 61,6% (2). Para isso, foram implementadas, inicialmente, as funcionalidades do projeto fazendo uso da tecnologia disponibilizada pelo *Google* — o *MediaPipe* —, onde diferentemente do método utilizado originalmente Bevilacqua; Engström; Backlund (2), o qual extrai 68 pontos faciais, no *MediaPipe* é possível extrair 468 pontos, permitindo que seja possível realizar uma extração mais precisa das zonas faciais.

Ao concluir essa etapa, iniciou-se a fase de testes e, posteriormente, as avaliações dos resultados. Essa última visou testar a viabilidade do método proposto bem como a complexidade das implementações, atentando-se para as fragilidades e potencialidades que a tecnologia apresentou. A partir disso, buscou-se apurar se os objetivos do projeto foram atingidos. Dentre as etapas de execução, a avaliação da precisão do método proposto foi a mais destacada. Nessa etapa, a partir dos materiais utilizados, realizou-se o levantamento dos resultados obtidos a serem comparados com o desfecho anteriormente alcançado (análise facial não baseada no *MediaPipe*); a pesquisa seguiu as etapas conforme a (Figura 12).

As seções a seguir descrevem as etapas necessárias.

5.1 DESENVOLVIMENTO USANDO MEDIAPIPE

O primeiro passo consiste na análise dos recursos disponibilizados pelo *MediaPipe* e, após a análise, separou-se os principais módulos que poderiam ser utilizados. Após isso, foram realizadas implementações com o intuito de ajustá-los de acordo com as necessidades do escopo do projeto, tratando as parametrizações e regras de negócio presente nos módulos utilizados.

5.2 PREPARAÇÃO DOS MATERIAIS, PROCESSAMENTO DOS VÍDEOS E EXTRAÇÃO DOS DADOS

Após concluídas as implementações, separou-se os materiais necessários para a execução do projeto. Dentre os materiais, haviam 16 vídeos obtidos da pesquisa (2) de participantes em um

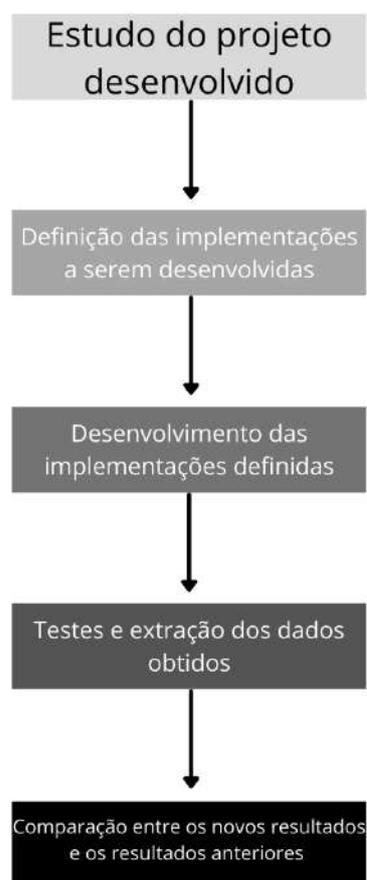


Figura 12 – Ilustração das etapas da pesquisa. Fonte: Elaborada pelo autor

ambiente controlado, no qual estavam jogando, sendo desses 16 vídeos: 2 vídeos do participante 1, 2 vídeos do participante 2, 2 vídeos do participante 3 e 8 vídeos do participante 4. Após a coleta dos materiais, foram realizadas as edições dos vídeos, onde removeu-se os *frames* nos quais os participantes estavam esperando para iniciar o jogo e também os *frames* situados após o término da sessão de jogos. Em seguida, foi utilizada a ferramenta *FFmpeg* para efetuar as normalizações dos vídeos, uniformizando-os para as dimensões de 640x480 e com 30 *frames per second* (FPS). Por fim, após a conclusão das implementações e com os materiais adequados às necessidades do projeto, se sucedeu os processamentos dos vídeos para a extração das *features*. Como resultado, após o processamento de cada vídeo, foi gerado um arquivo *CSV* contendo todos dados extraídos do vídeo; ao todo foi gerado um total de 16 arquivos *CSV*.

5.3 REDE NEURAL UTILIZADA

Inicialmente, o modelo de aprendizado de máquina continha uma rede neural do tipo *feedforward*, que continha apenas uma camada oculta; uma rede construída com essas características de arquitetura permite que seja analisado cada *input* de dados de forma singular (9). Como descrito no Capítulo 1, uma das características do projeto é a produção de um modelo de

aprendizado de máquina de forma única para cada indivíduo, sendo que durante os vídeos, cada jogo utilizado tem a finalidade de induzir o tédio no início e estresse no fim, realizando-se uma indução emocional de forma gradual. Dito isso, nas novas implementações desenvolvidas por Hilger; Ramisch; Bevilacqua (9), usou-se uma rede neural RNN (*Recurrent Neural Network - Rede Neural Recorrente*) com LSTM (*Long-Short Term Memory*) devido a sua tecnologia permitir sanar problemas com a arquitetura canônica (9), com isso esse seria o modelo de redes a ser adotado neste trabalho.

5.4 EXECUÇÃO E APURAÇÃO DOS RESULTADOS

Após finalizada a extração de todos os dados em conjunto com as novas *features* classificadas a partir do *Mediapipe*, o treinamento da rede neural seguiu as seguintes etapas: *a priori*, executou-se um pré-treinamento utilizando todos os dados apurados. Subsequente a isso, realizou-se novamente o treinamento nas coleções geradas, tendo como objetivo produzir um modelo singular para cada indivíduo, resultando em uma acurácia final calculada com base na média das acurácias de todos os conjuntos de validações adotados nos modelos gerados de forma individual para cada indivíduo.

6 RESULTADOS

Neste capítulo, será apresentado os resultados encontrados executando a metodologia descrita no capítulo 5.

6.1 CONFIGURAÇÕES DO EXPERIMENTO

Os resultados foram coletados usando um microcomputador com as respectivas configurações: AMD FX(tm)-8300 Eight-Core Processor × 4, NVIDIA Corporation GP106 [GeForce GTX 1060 6GB] e 32 GB de memória.

6.2 OBSERVAÇÕES DOS TESTES

Na arquitetura LSTM foram estudadas diversas variações, como o número de camadas escondidas, de neurônios e também diferentes tipos de neurônios, tendo como melhor resultado o uso de três camadas escondidas, sendo duas de 64 neurônios e uma de 16 neurônios (9). A acurácia média resultante foi de 62,51% — uma melhoria de quase 1,31% em comparação com os 61.2% do trabalho original.

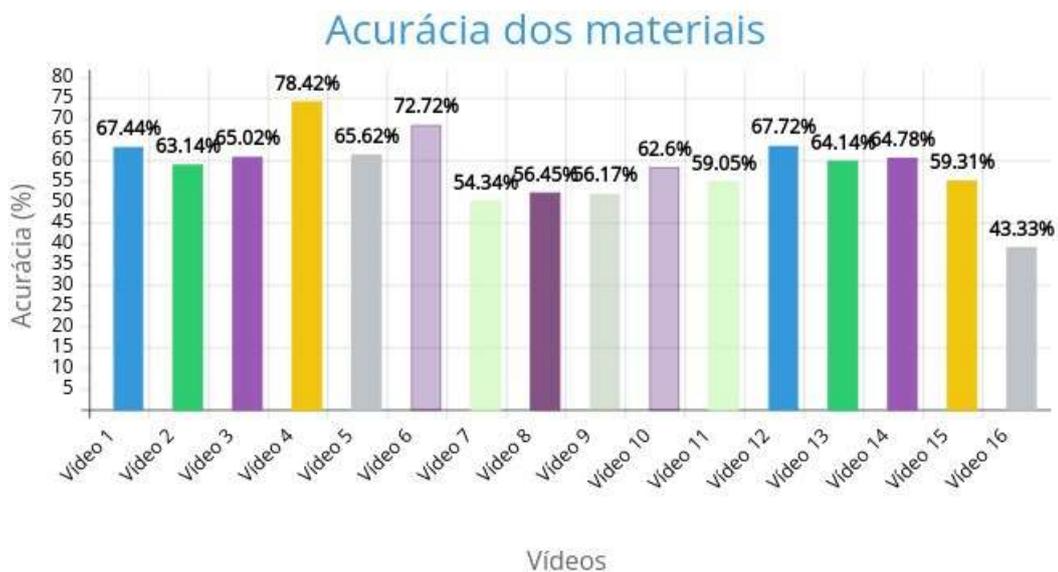


Figura 13 – Ilustração dos resultados obtidos. Fonte: elaborado pelo autor

Avaliando os resultados de forma mais aprofundada, constata-se que há uma grande carência na quantidade de materiais utilizados e também uma grande variação na acurácia dos modelos de cada indivíduo. Adicionado a isso, é notável a necessidade de avaliar mais a fundo o processo de extração das *features* e de aprendizagem juntamente com os vídeos originais, que não estavam todos disponíveis no decorrer desta pesquisa.

7 CONCLUSÃO

No início do trabalho, apresentou-se a importância da visão computacional no âmbito da detecção de emoções no contexto dos jogos digitais, as pesquisas anteriormente realizadas nesta vertente e também o objetivo central do desenvolvimento do projeto, buscando evoluir a pesquisa desenvolvida por Bevilacqua; Engström; Backlund (2). Além disso, explicaram-se algumas teorias relacionadas ao campo de pesquisa das emoções, sendo as duas principais a Teoria do Fluxo Nakamura; Csikszentmihalyi (13) e o modelo *Circumplex* de Afeto (MCA) Russell (15). Por fim foi exposta a ferramenta a ser utilizada no desenvolvimento do projeto — o *MediaPipe* Lugaresi et al. (12).

No capítulo 2 projetaram-se os objetivos da pesquisa, ressaltando a intenção do aperfeiçoamento dos métodos de análises faciais por meios não invasivos. Além disso, detalharam-se os objetivos gerais para o desenvolvimento do estudo.

O capítulo 3 apresentou a importância da pesquisa e o que inspirou o desenvolvimento do trabalho: o uso da Visão computacional para solucionar problemas envolvendo métodos de pesquisa dentro da área de emoções e jogos digitais.

No capítulo 4 foram descritos os trabalhos relacionados mais importantes, que definem conceitos como o reconhecimento de emoções humanas, expressões faciais e as tecnologias usadas no desenvolvimento do projeto, compondo assim parte da fundamentação teórica deste trabalho.

Nos capítulos 5 e 6 descreveu-se o fluxo do desenvolvimento e testes deste trabalho, e também apresentaram-se particularidades das implementações e das coletas dos dados utilizados na ferramenta desenvolvida. Por fim apurou-se a contabilização dos materiais obtidos e dos resultados gerados.

O último capítulo, fornece uma análise dos resultados obtidos a partir da execução das etapas apresentadas no decorrer do texto. O apuramento gerado a partir das comparações com os resultados anteriores da pesquisa desenvolvida por Bevilacqua; Engström; Backlund (3) nos mostra que, com base na tecnologia que foi utilizada — o *MediaPipe*, atingiram-se resultados positivos nos quais foi possível constatar um progresso na pesquisa. Por fim, conforme os dados apresentados no capítulo 6, é possível afirmar que os estudos projetados no decorrer da pesquisa contêm um futuro promissor caso sejam sanados os problemas apresentados ao fim deste estudo. Como sugestões para pesquisas futuras em aplicações reais em projetos de jogos digitais envolvendo ferramentas com a finalidade de detecção de estados emocionais, apontam-se duas possibilidades de desenvolvimento e estudo: análise de tecnologias mais atuais no âmbito da extração de *features* de um indivíduo durante a execução de um vídeo gravado em um ambiente controlado, e também a busca de novas tecnologias capazes de conseguir analisar os estados emocionais de um jogador de forma que não seja necessário a presença de um jogador em um ambiente controlado, de forma que isso não afete o resultado da acurácia final.

REFERÊNCIAS

- 1 ADOLPHS, Ralph. Recognizing emotion from facial expressions: psychological and neurological mechanisms. **Behavioral and cognitive neuroscience reviews**, v. 1, n. 1, p. 21–62, 2002. DOI: 10.1177/1534582302001001003.
- 2 BEVILACQUA, Fernando; ENGSTRÖM, Henrik; BACKLUND, Per. Automated Analysis of Facial Cues from Videos as a Potential Method for Differentiating Stress and Boredom of Players in Games. **Int. J. Comput. Games Technol.**, Hindawi Limited, London, GBR, v. 2018, p. 1, mar. 2018. ISSN 1687-7047. DOI: 10.1155/2018/8734540. Disponível em: <<https://doi.org/10.1155/2018/8734540>>.
- 3 _____. Game-Calibrated and User-Tailored Remote Detection of Stress and Boredom in Games. **Sensors**, v. 19, n. 13, 2019. ISSN 1424-8220. DOI: 10.3390/s19132877. Disponível em: <<https://www.mdpi.com/1424-8220/19/13/2877>>.
- 4 CHANEL, Guillaume et al. Boredom, Engagement and Anxiety as Indicators for Adaptation to Difficulty in Games. In: p. 13–17. DOI: 10.1145/1457199.1457203.
- 5 EKMAN P., Friesen W.V. **Unmasking the Face: A Guide to Recognizing Emotions from Facial**. [S.l.]: Prentice Hall Direct, 1975. ISBN 9780139381836.
- 6 GIANNAKAKIS, G. et al. Stress and anxiety detection using facial cues from videos. **Biomedical Signal Processing and Control**, v. 31, p. 89–101, 2017. ISSN 1746-8094. DOI: <https://doi.org/10.1016/j.bspc.2016.06.020>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1746809416300805>>.
- 7 GIANNAKAKIS, Giorgos; MARIAS, Kostas; TSIKNAKIS, Manolis. A stress recognition system using HRV parameters and machine learning techniques. In: 2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW). [S.l.: s.n.], 2019. p. 269–272. DOI: 10.1109/ACIIW.2019.8925142.
- 8 GIANNAKAKIS, Giorgos et al. Review on psychological stress detection using biosignals. **IEEE Transactions on Affective Computing**, p. 1–1, 2019. DOI: 10.1109/TAFFC.2019.2927337.
- 9 HILGER, Jean Carlo; RAMISCH, Junior Vitor; BEVILACQUA, Fernando. INTERAÇÃO HUMANO-COMPUTADOR: FUNDAMENTOS E APLICAÇÕES. **JORNADA DE INICIAÇÃO CIENTÍFICA E TECNOLÓGICA**, v. 1, n. 11, 2021.
- 10 J., Zhang et al. Detecting Negative Emotional Stress Based on Facial Expression in Real Time. **2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP)**, 2019. DOI: 10.1109/siprocess.2019.8868735.

- 11 JN., Bassili. Emotion recognition: the role of facial movement and the relative importance of upper and lower areas of the face. **J Pers Soc Psychol**, v. 37, n. 11, p. 2049–2058, 1979. DOI: 10.1037/0022-3514.37.11.2049.
- 12 LUGARESI, Camillo et al. MediaPipe: A Framework for Building Perception Pipelines. **CoRR**, abs/1906.08172, 2019. arXiv: 1906.08172. Disponível em: <<http://arxiv.org/abs/1906.08172>>.
- 13 NAKAMURA, Jeanne; CSIKSZENTMIHALYI, Mihaly. The Concept of Flow. In: **FLOW and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi**. Dordrecht: Springer Netherlands, 2014. p. 239–263. ISBN 978-94-017-9088-8. DOI: 10.1007/978-94-017-9088-8_16. Disponível em: <https://doi.org/10.1007/978-94-017-9088-8_16>.
- 14 PANTIC, M.; ROTHKRANTZ, L.J.M. Automatic analysis of facial expressions: the state of the art. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 22, n. 12, p. 1424–1445, 2000. DOI: 10.1109/34.895976.
- 15 RUSSELL, J. A. A circumplex model of affect. **Journal of Personality and Social Psychology**, v. 39, n. 6, p. 1161–1178, 1980. DOI: 10.1037/h0077714.
- 16 SA., Fahlman et al. Development and validation of the multidimensional state boredom scale. **Assessment**, v. 20, n. 1, p. 68–85, 2013. DOI: 10.1177/1073191111421303.