

Utilização de Técnicas da Mineração de Dados para a Classificação de Jogadores por Padrões em suas Expressões Faciais

Using Data Mining Techniques to Classify Players by Patterns within Their Facial Expressions

Nicolau Calado Jofilsan¹

 orcid.org/0000-0002-5227-7998

Sidney M. L. Lima²

 orcid.org/0000-0002-4350-9689

Renan Costa Alencar¹

 orcid.org/0000-0003-0028-558X

Sthéfano H. M. T. Silvas¹

 orcid.org/0000-0003-1715-7525

João V. O. Albuquerque¹

 orcid.org/0000-0002-4528-8963

Odilon H. Soares Filho¹

 orcid.org/0000-0001-5584-169X

Alexandre M. A. Maciel¹

 orcid.org/0000-0003-4348-9291

Wellington P. dos Santos¹

 orcid.org/0000-0003-2558-6602

¹Escola Escola Politécnica de Pernambuco, Universidade de Pernambuco, Recife, Brasil. E-mail: ncj@ecompp.poli.br

²Departamento de Eletrônica e Sistemas, Universidade Federal de Pernambuco UFPE Recife, Brasil.

³Escola Politécnica de Pernambuco, Universidade de Pernambuco, Recife, Brasil.

Artigo recebido em:
Artigo aceito em:

DOI: 10.xxxx/s11468-014-9759-3

Esta obra apresenta Licença Creative Commons Atribuição-Não Comercial 4.0 Internacional.

Como citar este artigo pela NBR 6023/2018: AUTOR 1; AUTOR 2; AUTOR 3. Título do artigo. Revista de Engenharia e Pesquisa Aplicada, Recife, v. , n. , p. , mês, ano. DOI: Disponível em: . Acesso em: .

RESUMO

Este trabalho apresenta como principais contribuições um modelo de mineração dos dados das expressões faciais de jogadores de jogos digitais e o agrupamento desses jogadores por padrões. A base de dados analisada foi gerada utilizando o sistema de visão computacional *clmtrackr.js Face tracker*, que compreende as expressões faciais como linhas formadas por pontos. As técnicas de mineração utilizadas foram aplicadas especificamente nos dados com as variações das 6 emoções básicas felicidade, tristeza, medo, surpresa, raiva e nojo, que representam a variação das expressões faciais no decorrer do uso dos jogos. Após a extração e o tratamento desses dados, foi possível agrupar os jogadores de acordo com padrões em suas expressões faciais. Por fim, como forma de visualização de cada grupo de jogadores, os resultados gerados a partir dos centroides dos valores das suas 6 emoções básicas foram inseridos em um simulador de expressões faciais.

PALAVRAS-CHAVE: Mineração de Dados; Jogos Digitais; Interação Humano Computador; Computação Afetiva;

ABSTACT

This work presents as main contributions a model of data mining of facial expressions of digital game players and the grouping of these players by standards. The analyzed database was generated using the computer vision system *clmtrackr.js Face tracker*, which comprises facial expressions as lines formed by points. The mining techniques used were applied specifically to the data with the variations of the 6 basic emotions happiness, sadness, fear, surprise, anger and disgust, which represent the variation of facial expressions during the use of the games. After extracting and processing this data, it was possible to group the players according to patterns in their facial expressions. Finally, as a way of visualizing each group of players, the results generated from the centroides of the values of their 6 basic emotions were inserted into a facial expression simulator.

KEY-WORDS: Data Mining; Digital Games; Human Computer Interaction; Affective Computing;

1 INTRODUÇÃO

A utilização de uma abordagem de coleta de dados baseada em sensoriamento remoto para adquirir sinais psicofisiológicos de usuários é mencionada na literatura como uma possível solução para o problema de coleta de dados sobre as expressões faciais dos jogadores. A análise facial é uma abordagem discreta para detecção de emoções dos jogadores [1]. Avanços na visão computacional permitem captura detalhada da atividade facial, tornando a análise automatizada o suficiente para inferir sobre as emoções dos jogadores [2,3]. Assim, expressões faciais detectadas automaticamente têm sido utilizadas no contexto dos jogos, inclusive para correlacionar dimensões da experiência do jogador e para aprimorar os jogos digitais [4-6].

O processo de mapear sinais psicofisiológicos adquiridos física ou remotamente envolve a identificação dos sinais que melhor predizem os estados emocionais [7], bem como uma definição adequada dos próprios estados emocionais [8]. Uma abordagem comum usada para realizar esse mapeamento é a aplicação de aprendizado de máquina. O uso de aprendizado de máquina geralmente começa com um grupo de usuários sendo exposto a materiais de elicitación de emoção como imagens e vídeos com rótulos emocionais conhecidos, como estresse e tédio. Os sinais desses usuários como expressões faciais são medidos durante a interação e usados para treinar um modelo de aprendizado de máquina de acordo com os materiais de descobrimento rotulados. Consequentemente, os modelos preditivos são baseados em uma perspectiva coletiva. Como um modelo geralmente é treinado a partir de dados de vários usuários, na prática, ele potencialmente descreve o comportamento médio do grupo, sem identificar diferentes perfis dentro desse mesmo grupo.

Trabalhos anteriores se concentram na utilização de sensores físicos intrusivos para adquirir sinais do usuário para detecção de emoções. Os estudos mencionados usam sensores físicos, materiais de elicitación de emoções não interativos como imagens, instruções para os participantes ficarem quietos e amostras altamente correlacionadas para treinar e testar qualquer modelo de classificação de emoção. Mesmo quando essas abordagens não são obstrutivas, e é utilizada análise facial com base em visão computacional, os modelos são treinados com emoções geradas a partir de interatividade limitada,

expondo o participante a imagens e vídeos. Quando expostos a tais materiais, os participantes enfrentam um papel passivo com oportunidades limitadas de interação.

Diante disso, faltam trabalhos com modelos de análise de dados gerados a partir da detecção de emoções de forma não intrusiva e customizada para o usuário no contexto dos estímulos em jogos digitais.

1.1 DESCRIÇÃO DO PROBLEMA

A análise facial através de métodos visuais utilizando visão computacional, em geral são mais rápidos e fáceis de implantar se comparados aos métodos que utilizam sensores físicos. No entanto, os trabalhos anteriores geralmente se concentram na análise de imagens ou vídeos cujos sujeitos realizaram expressões faciais sob orientação. Essas são circunstâncias artificiais que não retratam as interações naturais de usuários e jogos digitais [1]. Quando a análise é realizada em vídeos de usuários interagindo com jogos digitais, geralmente o objetivo é detectar um conjunto muito específico de expressões faciais como as 6 expressões faciais universais, desconsiderando as mudanças sutis no comportamento da face.

O presente trabalho se destaca, ao realizar análise de base de dados autoral com 68.299 amostras, geradas por 232 jogadores e coletadas em situação contextualizada no uso de jogos digitais. A partir da análise dos dados, esse estudo priorizou a busca de nuances nas expressões faciais, microexpressões, ao invés de procurar classificar expressões faciais predefinidas, alheias ao contexto do uso de jogos digitais.

1.2 OBJETIVOS

O objetivo deste trabalho foi a identificação dos perfis de jogadores através do uso de técnicas da mineração de dados. Esses perfis foram definidos por meio da análise e agrupamento por padrão das expressões faciais dos jogadores. Para tanto, em meio a investigação sobre a natureza dos dados, foi criado um modelo de mineração de dados com o objetivo de explicitar as etapas metodológicas realizadas neste trabalho e auxiliar nas análises futuras.

1.3 JUSTIFICATIVA

A detecção de emoções do usuário é considerada importante elemento da computação afetiva [9], possui diversas aplicações no campo da interação humano-computador, especialmente em pesquisas ligadas a jogos digitais. As técnicas mais comuns utilizadas para obter dados sobre o estado emocional dos jogadores, enquanto jogam, são medidas psicofisiológicas [10]. Estas geralmente se dão por sensores físicos que fornecem medição ininterrupta dos sinais fisiológicos do usuário enquanto joga [11-14]. Uma quantidade relevante de informações psicofisiológicas pode ser lida a partir de sinais emitidos pelo corpo humano, como frequência cardíaca, respiratória e expressões faciais, dentre outros. A partir da combinação desses sinais é possível inferir sobre o estado emocional do usuário [15,16]. Em um modo de captura não obstrutivo, de forma remota, não é permitido o uso de sensores físicos, o que pode gerar interrupções na coleta dos dados e ruídos, já que a webcam precisa estar adequadamente posicionada, com iluminação suficiente e devidamente conectada à Internet. Os sensores também podem restringir os movimentos do usuário enquanto joga, um sensor ligado a um dedo impede o uso desse dedo, por exemplo. A utilização da captura remota ajuda na não consciência do usuário de ser monitorado, melhorando a obtenção das informações pela espontaneidade, mas afeta os resultados no processo de captura dos sinais corporais [17-19].

Dessa forma, faltam estudos na área que analisem base de dados coletada remotamente e de forma contextualizada. A investigação dessa base de dados pode gerar informações e *insights* sobre o comportamento do jogador ainda desconhecidos na academia e na indústria.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 TEORIA DO FLUXO (*Flow*)

Na indústria, a detecção discreta de níveis de estresse e tédio pode ser usada para avaliar jogos, uma vez que essas emoções são componentes chave no ciclo de desafio recompensa, que torna um jogo divertido e engajador [20]. Uma avaliação sobre estresse e tédio, por exemplo, permite que desenvolvedores e editores otimizem seus jogos antes de lançá-los ao mercado, aumentando suas chances de sucesso financeiro. Assim, os

desenvolvedores podem explorar novas maneiras de criar jogos que detectam níveis de estresse e tédio sem hardwares ou sensores adicionais, fazendo com que os jogos autoajustem o nível de dificuldade durante o tempo de execução para melhorar a experiência dos jogadores.

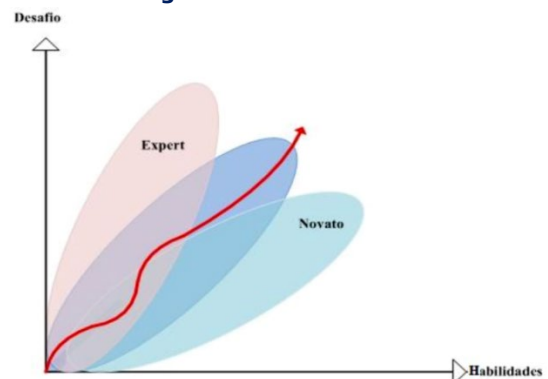
Sweetser e Wyeth [21] fizeram uma revisão da literatura sobre a experiência de usuários em jogos para determinar como os elementos do Flow se manifestam nos jogos de computadores. Um modelo de satisfação em jogos foi construído a partir da literatura baseado nos elementos do Flow e nas evidências da experiência do Flow em jogos.

O resultado conseguido foi o modelo GameFlow, que consiste em oito elementos principais: concentração, desafio, habilidades, controle, regras claras, feedback, imersão e interação social. Cada elemento consiste em um número variável de critérios que se relacionam com os elementos de Flow proposto por Csikszentmihalyi, afirmam Sweetser e Wyeth [21].

Descrições da experiência de Flow são idênticas aos que os jogadores experimentam quando imersos em jogos, perdendo a noção do tempo e pressões externas, juntamente com outros interesses [22].

Um conceito inspirado na teoria de Flow é a Zona do Flow. Para manter a experiência de Flow do usuário a atividade deve balancear o desafio inerente a essa atividade e a habilidade do jogador para superá-lo. Se o desafio está além das habilidades do jogador, a atividade se torna intransponível gerando estresse excessivo. Se o desafio falha em engajar o jogador, ele rapidamente perde o interesse e tende a deixar o jogo. O jogo deve refletir o balanceamento correto entre desafio e habilidade para manter o jogador na Zona do Flow [22].

Figura 1 – Zonas de Flow



Fonte: Adaptado de CHEN [22].

Considerando o contexto do uso de jogos digitais, diferentes jogadores têm diferentes habilidades e esperam diferentes desafios [22]. Diferentes jogos são atrativos para diferentes personalidades, e não apenas por causa de um problema particularmente atraente para o cérebro. Isto ocorre também por causa de uma solução particular que é atraente para um tipo de cérebro, e quando algo está dando certo o jogador procura não alterá-la. Para manter o jogador na Zona de Flow é importante oferecer diversas maneiras de experiência [23].

Muitos jogos oferecem apenas uma experiência simples e estática mantendo apenas um jogador típico no Flow, mas não será divertido para novatos ou experts, conforme ilustra a Figura 1 [22]. Nesse caso, identificar diferentes perfis de jogadores pode ser de suma importância para projetar experiências interessantes que podem ser oferecidas através dos jogos digitais.

2.2 EXPRESSÕES FACIAIS DOS JOGADORES

A análise do comportamento facial comumente se baseia em dados obtidos de sensores físicos como eletromiografia (EMG), ou da aplicação de métodos visuais para avaliar o rosto através de extração de características via visão computacional [24]. A abordagem baseada em dados EMG usa sensores físicos acoplados aos indivíduos para medir a atividade elétrica dos músculos faciais, como o zigomático, o orbicular dos olhos e o músculo corrugador do supercílio. O estudo realizado por Hazlett [25] apresenta evidências de atividade do músculo corrugador mais frequente quando ocorrem eventos positivos nos jogos. Por sua vez, Tijs et al. [14] mostram aumento da atividade do músculo zigomático associado a emoções positivas. Da mesma forma, Ravaja et al. [26] mostram que eventos positivos e gratificantes no jogo estão ligados ao aumento da atividade EMG zigomática e orbicular do olho. As abordagens baseadas em EMG são mais resistentes às variações das condições de iluminação e oclusão facial, no entanto, são intrusivas, uma vez que sensores físicos devem ser fixados ao rosto do sujeito. Ao contrário das abordagens baseadas em EMG, intrusivas e obstrutivas, a análise do

comportamento facial baseada em métodos visuais automatizados pode ser realizada remotamente e sem contato físico. O processo

geralmente envolve detecção do rosto, localização de características faciais, também conhecidas como pontos fiduciais, e classificação de tais informações em expressões faciais [27]. Uma abordagem de classificação comum é baseada nas distâncias e ângulos dos pontos de referência. Samara et al. [28] usam a distância euclidiana entre os pontos de face para treinar um modelo de máquina de vetores suporte (SVM) para detectar expressões. Da mesma forma, Chang et al. [29] usam 12 distâncias calculadas a partir de 14 marcos para detectar medo, amor, alegria e surpresa. Os pesquisadores Hammal et al. [30] usam 5 distâncias faciais calculadas a partir de linhas em regiões chave do rosto, derivadas do padrão de animação MPEG-4 [31], para classificação de expressões. Tang e Huang [32, 33] usam até 30 distâncias euclidianas entre os marcos faciais também obtidos a partir de modelos de face 3D baseados em MPEG-4 para reconhecer as 6 expressões faciais universais. Da mesma forma, Hupont et al. [34] classificam as mesmas emoções usando uma técnica de seleção de características baseada em correlação para selecionar as distâncias e ângulos mais significativos dos pontos faciais. Finalmente Akakn e Sankur [35] usam as trajetórias de marcos faciais para reconhecer gestos de cabeça e expressões faciais.

2.3 TRABALHOS RELACIONADOS

Alguns métodos visuais contam com análise manual ou automatizada baseada em Facial Action Coding System (FACS) como um padrão para categorização e medição de expressões emocionais [36]. Kaiser et al. [37] demonstram que mais Action Units (AUs) foram relatadas por codificadores FACS manuais durante a análise de gravações de vídeos de usuários jogando a parte estressante de um jogo, quando comparados à sua parte neutra. Além disso, os autores relatam o canto do lábio e o aumento da sobrancelha como AUs mais frequentes durante as sessões de jogo. Wehrle e Kaiser [38] usam uma análise facial automatizada baseada em FACS agregada com dados de eventos de jogo para fornecer uma análise da avaliação do estado emocional do usuário. Da mesma forma, Grafsgaard et al. [39] usam uma análise automatizada baseada em FACS para relatar uma relação entre a expressão facial e aspectos de envolvimento, frustração e aprendizagem em sessões de tutoria. Ao contrário

do trabalho anterior, Heylen et al. [40] não contam com FACS, mas em vez disso, usam uma análise facial empírica e manual com base na interpretação dos autores sobre o contexto. Heylen et al. [40] descobriram que na maioria das vezes os sujeitos permanecem com uma face neutra.

A análise facial com base em sensores físicos (EMG), fornece monitoramento contínuo dos usuários e não é afetada pelas condições de iluminação ou oclusão de pose gerada pelos movimentos do usuário. No entanto, os sensores são intrusivos e o uso desses sensores aumenta a consciência do usuário de estar sendo monitorado [41-43]. Abordagens baseadas em análise de vídeo utilizando FACS e visão computacional, são menos intrusivas. Apesar do FACS ter provado ser uma abordagem útil e quantitativa para medir expressões faciais [36], sua aplicação manual é trabalhosa e demorada e requer profissionais certificados para inspecionar as gravações de vídeo. A aplicação de FACS também possui desvantagens, incluindo a má interpretação de expressões faciais de indivíduos de diferentes culturas e etnias [44].

Nesse caso, a análise facial através de método visual utilizando visão computacional pode ser menos subjetiva se considerarmos sua natureza discreta.

3 MATERIAIS E MÉTODOS

A base de dados utilizada neste trabalho foi constituída a distância, através do experimento disponível no endereço emotiondetect.com.br. A coleta dos dados seguiu a mesma metodologia desenvolvida no trabalho de Bevilacqua [1], que possui foco na coleta desses dados presencialmente, em ambiente controlado. Mesmo realizando a coleta em campo, a distância, a base de dados autoral não apresentou prejuízo significativo na qualidade dos dados adquiridos. A privacidade do usuário também foi garantida, já que suas expressões não foram gravadas em vídeo, apenas as linhas de expressão, formadas por pontos, foram coletadas através do *clmtrackr.js face tracker* [45], como mostra a figura 2.

Figura 2 – Exemplo das linhas de expressão capturadas pelo sistema de visão computacional *clmtrackr.js face tracker*

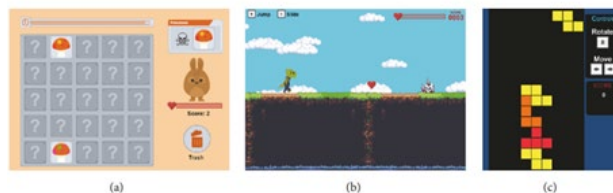


Fonte: [45].

A base de dados utilizada neste artigo foi construída a partir da detecção das expressões faciais de 232 jogadores durante a interação com três jogos digitais. A abordagem utilizada é eficiente em detectar as emoções dos jogadores de forma discreta a partir da análise facial destes. O método automatizado de análise utiliza a visão computacional por meio de webcam e a extração, feita de forma discreta de características faciais como mostra a figura 2. Esse sistema permite detectar a atividade de um conjunto de músculos faciais como um método de aferição das 6 emoções básicas obtidos naturalmente durante os jogos.

O experimento foi realizado com o uso de três jogos digitais customizados, com progressão linear e constante, de um estado enfadonho de tédio a um estressante, sem níveis predefinidos ou condições de parada. Esses jogos foram criados como uma abordagem válida na exploração de comportamentos faciais e sinais fisiológicos, ao considerar sua conexão com os estados emocionais que se deseja capturar.

Figura 3 – Figura 3: Jogos utilizados no experimento: a) Cogumelo, b) Plataforma, c) Tetris.

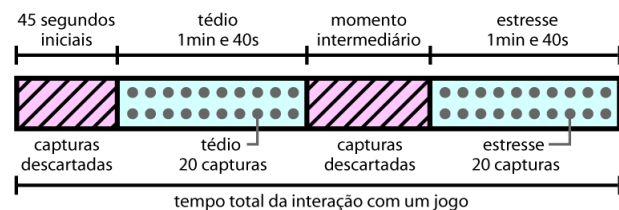


Fonte: [1].

O tempo médio da participação no experimento foi de 25 minutos. Durante esse tempo, os participantes jogaram três jogos diferentes na seguinte ordem a) *Jogo do cogumelo*, onde o jogador deve separar os cogumelos bons dos ruins analisando as cores; b) *Jogo estilo plataforma*, em que o jogador deve saltar ou deslizar por baixo de

obstáculos enquanto coleta corações; c) Tetris, clone da versão clássica do jogo, com a adição das dicas sobre a próxima peça a cair.

Figura 4 – Critério utilizado na extração inicial dos dados das expressões faciais em momentos de tédio e estresse



Fonte: Adaptado de [1].

Essa extração inicial teve o objetivo de separar os dados que representaram os momentos de maior intensidade das reações nas expressões faciais. Em cada um dos 3 jogos jogados pelo participante, 20 persistências dos dados no momento de tédio e 20 no momento de estresse deveriam ser extraídas de todo o conjunto de dados. Isso totalizou 60 momentos com a persistência das expressões faciais de um jogador no momento de tédio e 60 no momento de estresse. Essas persistências dos dados gerados pelas capturas aconteceram de 5 em 5 segundos.

3.1 MODELO DE MINERAÇÃO DOS DADOS DAS EXPRESSÕES FACIAIS

O modelo de mineração dos dados das expressões faciais dos jogadores de jogos digitais proposto nesse trabalho, surgiu após análise do contexto e da natureza dos dados armazenados na base.

Em relação a etapa 1, por consistir em um experimento realizado a distância, uma série de vantagens podem ser atribuídas a essa solução como escalabilidade, facilidade de acesso, dispensar a necessidade de deslocamento e de um ambiente controlado, por exemplo. Mas como a maioria das soluções de design, ao resolver alguns problemas, é comum que outros surjam, como a falta de engajamento na realização de experimentos a distância, obstruções na captura da câmera, ruídos e distrações no ambiente e instabilidade na internet.

Isso gerou, ao mesmo tempo, uma base de dados 2. rica ao considerarmos o contexto e a diversidade dos estilos de jogadores que participaram do experimento, mas sensível no tocante a incompletude dos dados em uma quantidade considerável de participações.

Considerando esses possíveis problemas, a etapa 3. Extração das características, atuou como pré-

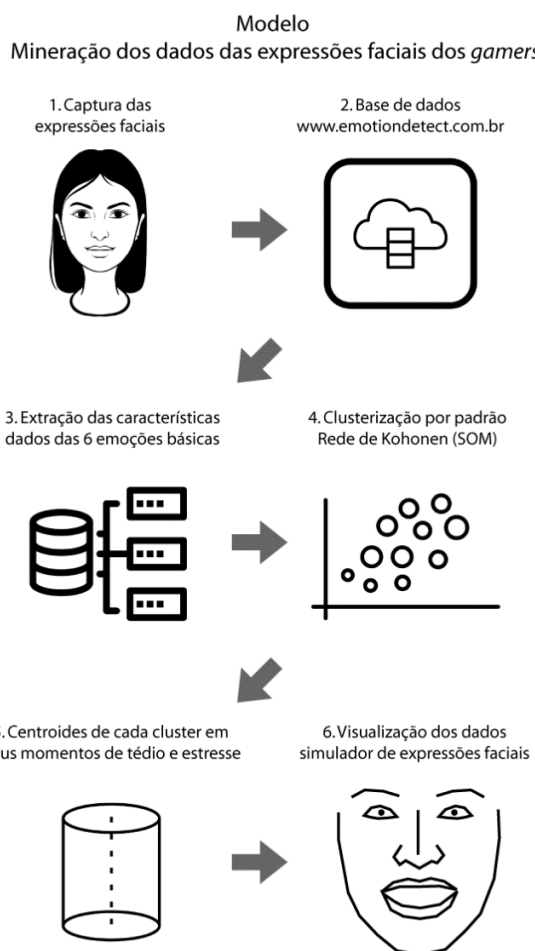
processamento onde, além de extrair os dados de interesse dessa análise, as variações das 6 emoções básicas, também organizou o restante dos dados em linhas e colunas por tipo de dado.

Na etapa 4, foi utilizada a clusterização com rede neural não supervisionada do tipo SOM em seis grupos gerados de acordo com padrões nas expressões faciais.

A etapa 5. Consistiu em gerar centroides dos dados com as variações das 6 emoções básicas de cada cluster em seus momentos de tédio e estresse respectivamente.

Finalmente, na etapa 6. Esses valores dos centroides foram inseridos em um simulador de expressões faciais para viabilizar a visualização dos dados tratados.

Figura 5 – Modelo – Mineração dos dados das expressões faciais dos gamers



Fonte: Os Autores.

3.2 DESCRIÇÃO DA BASE DE DADOS

A base de dados constitui a leitura de expressões faciais e respostas de questionários durante a realização dos três jogos por 232 jogadores. As diferentes experiências de jogo deram à base uma composição 68.299 linhas de dados. Presume-se que os jogadores são de ambos os gêneros, detectado pelo processo de captura, e de diferentes idades. Tendo em vista que todo o processo do experimento é realizado de forma remota e com total isonomia.

A base de dados bruta apresenta inicialmente quatro campos `fk_game`, `timestamp`, `uuid` e `data` que representam respectivamente a identificação do jogo, o tempo, a identificação do usuário e os dados obtidos durante a captura.

O campo `data` apresenta um array de JSON com persistência dos dados de 5 em 5 segundos. Cada linha da nova base de dados, com a abertura do campo `data`, representa o cálculo da distância euclidiana de pontos de referência faciais. Estas variações determinam a probabilidade das seis emoções básicas medo, tristeza, surpresa, raiva, alegria, nojo.

Após a abertura do campo `data` a base passou a possuir 34 variáveis, dentre elas 24 pontos de leituras das distâncias euclidianas indo do ponto 1 (`p1`) ao ponto 24 (`p24`). Além da distância euclidiana, o banco conta com os campos `tempo` (`timestamp`), `identificação do usuário` (`uuid`) e `identificação do jogo` (`game_id`) que são compostos por números inteiros e os demais campos são formados por dados do tipo `float`.

3.2.1 Extração das Características – Pré Processamento

Os dados de capturas dos eventos foram armazenados como um array de objetos JSON dentro do banco de dados. Para realizar a extração das características e descartar os dados irrelevantes para o estudo, foram desenvolvidos scripts de separação de dados usando a linguagem de programação Python, possibilitando que além fazer a extração dos dados fossem aplicados os filtros necessários para excluir dados irrelevantes e aplicar as regras utilizadas no estudo para seleção das amostras de dados.

Por se tratar de dados estruturados não indexados, foi utilizado o algoritmo de árvores binárias de busca, para realizar uma busca

incremental e interativa em todos os elementos do array de objetos JSON dentro do campo `data` do banco de dados. Todos os objetos dentro do array de JSON possuem a mesma estrutura de dados, o elemento `"t"` que guarda a fração de milissegundos em que a captura foi realizada, e o elemento `"d"` que guarda todas as variáveis do evento de captura.

A estratégia de busca em árvores binárias iterativas foi escolhida porque para realizar a leitura dos registros do banco de dados é necessário fazer a varredura em todo array de objetos JSON e verificar cada elemento e o seu respectivo valor, isso possibilita que o script de extração consiga classificar esses dados como dado válido para o resultado estudo, como por exemplo as distâncias euclidianas, ou como dado inútil para o resultado do estudo, um exemplo de dado inútil seriam as capturas de eventos de pressionamento de botões do teclado ou movimento do mouse. Quando o script termina a classificação e indexação de todos os dados relevantes, é realizada a inclusão desses dados como resultado do experimento do jogador e jogo em que a captura foi realizada.

Após a extração de características, há a mineração das amostras através de análise utilizando scripts autorais. A mineração diz respeito à eliminação de amostras com números nulos/zeros nos sentimentos padrões e nos pontos de detecção facial. A nulidade dos atributos de entrada pode ter ocorrido por distintos fatores, visto que os experimentos foram realizados em ambientes não controlados. Algumas hipóteses são a retirada do rosto do enquadramento da webcam, clarão ou escuridão temporária alterando a luminosidade do ambiente, a webcam ter falhado temporariamente durante o experimento ou ainda a possível queda repentina da conexão com a internet. Ao final, a base de dados autoral emprega 36.264 amostras de tédio e outras 32.035 amostras de estresse.

3.2.2 Clusterização

A literatura aponta grupos de usuários com respostas emotivas completamente distintas entre si. Por exemplo, os nativos digitais processam diferentes informações de forma muito rápida e paralela, o que implica em reações contendo múltiplas emoções, por outro lado os imigrantes digitais costumemente são despreparados para lidar com essas multidimensões [46]. Em termos técnicos, a presente proposta visa analisar as

expressões faciais produzidas pelos jogadores durante o decorrer da partida de modo a criar grupos de jogadores de forma não-supervisionada.

A Clusterização proposta adota a rede neural do tipo SOM (Self Organizing Map - Mapa Auto-organizável). As ligações sinápticas (ajuste de pesos) da rede SOM são baseadas no aprendizado competitivo. O neurônio vencedor é aquele que alcançar menor distância euclidiana à resposta obtida pela rede neural também responsável por organizar automaticamente os dados topologicamente.

Uma das principais vantagens oferecidas pela clusterização é o fornecimento de uma melhor representação dos dados, pois todos os objetos em um cluster têm menos variabilidade em seus atributos e podem ser resumidos com eficiência [47]. Em acréscimo, a clusterização é empregada como mineração de dados visando a segregação de jogadores os quais tiveram algum problema de captura de imagens durante o experimento. Os dados deste trabalho bem como o notebook em Python encontram-se no repositório GitHub no endereço <https://github.com/renanalencar/projeto-md-emocoes> para replicação deste experimento.

4 RESULTADOS

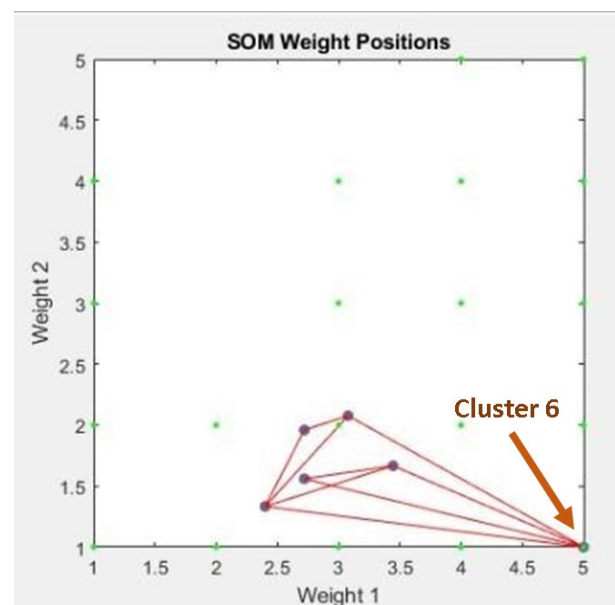
Após o pré-processamento, utilizando a estratégia de busca em árvores binárias iterativas com a linguagem de programação Python, os dados passaram de uma organização em linhas com diferentes tipos de dados agrupados, para uma tabela onde as colunas representavam o tipo de dado e as linhas seu momento de persistência no tempo. Após esse pré-processamento, das 232 participações registradas, apenas 107 mostraram-se válidas de acordo com a metodologia e as técnicas de mineração adotadas.

Esse tratamento simplificou a extração dos dados com as variações das 6 emoções básicas dos jogadores nos momentos de tédio e estresse.

Assim, de posse dos dados com as variações das emoções, que foram extraídos das tabelas, deu-se início ao processo de clusterização utilizando a rede não supervisionada SOM.

Uma das principais vantagens oferecidas pela clusterização é o fornecimento de uma melhor representação dos dados, pois todos os objetos em um cluster têm menos variabilidade em seus atributos e podem ser resumidos com eficiência.

Figura 6 – Resultado da clusterização em 6 grupos por padrões nas expressões faciais – cluster 6, agrupamento dos participantes que apresentaram problemas na realização do experimento.



Fonte: Os Autores.

Assim, além de separar os conjuntos de expressões faciais por padrões, a rede neural realizou mineração não convencional separando os jogadores que apresentaram problemas para completar sua participação no experimento agrupando-os no cluster 6. Esse cluster recebeu as participações com excessivos zeros e nulos nas variações das emoções no decorrer do tempo.

4.1 NORMALIZAÇÃO E CENTROIDES

Os valores das emoções gerados pelo sistema de visão computacional originalmente variam entre 0 e 1. Considerando que nesse contexto as expressões naturalmente teriam valores baixos, por se tratar de microexpressões, iniciou-se a análise dos clusters válidos. Em busca de compreender o comportamento de cada um desses 5 grupos de jogadores, foi calculado um centroide para cada uma das seis emoções em seus respectivos momentos de tédio e de estresse. Os resultados gerados com esse cálculo apresentaram valores muito baixos como previsto. Isso mostrou que seria necessário normalizar esses valores de forma que possibilitasse uma melhor análise dos dados.

Para tanto, foram utilizados os intervalos do simulador de emoções que possuem valores entre

0 e 2000. Assim, a geração dos centroides com os valores normalizados, também poderiam ser aplicados diretamente no simulador.

A **Quadro 1** - Centroides das 6 Emoções Básicas no Momento Classificado como Tédio.

Momento tédio - Centroides das 6 emoções básicas normalizados de 0 a 2000						
Clus ter	Raiva	Nojo	Medo	Triste za	Surpre sa	Felicida de
1	290.17	305.82	117.09	470.84	310.14	432.73
2	607.52	659.03	237.95	344.23	355.31	629.86
3	286.02	307.1	113.2	385.61	263.19	361.05
4	205.78	220.45	222.58	486.02	476.61	441.02
5	273.52	266.75	136.83	566.54	326.43	442.73

Fonte: Os Autores.

Quadro 2 - Centroides das 6 Emoções Básicas no Momento Classificado como Estresse.

Momento Estresse - Centroides das 6 emoções básicas normalizados de 0 a 2000						
Clus ter	Raiva	Nojo	Medo	Triste za	Surpre sa	Felicida de
1	342.02	356.32	146.59	477.39	277.22	409.26
2	556.2	419.33	487.83	376.65	344.55	668.97
3	332.59	297.02	121.27	304.79	212.5	400.84
4	217.76	283.15	211.44	467.63	495.81	414.31
5	396.85	377.83	118.43	466.7	275.03	424.19

Fonte: Os Autores.

O quadro 1 mostra os valores dos centroides de cada emoção dos 5 clusters no momento tédio.

E o quadro 2 apresenta os centroides gerados a partir dos valores das emoções dos 5 clusters no momento classificado como estresse.

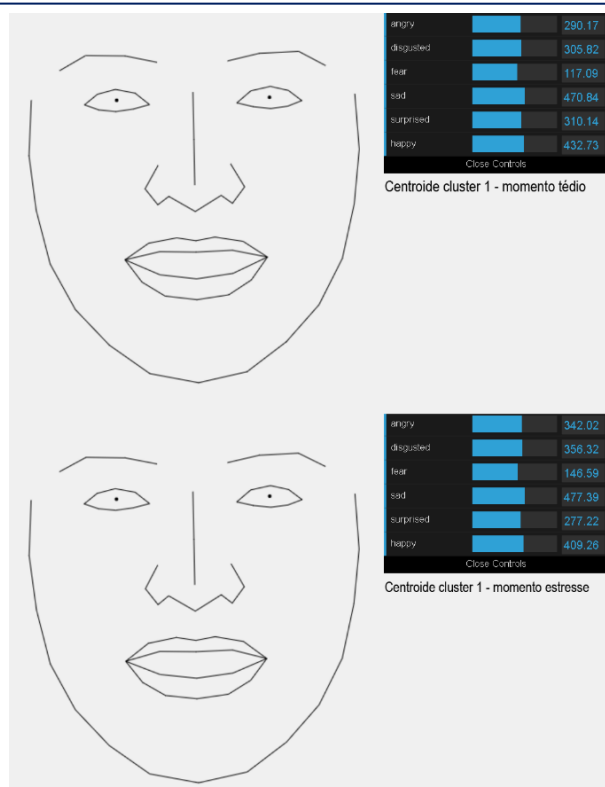
4.2 VISUALIZAÇÃO DOS DADOS

Os valores dos centroides dos clustres foram inseridos no simulador facial para que pudéssemos visualizar suas expressões e inferir sobre seus comportamentos.

O cluster 1 apresenta centroides com maiores valores nas emoções tristeza e felicidade. Seu comportamento demonstra aumento da raiva, do nojo, do medo e da tristeza, e a diminuição da surpresa e da felicidade no decorrer do tempo. Esse perfil de jogador possui valores maiores das emoções surpresa e felicidade enquanto se encontra no momento tédio.

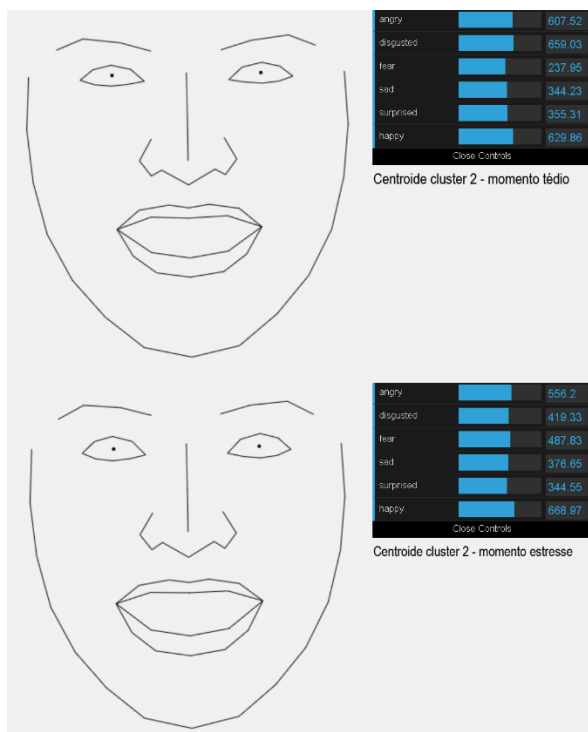
Em geral, o cluster 2 possui os valores mais altos nas emoções felicidade e raiva. Seu comportamento apresenta diminuição da raiva, do nojo e da surpresa, e o aumento do medo, da tristeza e da felicidade com o passar do tempo. Esse perfil de jogador possui os valores mais altos na emoção felicidade enquanto se encontra no momento estresse.

Figura 7 – Simulador de expressões faciais – centroides do cluster 1 nos momentos de tédio e estresse



Fonte: os autores.

Figura 8 – Simulador de expressões faciais – centroides do cluster 2 nos momentos de tédio e estresse



Fonte: os autores.

O cluster 3 possui os valores próximos entre as emoções raiva, nojo, tristeza e felicidade, girando em torno de 300 em uma escala de 0 a 2000. O destaque vai para as emoções com valores mais baixos, medo e surpresa, girando em torno de 121 e 212 no momento de estresse, respectivamente.

No caso do cluster 4 é possível perceber o aumento das emoções nojo e surpresa, e a diminuição da felicidade no decorrer do tempo. Seu maior valor registrado vai para surpresa no momento de estresse.

Finalmente, o perfil 5 apresenta diminuição das emoções medo, surpresa e felicidade, e aumento das emoções raiva, nojo e tristeza no decorrer do tempo. Sua emoção de valor mais alto é o medo no momento de tédio.

Figura 9 – Simulador de expressões faciais – centroides do cluster 3 nos momentos de tédio e estresse

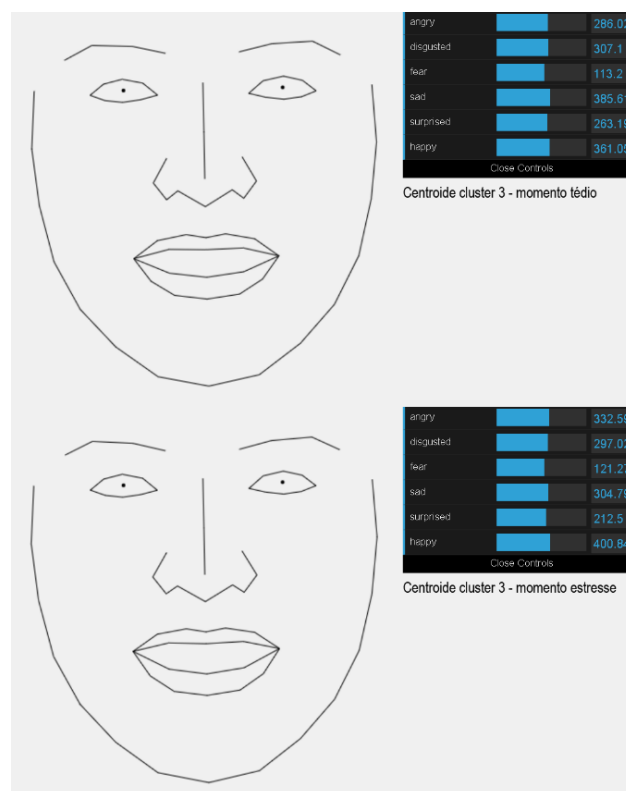
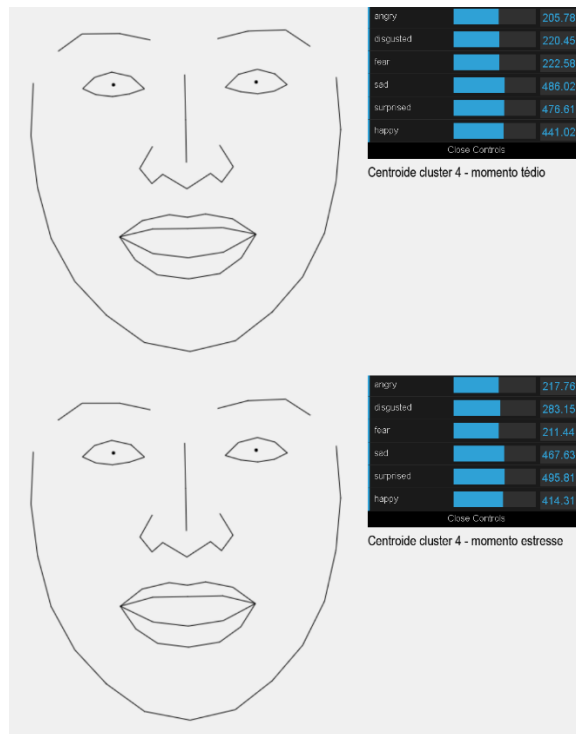


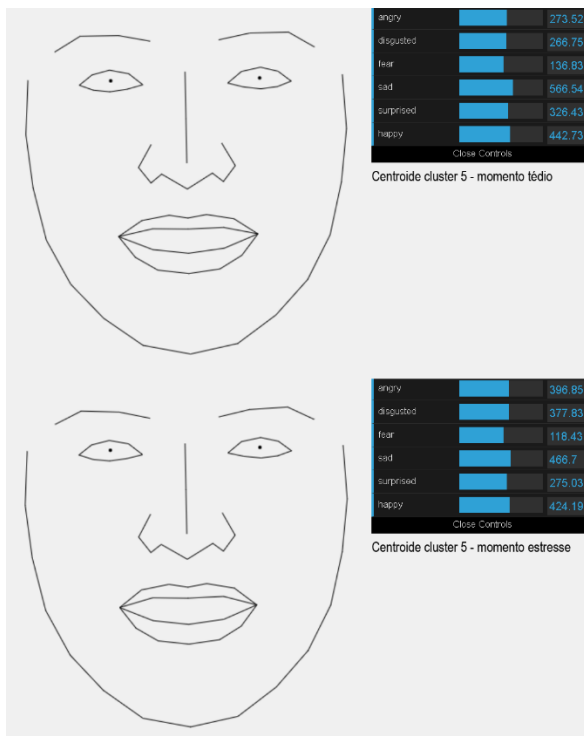
Figura 10 – Simulador de expressões faciais – centroides do cluster 4 nos momentos de tédio e estresse

Fonte: os autores.



Fonte: os autores.

Figura 11 – Simulador de expressões faciais – centroides do cluster 5 nos momentos de tédio e estresse



Fonte: os autores.

5 CONCLUSÃO

Este estudo se propôs a realizar investigação preliminar na base de dados autoral constituída através de experimento realizado em campo, a distância. As tecnologias e métodos utilizados refletem o estado-da-arte na área do estudo das expressões faciais e comportamento dos jogadores de jogos digitais. Através das técnicas de mineração de dados foi possível desvendar 5 perfis de jogadores distintos por meio da análise de suas microexpressões e ainda isolar um conjunto considerável de dados inválidos. Dessa forma, buscou-se correlacionar dimensões da experiência do jogador com o intuito de aprimorar os jogos digitais. Os indícios sobre as definições dos perfis dos jogadores, ainda necessita de aprimoramento, pois as variações em suas expressões faciais se mostraram bastante sutis. Como trabalhos futuros pretende-se continuar a investigação sobre os dados desta base, que se transforma e aumenta periodicamente.

REFERÊNCIAS

- [1] BEVILACQUA, F. H.; ENGSTRÖM; BACKLUND, P. "Game-calibrated and user-tailored remote detection of stress and boredom in games," *Sensors (Switzerland)*, vol. 19, no. 13, Jul. 2019, doi: 10.3390/s19132877.
- [2] COHN, J. F.; DE LA TORRE, F. Automated face analysis for affective. In *The Oxford Handbook of Affective Computing*; Oxford University Press: Oxford, UK, 2014; p. 131.
- [3] TAN, C. T.; BAKKES, S.; PISAN, Y. Inferring Player Experiences Using Facial Expressions Analysis. In *Proceedings of the 2014 Conference on Interactive Entertainment—IE2014, Newcastle, NSW, Australia, 2–3 December 2014*; ACM Press: Boston, MA, USA, 2014; pp. 1–8.
- [4] TAN, C. T.; BAKKES, S.; PISAN, Y. Correlation between Facial Expressions and the Game Experience Questionnaire. In *Proceedings of the Entertainment Computing-ICEC 2014: 13th International Conference, ICEC 2014, Sydney, Australia, 1–3 October 2014*; Springer: Berlin, Germany, 2014; Volume 8770, p. 229.
- [5] ZHOU, X.; HUANG, X.; WANG, Y. Real-time facial expression recognition in the interactive game based on embedded hidden markov model. In *Proceedings of the International Conference on Computer Graphics, Imaging and Visualization, 2004 (CGIV 2004), Penang, Malaysia, 2 July 2004*;

- IEEE: Piscataway Township, NJ, USA, 2004; pp. 144–148.
- [6] ZHAN, C.; LI, W.; OGUNBONA, P.; SAFAEI, F. A Real-Time Facial Expression Recognition System for Online Games. *Int. J. Comput. Games Technol.* 2008, 2008, 10.
- [7] JERRITTA, S.; MURUGAPPAN, M.; NAGARAJAN, R.; WAN, K. Physiological signals based human emotion Recognition: A review. In *Proceedings of the 2011 IEEE 7th International Colloquium on Signal Processing and its Applications*, Penang, Malaysia, 4–6 March 2011; Institute of Electrical and Electronics Engineers (IEEE): Piscataway Township, NJ, USA, 2011; pp. 410–415.
- [8] MANDRYK, R. L.; ATKINS, M. S.; INKPEN, K. M. A continuous and objective evaluation of emotional experience with interactive play environments. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, Montréal, QC, Canada, 22–27 April 2006; ACM: New York, NY, USA, 2006; pp. 1027–1036.
- [9] PICARD, R.W. *Affective Computing*; MIT Press: Cambridge, MA, USA, 2000.
- [10] MEKLER, E. D.; BOPP, J. A.; TUCH, A. N.; OPWIS, K. A systematic review of quantitative studies on the enjoyment of digital entertainment games. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems—CHI'14*, Toronto, ON, Canada, 26 April–1 May 2014; Association for Computing Machinery (ACM): New York, NY, USA, 2014; pp. 927–936.
- [11] BOUSEFSAF, F.; MAAOUI, C.; PRUSKI, A. Remote assessment of the Heart Rate Variability to detect mental stress. In *Proceedings of the 2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops*, Venice, Italy, 5–8 May 2013; pp. 348–351.
- [12] YUN, C.; SHASTRI, D.; PAVLIDIS, I.; DENG, Z. O'game, can you feel my frustration? In *Proceedings of the 27th International Conference on Human Factors in Computing Systems—CHI 09*, Boston, MA, USA, 4–9 April 2009; Association for Computing Machinery (ACM): New York, NY, USA, 2009; pp. 2195–2204.
- [13] RANI, P.; LIU, C.; SARKAR, N.; VANMAN, E. An empirical study of machine learning techniques for affect recognition in human-robot interaction. *Pattern Anal. Appl.* 2006, 9, 58–69.
- [14] TIJS, T. J.; BROKKEN, D.; IJSSELSTEIJN, W. A. "Dynamic game balancing by recognizing affect," in *Fun and Games*, vol. 5294 of *Lecture Notes in Computer Science*, pp. 88–93, Springer, Berlin, Germany, 2008.
- [15] Kukolja, D.; Popović, S.; Horvat, M.; Kovač, B.; Čosić, K. Comparative analysis of emotion estimation methods based on physiological measurements for real-time applications. *Int. J. Hum. Comput. Stud.* 2014, 72, 717–727.
- [16] ZACHARATOS, H.; GATZOULIS, C.; CHRYSANTHOU, Y. L. Automatic emotion recognition based on body movement analysis: A survey. *IEEE Comput. Graph. Appl.* 2014, 34, 35–45.
- [17] YAMAKOSHI, T.; YAMAKOSHI, K.; TANAKA, S.; NOGAWA, M.; SHIBATA, M.; SAWADA, Y.; ROLFE, P.; HIROSE, Y. A Preliminary Study on Driver's Stress Index Using a New Method Based on Differential Skin Temperature Measurement. In *Proceedings of the 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Lyon, France, 22–26 August 2007; Institute of Electrical and Electronics Engineers (IEEE): Piscataway Township, NJ, USA, 2007; pp. 722–725.
- [18] YAMAGUCHI, M.; WAKASUGI, J.; SAKAKIMA, J. Evaluation of driver stress using biomarker in motor-vehicle driving simulator. In *Proceedings of the 2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, New York, NY, USA, 30 August–3 September 2006; Institute of Electrical and Electronics Engineers (IEEE): Piscataway Township, NJ, USA, 2006; pp. 1834–1837.
- [19] HEALEY, J.; PICARD, R.W. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Trans. Intell. Transp. Syst.* 2005, 6, 156–166.
- [20] SCHELL, J. *The Art of Game Design: A book of Lenses*; CRC Press: Boca Raton, FL, USA, 2014.
- [21] SWEETSER, P.; WYETH, P. GameFlow: a model for evaluating player enjoyment in games. *Comput. Entertain.* 2005. DOI=10.1145/1077246.1077253. Disponível em: <<http://doi.acm.org/10.1145/1077246.1077253>> Acesso em 03/05/2021.
- [22] CHEN, J. 2007.Flow in games (and everything else). *Commun. ACM*, p. 31–34. 2007.
- [23] KOSTER, R. *The Theory of Fun for Game Design*. 2a. ed. O'Reilly Media, 2013. p. 300.
- [24] SCHRADER, C.; BRICH, J.; FROMMEL, J.; RIEMER, V.; ROGERS, K. "Rising to the challenge: an emotion-driven approach toward adaptive serious games," in *Serious Games and Edutainment Applications*, pp. 3–28, Springer, 2017. View at: Google Scholar
- [25] HAZLETT, R. L.; "Measuring emotional valence during interactive experiences: boys at video game play," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1023–1026, 2006.

- [26] RAVAJA, N.; SAARI, T.; SALMINEN, M.; LAARNI, J.; KALLINEN, K. "Phasic emotional reactions to video game events: a psychophysiological investigation," *Media Psychology*, vol. 8, no. 4, pp. 343–367, 2006.
- [27] SALAH, A. A.; SEBE, N.; GEVERS, T. "Communication and automatic interpretation of affect from facial expressions," *Affective Computing and Interaction: Psychological, Cognitive and Neuroscientific Perspectives*, pp. 157–183, 2010.
- [28] SAMARA, A.; GALWAY, L.; BOND, R.; WANG, H. "Sensing affective states using facial expression analysis," in *Ubiquitous Computing and Ambient Intelligence, Lecture Notes in Computer Science*, pp. 341–352, Springer International Publishing, 2016.
- [29] CHANG, C.-Y.; TSAI, J.-S.; WANG, C.-J.; CHUNG, P.-C. "Emotion recognition with consideration of facial expression and physiological signals," in *Proceedings of the 2009 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology, CIBCB 2009*, pp. 278–283, April 2009.
- [30] HAMMAL, Z.; COUVREUR, L.; CAPLIER, A.; ROMBAUT, M. "Facial expression classification: An approach based on the fusion of facial deformations using the transferable belief model," *International Journal of Approximate Reasoning*, vol. 46, no. 3, pp. 542–567, 2007.
- [31] ABRANTES, G. A.; PEREIRA, F. "MPEG-4 facial animation technology: survey, implementation, and results," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no. 2, pp. 290–305, 1999.
- [32] TANG, H.; HUANG, T. S. "3D Facial expression recognition based on automatically selected features," in *Proceedings of the 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops*, pp. 1–8, June 2008.
- [33] TANG, H.; HUANG, T. S. "3d facial expression recognition based on properties of line segments connecting facial feature points," in *Proceedings of the Automatic Face & Gesture Recognition, 8th IEEE International Conference on IEEE*, pp. 1–6, 2008.
- [34] HUPONT, I.; BALDASSARRI, S.; CERREZO, E. "Facial emotional classification: from a discrete perspective to a continuous emotional space," *PAA. Pattern Analysis and Applications*, vol. 16, no. 1, pp. 41–54, 2013.
- [35] AKAKN, H. Ç.; SANKUR, B. "Spatiotemporal-boosted DCT features for head and face gesture analysis," in *Human Behavior Understanding*, pp. 64–74, Springer Nature, 2010.
- [36] BARTLETT, M. S.; HAGER, J. C.; EKMAN, P.; SEJNOWSKI, T. J. "Measuring facial expressions by computer image analysis," *Psychophysiology*, vol. 36, no. 2, pp. 253–263, 1999.
- [37] KAISER, S.; WEHRLE, T.; EDWARDS, P. "Multi-modal emotion measurement in an interactive computer game: A pilot-study," in *Proceedings of the VIII conference of the international society for research on emotions*, pp. 275–279, ISRE Publications Storrs, 1994.
- [38] WEHRLE T.; KAISER, S. "Emotion and facial expression," in *Affective Interactions*, vol. 1814 of *Lecture Notes in Computer Science*, pp. 49–63, Springer, Berlin, Germany, 2000.
- [39] GRAFSGAARD, J. F.; WIGGINS, J. B.; BOYER, K. E.; WIEBE, E. N.; LESTER, J. C. "Automatically recognizing facial expression: Predicting engagement and frustration," in *EDM*, pp. 43–50, 2013.
- [40] HEYLEN, D.; GHIJSEN, M.; NIJHOLT, A.; OP DEN AKKER, R. "Facial signs of affect during tutoring sessions," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*: Preface, vol. 3784, pp. 24–31, 2005.
- [41] YAMAKOSHI T.; YAMAKOSHI, K.; TANAKA S. et al. "A preliminary study on driver's stress index using a new method based on differential skin temperature measurement," in *Proceedings of the 29th Annual International Conference of IEEE-EMBS, Engineering in Medicine and Biology Society, EMBC'07*, pp. 722–725, August 2007.
- [42] YAMAGUCHI, M.; WAKASUGI, J.; SAKAKIMA, J. "Evaluation of driver stress using biomarker in motor-vehicle driving simulator," in *Proceedings of the Conference Proceedings. Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1834–1837, New York, NY, August 2006.
- [43] HEALEY, J. A.; PICARD, R. W. "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 2, pp. 156–166, 2005.
- [44] JACK, R. E. "Culture and facial expressions of emotion," *Visual Cognition*, vol. 21, no. 9-10, pp. 1248–1286, 2013.
- [45] SARAGIH, Jason M.; LUCEY, Simon; COHN, Jeffrey F. (2011) Deformable model fitting by regularized landmark mean-shift. *Int J Comput Vis* 91(2):200–215. Disponível em: <https://dl.acm.org/doi/10.1007/s11263-010-0380-4>. Acesso em: 27 abr. 2021 às 14hs50min.
- [46] PRENSKY, M. *Digital Natives, Digital Immigrants, On the Horizon*, Vol. 9 No. 5, pp. 1–6.

<https://doi.org/10.1108/10748120110424816>, 2001.

- [47]** NERURKAR, P. et al. Empirical Analysis of Data Clustering Algorithms, *Procedia Computer Science*. Volume 125, Pages 770-779. doi: <https://doi.org/10.1016/j.procs.2017.12.099>, 2018.