

Affective computing in the context of music therapy: a systematic review

Computação afetiva no contexto da musicoterapia: uma revisão sistemática

Computación afectiva en el contexto de la musicoterapia: una revisión sistemática

Received: 11/04/2021 | Reviewed: 11/11/2021 | Accept: 11/18/2021 | Published: 11/28/2021

Maíra Araújo de Santana

ORCID: <https://orcid.org/0000-0002-1796-7862>
University of Pernambuco, Brazil
E-mail: mas2@ecomp.poli.br

Clarisse Lins de Lima

ORCID: <https://orcid.org/0000-0003-1198-8627>
University of Pernambuco, Brazil
E-mail: cll@ecomp.poli.br

Ariane Sarmiento Torcate

ORCID: <https://orcid.org/0000-0003-2779-873X>
University of Pernambuco, Brazil
E-mail: ast@ecomp.poli.br

Flávio Secco Fonseca

ORCID: <https://orcid.org/0000-0003-4956-1135>
University of Pernambuco, Brazil
E-mail: fsf2@ecomp.poli.br

Wellington Pinheiro dos Santos

ORCID: <https://orcid.org/0000-0003-2558-6602>
Federal University of Pernambuco, Brazil
E-mail: wellington.santos@ufpe.br

Abstract

Music therapy is an effective tool to slow down the progress of dementia since interaction with music may evoke emotions that stimulates brain areas responsible for memory. This therapy is most successful when therapists provide adequate and personalized stimuli for each patient. This personalization is often hard. Thus, Artificial Intelligence (AI) methods may help in this task. This paper brings a systematic review of the literature in the field of affective computing in the context of music therapy. We particularly aim to assess AI methods to perform automatic emotion recognition applied to Human-Machine Musical Interfaces (HMMI). To perform the review, we conducted an automatic search in five of the main scientific databases on the fields of intelligent computing, engineering, and medicine. We search all papers released from 2016 and 2020, whose metadata, title or abstract contains the terms defined in the search string. The systematic review protocol resulted in the inclusion of 144 works from the 290 publications returned from the search. Through this review of the state-of-the-art, it was possible to list the current challenges in the automatic recognition of emotions. It was also possible to realize the potential of automatic emotion recognition to build non-invasive assistive solutions based on human-machine musical interfaces, as well as the artificial intelligence techniques in use in emotion recognition from multimodality data. Thus, machine learning for recognition of emotions from different data sources can be an important approach to optimize the clinical goals to be achieved through music therapy.

Keywords: Affective computing; Emotion recognition; Auditory stimulation; Recommendation system; Music therapy.

Resumo

A musicoterapia é uma ferramenta eficaz para retardar o progresso da demência, uma vez que a interação com a música pode evocar emoções que estimulam as áreas do cérebro responsáveis pela memória. Essa terapia é mais bem-sucedida quando o terapeuta fornece estímulos adequados e personalizados para cada paciente. Essa personalização costuma ser difícil. Assim, métodos de Inteligência Artificial (IA) podem auxiliar nessa tarefa. Este artigo traz uma revisão sistemática da literatura da área de computação afetiva no contexto da musicoterapia. Em particular, pretendemos avaliar métodos de IA para realizar o reconhecimento automático de emoções aplicados a Interfaces Musicais Homem-Máquina (HMMI). Para realizar a revisão, realizamos uma busca automática em cinco das principais bases de dados científicas nas áreas de computação inteligente, engenharia e medicina. Procuramos todos os artigos publicados entre 2016 e 2020, cujos metadados, título ou resumo contenham os termos definidos na string de pesquisa. O protocolo de revisão sistemática resultou na inclusão de 144 trabalhos das 290 publicações retornadas da pesquisa. Através desta revisão do estado da arte, foi possível elencar os desafios atuais no reconhecimento automático de emoções. Também foi possível perceber o potencial do reconhecimento automático de emoções para construir soluções assistivas não invasivas baseadas em interfaces musicais homem-máquina, bem como as técnicas

de inteligência artificial em uso no reconhecimento de emoções a partir de dados multimodais. Assim, o aprendizado de máquina para reconhecimento de emoções de diferentes fontes de dados pode ser uma abordagem importante para otimizar os objetivos clínicos a serem alcançados por meio da musicoterapia.

Palavras-chave: Computação afetiva; Reconhecimento de emoções; Estimulação acústica; Sistema de recomendação; Musicoterapia.

Resumen

La musicoterapia es una herramienta eficaz para ralentizar el progreso de la demencia, ya que la interacción con la música puede evocar emociones que estimulan las áreas del cerebro responsables de la memoria. Esta terapia tiene más éxito cuando el terapeuta proporciona estímulos adecuados y personalizados para cada paciente. Esta personalización suele ser difícil. Por lo tanto, los métodos de Inteligencia Artificial (IA) pueden ayudar en esta tarea. Este artículo trae una revisión sistemática de la literatura en el campo de la computación afectiva en el contexto de la terapia musical. En particular, nuestro objetivo es evaluar los métodos de inteligencia artificial para realizar el reconocimiento automático de emociones aplicado a las interfaces musicales hombre-máquina (HMMI). Para realizar la revisión, realizamos una búsqueda automática en cinco de las principales bases de datos científicas en los campos de la computación inteligente, la ingeniería y la medicina. Buscamos todos los artículos publicados entre 2016 y 2020, cuyos metadatos, título o resumen contengan los términos definidos en la cadena de búsqueda. El protocolo de revisión sistemática resultó en la inclusión de 144 trabajos de las 290 publicaciones devueltas de la búsqueda. A través de esta revisión del estado del arte, fue posible enumerar los desafíos actuales en el reconocimiento automático de emociones. También fue posible darse cuenta del potencial del reconocimiento automático de emociones para construir soluciones de asistencia no invasivas basadas en interfaces musicales hombre-máquina, así como las técnicas de inteligencia artificial que se utilizan en el reconocimiento de emociones a partir de datos multimodal. Por lo tanto, el aprendizaje automático para el reconocimiento de emociones a partir de diferentes fuentes de datos puede ser un enfoque importante para optimizar los objetivos clínicos que se deben lograr a través de la musicoterapia.

Palabras clave: Computación afectiva; Reconocimiento de emociones; Estimulación acústica; Sistema de recomendación; Musicoterapia.

1. Introduction

1.1 Motivation

With the aging of the world population, several countries have been facing considerable changes in the last decades (UN, 2020; EC, 2020). The elderly population has increased, and so is the prevalence of diseases associated with old age, such as osteoporosis, hypertension and dementia (NRC, 2001; WHO, 2019, 2017; Ricci, 2019). Alzheimer's Disease and cerebrovascular ischemia are the two most important causes of dementia worldwide (Nichols et al., 2019; Rizzi et al., 2014).

Several studies show that music therapy has the ability to slow down the progress of dementia through musical stimuli and music education (Gallego & Garcia, 2017; de Souza et al., 2017; Fang et al., 2017; King et al., 2019). Interaction with music stimulates the areas of the brain responsible for memory through emotions. However, the effectiveness of music therapy is closely related to the recognition and correct stimulation of emotions in the patient by the therapist. This task of recognizing emotions is often arduous, especially for less experienced professionals.

Artificial Intelligence (AI) algorithms have been shown to be effective in solving complex classification problems, including the ones related to emotion recognition (Poria et al., 2017; C.-H. Wu et al., 2016; Cambria, 2016). Thus, emotion recognition performed by AI methods has the potential to contribute to the construction of an interface capable of helping music therapists in determining musical genres, styles and rhythms. This interface may be of great help to optimize intervention in the context of the treatment of Alzheimer's disease and mild cognitive impairment.

Thus, this document proposes a systematic review of the literature in the field of Computational Intelligence applied to emotion recognition in Electroencephalography (EEG), voice signals and facial expression. From this review, we particularly aim to identify AI methods to perform automatic emotion recognition applied to human-machine musical interfaces. Particularly, this review seeks to answer the following research questions: (1) What are the current challenges in automatic emotion recognition? (2) How useful is automatic emotion recognition for non-invasive assistive solutions based on human-machine musical interfaces? (3) Which AI techniques are being used in emotion recognition? (4) Which deep networks

architectures are used for music recommendation based on emotion recognition from multimodalities data?

Next topic presents some background information regarding to aging, cognitive impairment, affective computing, and music therapy. In Methodology section we describe the systematic search protocol. Then, the results are showed and discussed. Finally, in the last section we highlight some conclusions and limitations of this work.

1.2 Background

This topic presents the main theoretical references which are the basis for carrying out and understanding this research. Thus, we explore the concept of cognitive deficits, focusing on Alzheimer's Disease. Then, we give a brief introduction about emotions. We point out emotion recognition, considering the forms and stimuli for identification, emphasizing the field of study of Affective Computing. Finally, we present a brief review about therapies and interventions. Additionally, we discuss how they are being assisted by technologies and enhanced with musical techniques.

To introduce the discussion concerning cognitive deficits and dementia, some aspects need to be outlined. It is important to clarify that cognition refers to the way our brain learns, perceives, remembers, and also processes the information absorbed by the different senses (Sheffield et al., 2018). Therefore, cognitive deficit is characterized as an obstacle in this development, especially with respect to learning and intellectual limitations.

Also known as mild cognitive impairment, cognitive deficit can cause slight loss of memory, attention and difficulties in logical reasoning. However, it does not present dementia (Hamdan, 2008). In the elderly, these symptoms are often confused with natural causes of aging. As yet, these are signs that alert both individuals and their families, since in a few cases, they can evolve or be the beginning of Alzheimer's disease (AD) in its early stages (Pais et al., 2020).

Despite the similarities, the differentiation between cognitive decline and AD is essential to avoid confusion between both conditions. AD is the most common type of dementia and is also a general term used to describe conditions that occur when the brain can no longer function properly. In the literature, AD is classified as a chronic, degenerative, and progressive dementia (Lourinho & Ramos, 2019). According to Caetano et al. (2017), neurodegenerative diseases are those which cause irreversible neuron degeneration. According to research carried out by Bertazone et al. (2016), Alzheimer's Disease is characterized by changes in memory, nonetheless, it is rarely the most detectable symptom. Failures in cognition, motor skills and language can also appear as the first symptoms and tend to worsen as the disease progresses.

Although there is no cure for AD, other forms of intervention were also developed to promote an improvement on the Alzheimer's patient's quality of life (Caetano et al., 2017). In order to contemplate the individual in a holistic way, the treatment requires a multidisciplinary team - with an interdisciplinary approach - to combine pharmacological and non-pharmacological measures. de Souza et al. (2017) cites psycho-corporal and biological therapies, and music therapy as examples of these interventions. Nonetheless, several other approaches have also been emerging in the field of affective computing and computational intelligence.

Defining emotions is not a trivial task, as this term is frequently used in different contexts, and it is also present in everyday situations (Paxiuba & Lima, 2020). It is worth emphasizing that emotions play an essential role in the social formation of any human being. Studies clarify that emotional expressions are composed of variables that can be directly related to cognitive aspects (Le & Provost, 2013; Izard, 1977). Hence, emotions can manifest in different ways in each individual which include sensations, facial expressions, and body movement. For this reason, emotions are one of the most important experiences, because they guide choices, motivations, decisions, among other aspects. Additionally, they are essential for the process of verbal and non-verbal communication (Marosi-Holzberger et al., 2013; Dorneles et al., 2020).

The literature distinguishes emotions in, at least, two different categories (Scardua & Marques, 2018; Oliveira & Jaques, 2013; Ekman & Friesen, 1971). The first one is the universal or basic emotions defined by joy, sadness, anger, fear,

disgust, surprise, amazement. The second category comprises the social emotions such as guilt, shame, arrogance, and wonder. As already mentioned, for each emotion, there is a definition and a peculiar way of manifesting itself in each individual.

When it comes to recognizing these emotions, humans are often able to feel and/or identify each other's emotional state. This detection may be naturally noticeable to people, but it is still a difficult task for computers. It is in this specific context that we highlight a sub-area in Artificial Intelligence which studies emotions in computers, called Affective Computing (AC) (Paxiuba & Lima, 2020).

The term "Affective Computing" was proposed by Rosalind Picard in 1997 (Nalepa et al., 2019), and refers to a field of research that is totally interdisciplinary with other areas of knowledge (such as Biomedical Engineering, Psychology, and Computer Science). This field of research seeks to develop computational and emotion recognition methods for diverse purposes and applications. Briefly, AC studies how computers can recognize, model and express emotions (and other human psychological aspects), and how they respond to them (Picard, 1997).

In this scenario, there are several ways to investigate the recognition of emotions. But for that to happen, data must be used. According to González & McMullen (2020), data can come from different sources, such as voice, facial expressions, and physiological signals.

For the emotion classification through voice signal, cultural and language patterns can also be considered to the model development. Some attributes are considered more relevant, and they are well established in this type of prediction. Among them, we can mention pitch (Sondhi, 1968), energy or intensity (Ingale & Chaudhari, 2012), formants (Goudbeek et al., 2009), mel-cepstral frequency coefficients (MFCC) (Han & Chan, 2006) along with the characteristics common to most waveforms.

In the literature referring to audio recognition, Convolutional Neural Networks (CNN) (Zhao et al., 2019), Support Vector Machines (SVM) (Sonawane et al., 2017), and Generative Adversarial Networks (GANs) (Chatziagapi et al., 2019) are quite recurrent. They are considered the leading models for this type of classification. It is also worth emphasizing the popularity of the MFCC as one of the most used attributes. According to Han & Chan (2006), the MFCC is a parametric representation of the frequency spectrum of the voice signal. It is a scale very close to the human auditory system, which has a non-linear behavior in frequency.

Another way to perform human affective recognition is through its physiological signals. From the Peripheral Nervous System, the most evaluated signals are: the Galvanic Skin Response (GSR), the Respiration Range, the Skin Temperature, and the Electrocardiogram (ECG). From the Central Nervous System, the signal evaluated is the Electroencephalogram, commonly known as EEG. In this type of approach, models which combine the analysis of more than one signal tend to achieve better performances. Knowing that physiological and cognitive responses can be significantly impacted by emotions (Brosch et al., 2013), combined classifiers that use multimodal analysis become very reliable for the identification of these sensations. They are also an alternative to voice and image's recognition methods.

Another popular method for emotion recognition is through facial expressions (Y. Wang & Kosinski, 2018). This method is somehow more direct, intuitive and easy to identify. AI models based on this optics almost always have attributes that can be visually identified, such as dark hair, and light eyes. However, these features are extremely difficult to standardize and to extract. Thus, studies that take into account the varieties of physiognomy, age, gender, and different demographic groups are a trend among new researches (Buolamwini & Gebru, 2018).

AI models focused on images and videos have a high computational cost. This is due to their large amount of data throughout various pre-processing stages or classification. However, these deficits can be offset by better performance and accuracy for day-to-day applications (Jeong & Ko, 2018), whereupon other approaches would not be feasible or practical.

Lastly, it is worth taking into account the fact that emotions are brief psychophysiological phenomena. Thereby, collecting data on emotions spontaneously is still a challenging task. For this reason, it is often necessary, in research

environments, to place the individual in situations with the purpose of evoking certain emotions. The emotions can be evoked through (Meska et al., 2020) odors, visual or auditory stimuli. The visual and auditory stimuli can be images, videos, songs (Vicencio-Martínez & Garay-Jiménez, 2019), affective scenes using Virtual Reality (T. Xu et al., 2019; Teo & Chia, 2018), and others.

Considering what has been exposed so far, it is important to contextualize them in interdisciplinary fields which include music, education, health and technology. These fields, when combined, have potential for therapeutic applications and diversified interventions. The interventions include games, affective entertainment, and health rehabilitation assisted by Virtual Reality (Bulagang et al., 2021). Music, for example, is a resource that has the ability to stimulate and develop the brain (Kirana et al., 2018). It has a variety of applications for therapeutic purposes such as emotion regulation, social interaction (Agres et al., 2021), and motor functions rehabilitation (Dechenaud et al., 2019). Music can also be applied in physical therapy. Since the exercises are often monotonous and repetitive for the patients, the music can increase their motivation and involvement in the section (Colombo et al., 2019).

With regard to technologies for therapeutic purposes, we highlight games, which add entertainment to the therapy moment, and improve patient engagement (Fonteles et al., 2018). This is possible because games offer rewards and different levels of challenges. These characteristics are advantageous because they involve the individuals in the rehabilitation process. Simultaneously, they increase the individual's motivation during the exercises (Agres et al., 2021). Games can have different classifications depending on the goals for the rehabilitation process. For example, games like Exergames aim to improve the health and well-being of elderly people through stimuli to perform physical activities (Crespo et al., 2016). Another category is the serious games. According to Agres et al. (2021), serious games can be an alternative to train the skills of patients with Parkinson's Disease. It can also assist in treating traumatic brain injury, and in training for patients with dementia.

Virtual reality (VR) applications are pointed out by therapists as a high degree of immersion approach and as a non-threatening method for patients to practice multisensory integration in a real context (Lubetzky et al., 2019). As shown in the research carried out by Geraets et al. (2021), VR can make therapies more available and cost-effective for a larger group of patients. In addition, they can increase the intensity of treatment through home exercises that complement face-to-face therapies. The authors emphasize that with VR, innovative strategies can not only improve therapeutic interventions, but can also be used to investigate mechanisms involved in the persistence of mental disorders. As in games, VR applications are also diverse. As it is a more engaging and realistic approach involving and realistic, it contributes to senses and thought stimulation (T. Xu et al., 2019). These stimuli in VR can even be customized according to the purpose of rehabilitations, as in the study by Cameirão et al. (2017), where the stimuli were customized for the rehabilitation of stroke survivors with mild cognitive impairment (MCI).

Still in the context of technologies, therapies, and interventions, affective and social robotics emerges as an opportunity to carry out personalized therapies (Agres et al., 2021). These technologies can contribute to the rehabilitation process becoming less dependent on therapeutic expertise (Kikuchi et al., 2018). Thus, reducing the burden on the physicians who perform the interventions (Agres et al., 2021). In general, a robotic system be used for many purposes. It can be a bridge for interaction with autistic children, improving their social skills (HAN et al., 2018), assist in treatment of children with cancer (Ranjkar et al., 2019), contribute with wrist (English & Howard, 2017a) and upper limb (Kikuchi et al., 2018) rehabilitation. Furthermore, it can contribute to therapeutic interventions in order to support people with visual and intellectual disabilities (Wingerden et al., 2020).

It is worth reiterating that the studies mentioned in this section confirm that applications and inclusion of games, VR or robotics for therapeutic purposes - when integrated with musical strategies - can be a powerful tool. In fact, music makes all the difference in the rehabilitation process, and also influences the user's affinity with the systems. For this reason, the search

for musical solutions for applications in healthcare grows continuously. A topic that cannot be discarded is the importance of properly trained professionals to assist in the development of these technologies so that their effectiveness is truly beneficial.

2. Methodology

This systematic review selected primary studies based on keywords, search period, and both inclusion and exclusion criteria. As sources, we chose five of the main scientific databases on the fields of intelligent computing, engineering, and medicine (i.e., IEEE Xplore, MedLine/PubMed, SCOPUS, Science Direct and Springer Link). The systematic review was conducted with automatic search in the databases using keywords and period as filters. Particularly, the scope of this study was to focus on the methodological aspects of state-of-the-art works, especially issues related to the most explored computational and artificial intelligence methods.

Therefore, we search all papers released from 2016 to 2020, whose metadata, title or abstract contains the terms defined in the following search string: (“Artificial Intelligence” OR “Deep Learning” OR “Machine Learning” OR “Computational Intelligence” OR “Neural Network” OR “Deep Kernel”) AND (“Electroencephalography” OR “EEG” OR “Neural Signals” OR “Brain Signal”) AND (“Voice Signals” OR “Speech”) AND (“Emotion Recognition” OR “Recognition of Emotion” OR “Affective”) AND (“Alzheimer” OR “Dementia” OR “Degenerative Disorder” OR “Neurodegenerative Disease” OR “Neurodegenerative Disorder” OR “Cortical Disorder”) AND (“Human-Machine Musical Interfaces” OR “Affective Music” OR “Music Therapy” OR “Music Biofeedback”).

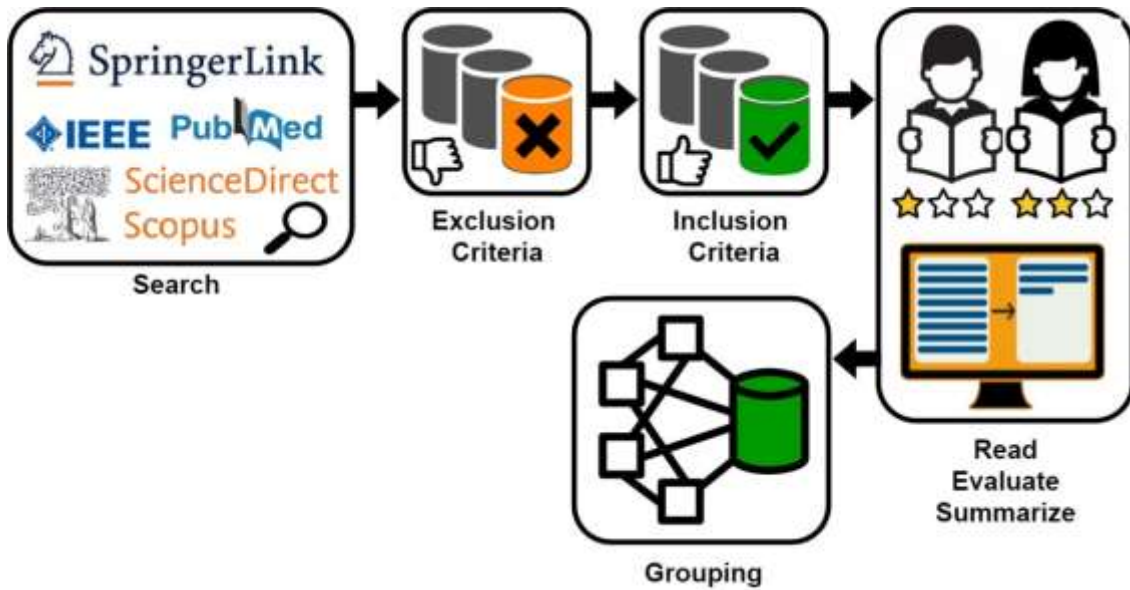
Article selection was performed in five phases (Figure 1). In the first phase, we identified the number of papers from each scientific database. Then, in the second step, we assessed its suitability to all four exclusion criteria (EC). As EC, we exclude duplicated works, studies that do not include some computational tool, documents classified as poster, tutorial, editorial, book, annals or reviews, and studies based on invasive techniques.

The third phase consisted in an evaluation according to introduction and conclusion, in order to select studies that meet at least one of the following inclusion criteria (IC): IC1) studies with computational tools applied to EEG signals; IC2) studies with computational methods applied to voice signals; IC3) studies with computational tools applied to physiological data (Galvanic Skin Response, Respiratory Frequency, Heart Rate, Electrocardiogram, Electrooculogram); IC4) studies with computational tools applied to image or video processing; IC5) studies that use emotion recognition in EEG signals; IC6) studies with emotion recognition in voice signals; IC7) studies with emotion recognition in physiological signals; IC8) studies that use emotion recognition in images or videos; IC9) studies that use Human-Machine Musical Interfaces (HMMI).

After reading the remaining papers, in the fourth phase, we assigned scores to each of them based on the quality checklist in Table 1. To assign these quality scores, the articles were carefully read and evaluated for their formal aspects. Thus, each article was assigned a score for each question presented in Table 1. The score could be 0 (zero) if the text did not meet the criterion, 0.5 if partially met, or 1 if the criterion was well presented in the text. Finally, in the fifth step, we group the studies according to their content.

Finally, to summarize the results from each article, we extracted the following information: (a) Article name; (b) Name of authors; (c) Year; (d) Institution; (e) Database; (f) Application of the study; (g) Methodology procedures; (h) AI classification methods; (i) Phase of study development (Concept Design, Testing, Maintenance of Commercial Product); (j) Results; (k) Advantages of the method; (l) Disadvantages of the method.

Figure 1: Search phases.



Source: Authors.

Table 1: Quality Criteria (QC) to evaluate the selected studies.

ID	Quality Criteria	Answer
QC1	Is there a description of the application domain in which the search was performed?	YES (Y) = 1.0; PARTIALLY (P) = 0.5; NO (N) = 0.
QC2	Are the objectives clearly stated?	Y/P/N
QC3	Are the methodology procedures, such as EEG or audio signals acquisition methods, databases or development environment tools clearly expressed?	Y/P/N
QC4	Is there a description of the sources of information?	Y/P/N
QC5	Are the computational tools, such as Computational Intelligence techniques detailed explained? Including architectures parameters and/or equations.	Y/P/N
QC6	Are the future works or contributions clearly described?	Y/P/N
QC7	Does the study exposes the main challenges and limitations of the proposed method?	Y/P/N

Source: Authors.

3. Results and Discussion

Four people performed papers assessment. Figure 2 shows all evaluation phases and the amount of articles included and excluded in each one. The search process resulted in 189 publications in IEEE Xplore database, 1 in MedLine/PubMed, 38 in SCOPUS, 19 in Science Direct, and 43 in Springer Link. The interdisciplinary nature of this work required searching on several different bases. However, the chosen databases have different emphases and include works from diverse areas of knowledge. Thus, it is natural that the search returns different amounts of papers in each database. From these works, 122 met at least one of the exclusion criteria and were removed. The 168 remaining studies went through title, abstract, and keywords evaluation, resulting in 24 excluded and 144 accepted works. Thus, from the 290 studies found by the search string, 144 were

actually included in this review.

It is important to mention that during search process, we needed to reduce the length of the search strings used in Science Direct and Springer Link, since these bases did not support the original amount of Boolean operators. Thus, for Science Direct we used the terms: (“EEG”) AND (“Speech”) AND (“Emotion Recognition”) AND (“Dementia”) AND (“Music”). In Springer Link, we used: (“Machine Learning”) AND (“EEG”) AND (“Speech”) AND (“Emotion Recognition”) AND (“Dementia”) AND (“Music”).

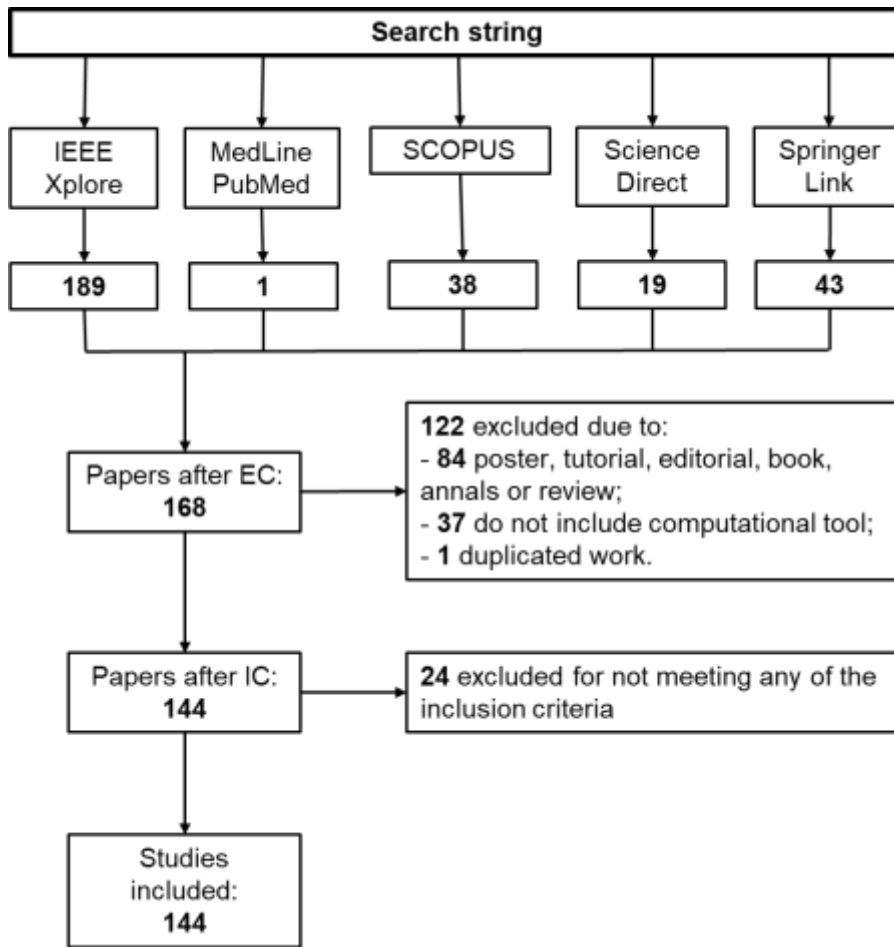
In the last phase of the review, we grouped the selected studies according to the division shown in Figure 3. The studies were combined in six groups: Games, Virtual Reality, Affective Robotics and Therapies (G1), Physiological and behavioral responses induced by acoustic signals (G2), Emotion recognition associated with acoustic stimuli (G3), Non-invasive assistive solutions based on human-machine musical interfaces (G4), Musical composition, recommendation and customization (G5), and other approaches (G6). This last group consists of some studies that does not fit any of the other groups.

As shown in Figure 3, the minority of the studies (8%) are in the 4th group, regarding to non-invasive assistive solutions based on human-machine musical interfaces. G1 and G3 together concentrate almost half the amount of studies (47%). Thus, its possible to see that in the last years many studies are being conducted in the areas of games, virtual reality, affective robotics, therapies, and emotion recognition associated with acoustic stimuli. 19% of the papers are related to physiological and behavioral responses induced by acoustic signals (G2). Finnaly, both G5 and G6 have 13% of the studies each.

All selected studies were evaluated according to the previously defined quality criteria. Then, we calculated their average quality score (AQS) as the mean value of the scores from all criteria. Figure 4a shows the average score of all quality criteria for each group. Most of the studies had good performance in describing their goals and methodology, however, few studies provided information regarding to their contribution and limitations. The AQS for all groups is found in Figure 4b. Overall, all groups showed AQS above 0.7. The studies on G3 achieved the best quality scores while less quality papers were found in G5. This may indicate that more qualified papers are related to studies on emotion recognition associated with acoustic stimuli. However, there are still few relevant studies on customization of musical content.

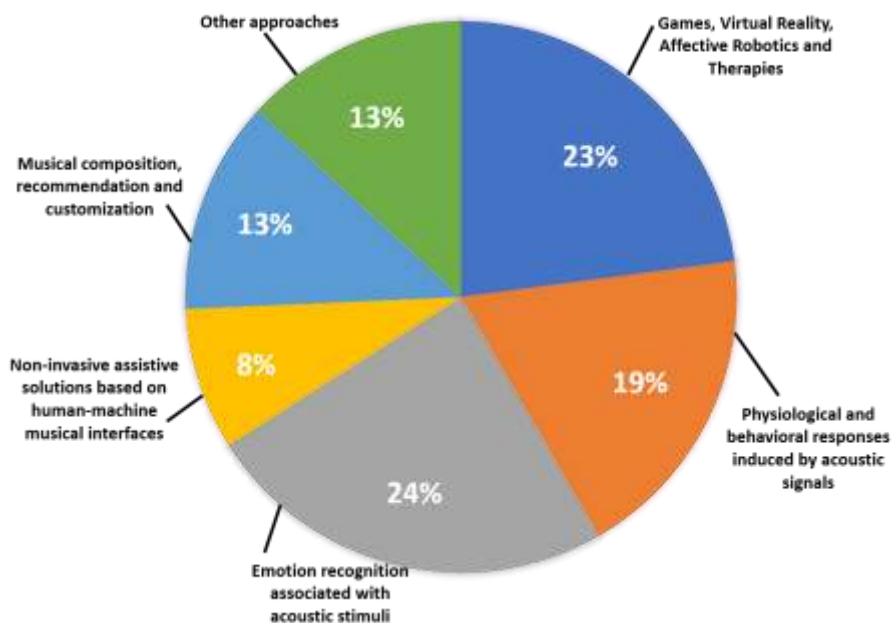
Regarding to the year of publication, Figure 5 shows the the distribution of papers published over the last five years for each group. We observe an increase in the number of studies in the last two years (2019 and 2020) specially in groups G2 and G3. This increase shows a growing interest in the effects of music and other acoustic stimuli in the human being.

Figure 2: Amount of works for each selection phase.



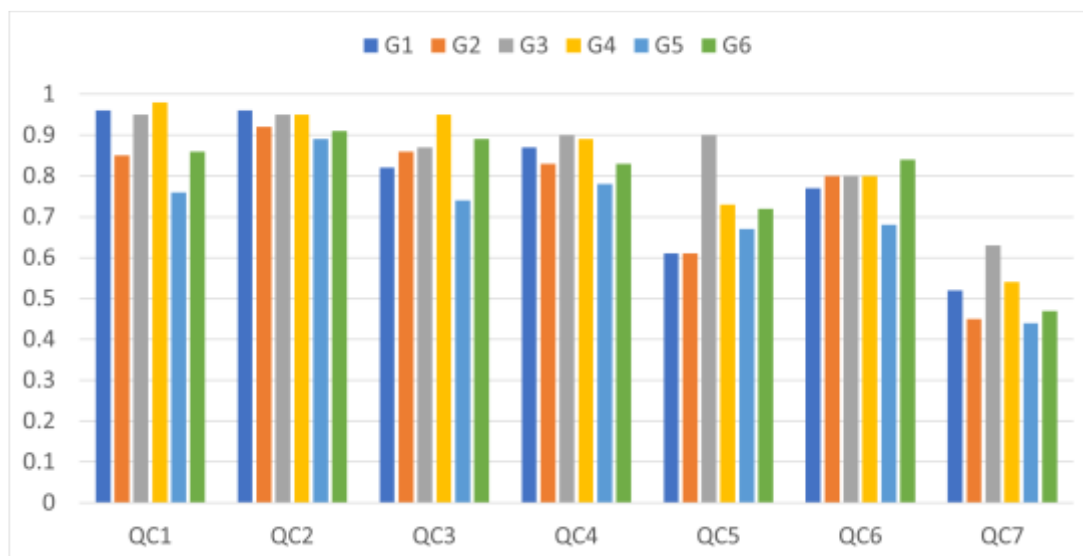
Source: Authors.

Figure 3: Groups distribution of the selected studies.

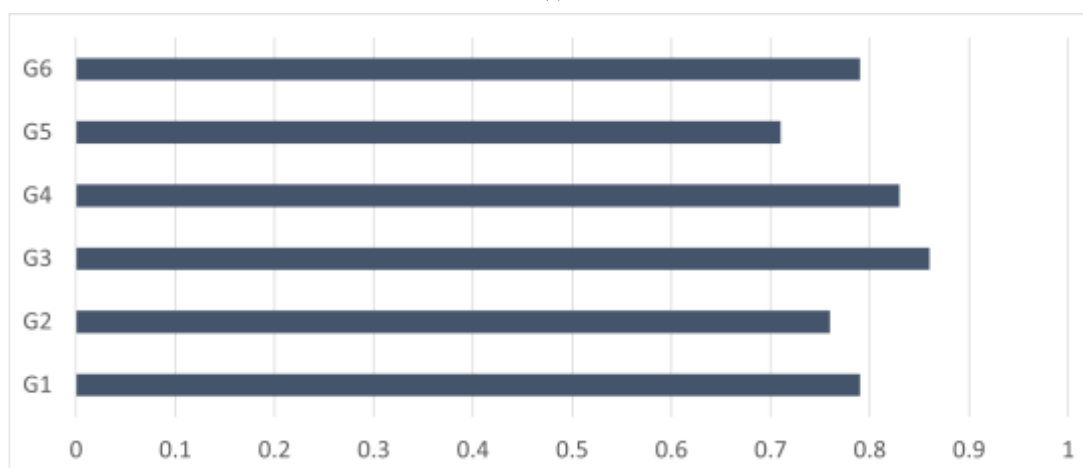


Source: Authors.

Figure 4: Quality analysis of the studies from each group. In (a) are the average score for each quality criteria and (b) shows the average quality score for all groups.



(a)



(b)

Source: Authors.

Figure 5: Year distribution of the articles in each group. The amount of studies per year for G1, G2, G3, G4, G5 and G6 are respectively shown in (a), (b), (c), (d), (e), and (f).



Source: Authors.

3.1 Games, virtual reality, affective robotics and therapies

In the present study, it was possible to identify that music brings several health benefits and contributes to the healthy development of the brain. In addition to music, we identified other gamified approaches such as Games, Virtual Reality (VR), and Affective Robotics. Such approaches were proposed by the articles listed in Table 2. All the aforementioned approaches, when applied to therapies for rehabilitation or treatment purposes, show excellent therapeutic potential. This occurs because they contribute to the patient's involvement in the exercises and contribute to increase their motivation.

It is evident that emotion recognition approaches mostly use music, videos, and images (static and dynamic) to stimulate emotions in individuals. Yet, studies already use affective scenes in Virtual Reality to stimulate and to induce emotions. Then, they perform emotion recognition methods through electroencephalographic signals (T. Xu et al., 2019; Teo & Chia, 2018). In another case, the authors use VR scenes along with music as stimuli to induce emotional responses in individuals (Sra et al., 2017). It emphasizes that, although VR were not designed for this purpose, it stands out positively when compared to traditional approaches. One of the reasons is because, with VR, the individuals involvement can be higher since

the VR is able to bring them closer to the reality in an immersive way. These aspects mentioned can be more promising when combined with music.

Affective virtual spaces have gained prominence for VR applications. It covers areas ranging from education, entertainment, art, well-being to health. Hence, VR applications are being widely used for therapeutic purposes. For example, to assist with rehabilitation of people with vestibular disorder (Lubetzky et al., 2019), and rehabilitation of stroke-affected people with mild cognitive impairment (Cameirão et al., 2017). Other applications include: psychological relaxation therapies (Zhu et al., 2019), and support to the Emotional self-awareness (Bermúdez i Badia et al., 2019). These therapies use audiovisual scenes with music which induces relaxation, and feedback with VR oriented affective states, respectively. Thus, considering the aforementioned therapies, researches are working towards the development of tools which assesses the user experience. Both the therapeutical context, and the individuals affective state are taken into account (Krüger et al., 2020).

Regarding to therapies, musical sonification appears as a potential technique. It has been helping in treatments and rehabilitation. The gamified approaches which have been developed involves adapted guitar for upper limbs rehabilitation for stroke-affected individuals (Dechenaud et al., 2019). It also includes gesture recognition tools for motor rehabilitation (Murad et al., 2017). Researches have also investigated and presented good results for therapies based on vibro-acoustics. They appraised its effects on physiological signals of young people and adults. The authors evaluated the biosignals to assess the therapy potential to induce a relaxation state after a stress situation (Delmastro et al., 2018; Cavallo et al., 2020).

Given the advances in therapies that use musical strategies, different applications arise. Among the application we exemplify: (1) the development of a system to measure the reaction of autistic children during music therapies, considering parameters such as frequency and intensity of sound; (2) protocol validation for training and assessment of manual function rehabilitation; (3) development of intelligent systems to detect stress through EEG signals, leading individuals to relaxation, (4) and music therapy applications to aid in memory recovery.

Physical therapy exercises often can cause discomfort, boredom or fatigue on patients. This is due to repetitive and monotonous activities. Games with musical techniques and VR applications have become an alternative strategy to involve patients on their activities. Musical games stand out for some contributions such as (1) in the promotion of the upper limb motor rehabilitation on stroke-affected individuals (Sanders et al., 2020); (2) in the motor and cognitive skill's exercises by touching sequences (English & Howard, 2017b); (3) and in hand exercises through the process of learning rhythms and melodic structures (Fonteles et al., 2018).

Exergames games based on VR are being used to treat low back pain (Ortegon-Sarmiento et al., 2020), and to assist in therapies for the elderly encouraging arm movement (Crespo et al., 2016). Also in the context of games, other approaches can be highlighted such as (1) the affective adaptation of games to motor rehabilitation by recognizing emotions by facial expressions (Aranha et al., 2017) (2) modulation of affective and physiological states of individuals using VR-based musical games to stimulate mental activation and relaxation (Arroyo-Palacios & Slater, 2016) (3) recognition of frustrations caused by games and their interfaces through multimodal signal analysis (Song et al., 2019) (4) learning components for a digital opera composition through a gamified app (Scott, 2017).

The use of affective robotics has also gained a lot of acceptance and prominence due to its diverse uses. It has been used mainly for therapeutic purposes, for example, robotic systems for upper limbs and wrists' motor rehabilitation using musical integration (Kikuchi et al., 2018; English & Howard, 2017a). Affective robots has also been used in health interventions in order to to assist individuals with intellectual disabilities (Shukla et al., 2019). Musical and social robots are also adopted for interaction with autistic children (HAN et al., 2018), and children undergoing cancer treatment (Ranjkar et al., 2019), respectively.

Affective robots using musical techniques report success in the rehabilitation process. These are being developed to

act as personal companions with a focus on the mood of individuals. They can also interact through scents, and music to generate positive emotions (Lui et al., 2017). Affective robots can also interact as robot dogs to cooperate in music therapies and identification of pneumonia in patients suffering from dementia (Lyu & Yuan, 2020). Moreover, they are able to interact as a singing robot in order to communicate through music, and evoke and assess emotional responses in humans (Wolfe et al., 2020).

Finally, it is important to point out that only 9 studies, among the 33 studies mentioned in this subsection, used Artificial Intelligence techniques (T. Xu et al., 2019; Lui et al., 2017; Aranha et al., 2017; Song et al., 2019; Murad et al., 2017; Cavallo et al., 2020; Teo & Chia, 2018; Lyu & Yuan, 2020; Sra et al., 2017). The mostly used techniques were Support Vector Machines (SVM), Random Forest, Convolutional Neural Networks (CNN), Decision Trees and K-Nearest Neighbor (KNN). Furthermore, we observed that a significant portion of 22 of the 33 studies are in the experimental phase. Six of them are in prototyping stage, three of them are in concept design stage, and only two in the validation stage.

3.2 Physiological and behavioral responses induced by acoustic signals

From this study, it we observed that almost 19% of the selected works deal with physiological and/or behavioral responses induced by sound signals (Table 3). In addition, we noted that there has been an increase in the number of publications involving this topic over the years, as most articles were published over the years 2019 and 2020. Only 9 out of the 27 works in this area were published between the years 2016 and 2018. This demonstrates a growth in scientific interest in this area of research.

We also observed that such studies have been carried out in several countries, especially in the European, Asian and American continents. As for the data sources, almost all the works were developed from their own databases. The data were collected locally, and with little dissemination to the scientific community. Another aspect is that the vast majority of the returned studies is in the experimental phase. However, three of them propose the design and the tool validation.

The returned works also presented a great diversity in the investigation of the effects of sound signals in the human organism. Most of these works sought to better understand the interventions of music and auditory stimuli in human brain waves (Qin et al., 2020; Dutta et al., 2020; Bhargava et al., 2020; Liao et al., 2017; Q. Li et al., 2020; Soysal et al., 2020; Ramdinwawii & Mittal, 2017; Rushambwa & Mythili, 2017; Chavan et al., 2016; O'Toole et al., 2019; Santhosh et al., 2020; Geethanjali et al., 2019; Liu et al., 2020; X. Wang et al., 2020; Jagiello et al., 2019; Ibrahim et al., 2019). Other works have studied ways in which factors such as musical preference can affect human physiology, mainly through the measurement of hemodynamic aspects (Yamada & Ono, 2019; Huang & Benjamin Knapp, 2017).

Several studies also acted in the investigation of behavioral changes favored by sound stimuli. In this sense, some works have successfully explored the effects of music in reducing anxiety, stress, and depression (Dragulin et al., 2019; Lin et al., 2020; Wang et al., 2016; Mitranont et al., 2017; Chavan et al., 2016; Chennafi et al., 2018; Liu et al., 2020; Geethanjali et al., 2019). Other groups studied the effects of these sound stimuli in increasing the state of relaxation, attention, and engagement in activities (Kanehira et al., 2018; Leslie et al., 2019; Ramdinwawii & Mittal, 2017; Stappen et al., 2019; Rushambwa & Mythili, 2017; Plut & Pasquier, 2019; Soysal et al., 2020; Rushambwa & Mythili, 2017).

Regarding the use of Artificial Intelligence (AI) tools, only nine works related to this topic made use of these tools (Yamada & Ono, 2019; Huang & Benjamin Knapp, 2017; Wang et al., 2016; Liu et al., 2020; Dutta et al., 2020; Bhargava et al., 2020; Q. Li et al., 2020; Ibrahim et al., 2019; Liu et al., 2020). Such computational approaches were mainly applied in works that sought to analyze cerebral and hemodynamic biosignals. Both supervised and unsupervised learning methods were used. Among these methods, it is worth highlight that Support Vector Machines (SVM) were applied in the majority of the studies. The ones which did not use AI tools, commonly used cognitive tests or performed purely qualitative analyzes through

observations and interviews with participants.

Table 2: Studies related to games, virtual reality, affective robotics and therapies.

Group	Description	Amount of Studies	Studies
G1	Games, Virtual Reality, Affective Robotics and Therapies	33	Krüger et al. (2020) Dechenaud et al. (2019) T. Xu et al. (2019) Sanders et al. (2020) Fonteles et al. (2018) Lubetzky et al. (2019) Kikuchi et al. (2018) Ortegon-Sarmiento et al. (2020) Lui et al. (2017) Shukla et al. (2019) Cameirão et al. (2017) Bermúdez i Badia et al. (2019) Tiwari & Tiwari (2017) Colombo et al. (2019) English & Howard (2017b) English & Howard (2017a) Kirana et al. (2018) HAN et al. (2018) Crespo et al. (2016) Aranha et al. (2017) Arroyo-Palacios & Slater (2016) Song et al. (2019) Murad et al. (2017) Mulla et al. (2017) Delmastro et al. (2018) Cavallo et al. (2020) Wolfe et al. (2020) Teo & Chia (2018) Lyu & Yuan (2020) Ranjkar et al. (2019) Sra et al. (2017) Scott (2017) Zhu et al. (2019)

Source: Authors.

Table 3: Studies related to physiological and behavioral responses induced by acoustic signals.

Group	Description	Amount of Studies	Studies
G2	Physiological and behavioral responses induced by acoustic signals	27	Yamada & Ono (2019) Dragulin et al. (2019) Kanehira et al. (2018) Stappen et al. (2019) Mitranont et al. (2017) Leslie et al. (2019) Liao et al. (2017) Lin et al. (2020) Ibrahim et al. (2019) Jagiello et al. (2019) X. Wang et al. (2020) Liu et al. (2020) Geethanjali et al. (2019) Santhosh et al. (2020) O'Toole et al. (2019) Chavan et al. (2016) Wang et al. (2016) Qin et al. (2020) Dutta et al. (2020) Bhargava et al. (2020) Q. Li et al. (2020) Plut & Pasquier (2019) Soysal et al. (2020) Ramdinwawii & Mittal (2017) Rushambwa & Mythili (2017) Chennafi et al. (2018) Huang & Benjamin Knapp (2017)

Source: Authors.

3.3 Emotion recognition associated with acoustic stimuli

Among the 144 studies selected through this systematic review, 35 are related to emotion recognition (listed in Table 4). Considering the years of publication, we noted that most of the works were published in the years 2019 and 2020, both presenting nine studies.

Considering the types of approach for the recognition of emotions, we observed that a significant number of studies uses the analysis of physiological signals to achieve this objective. Thus, studies using biosignals such as EEG (Bankar et al., 2018; Shen et al., 2020; Dutta et al., 2020; Marimpis et al., 2020; Bo et al., 2017; Rahman et al., 2020), ECG (Hsu et al., 2020) and skin's electrodermal activity (Rahman et al., 2019) were returned. Other studies combined more than one parameter, in addition to biosignals, to assess emotion. In this context, we highlight studies using more than one biosignal for emotional assessment. In Ramírez et al. (2020)'s approach, they combined functional magnetic resonance's and EEG's parameters. Daly et al. (2019), on the other hand, combined ECG signal, and the skin's electrodermal activity to the emotion recognition task. In works regarding the evaluation of biosignals, we also observe that in the vast majority the authors collect their own data. Although they described the processes in the papers, the lack of availability of open datasets makes it difficult to reproduce and improve the models which were tested.

Sereval works also used other parameters instead of biosignals for emotion recognition. Among them, we can highlight the groups that assess emotion through audio analysis (Greer et al., 2020; Lv et al., 2018; Mo & Niu, 2019; Lopes et al., 2019; Kumar et al., 2016; Panda et al., 2020; Chapaneri & Jayaswal, 2018). Textual aspects also have great potential in the field of emotion recognition in music. Thus, some works explore the use of song lyrics (Malheiro et al., 2018; Matsumoto & Sasayama, 2018), as well as their association with audiovisual content to improve the performance of models in this field of study (Nemati & Naghsh-Nilchi, 2017). As for data, most studies that involve only audio analysis, use data from public

databases such as Soundtracks, MTV database, MediaEval. In Panda et al. (2020)'s study, the authors made available to the scientific community the database which they elaborated. For those studies that use only video content, and audiovisual content, the authors assemble their own dataset. Therefore, it affects the study's reproducibility.

Finally, regarding the techniques involving AI, only nine works did not use this approach (Bankar et al., 2018; Greer et al., 2020; Daly et al., 2019; Kumar et al., 2016; Bortz et al., 2019). Among those, 15 papers tackled classification problems. On the other hand, 8 papers worked towards the prediction of continuous emotion using regression algorithms. Malheiro et al. (2018), in turn, involved both types of problem. Another aspect observed was that the SVM was the algorithm that was most present in the returned studies (Wardana et al., 2018; Hsu et al., 2020; Lv et al., 2018; Dutta et al., 2020; Mo & Niu, 2019; Bakhtiyari et al., 2019; Lopes et al., 2019; Bo et al., 2017; Nemati & Naghsh-Nilchi, 2017; Rahman et al., 2020; Nawa et al., 2018; Dantcheva et al., 2017), followed by Artificial Neural Networks (Zhang et al., 2019; Lv et al., 2018; Dutta et al., 2020; Goyal et al., 2016; Marimpis et al., 2020; Bakhtiyari et al., 2019; Rahman et al., 2019; Rahman et al., 2020; Chapaneri & Jayaswal, 2018).

3.4 Non-invasive assistive solutions based on human-machine musical interfaces

Musical and sound stimuli are increasingly being explored in interfaces of the most diverse orders. In this review, 12 of the 144 articles returned by the research deal with solutions based on human-machine musical interfaces (Table 5). Seven of them are in the testing phase and five in the implementation and clinical validation phase of the solution. All these studies used their own databases collected locally in the field or in the laboratory. Most of the works were developed in European and Asian countries, with special emphasis on Japan. This country showed a great advance in the development, and application of human-computer interfaces compared to other countries in the world.

Table 4: Studies related to emotion recognition associated with acoustic stimuli.

Group	Description	Amount of Studies	Studies
G3	Emotion recognition associated with acoustic stimuli	35	Bankar et al. (2018) Zhang et al. (2019) Malheiro et al. (2018) Navarathna et al. (2019) Shen et al. (2020) Dantcheva et al. (2017) Wardana et al. (2018) Greer et al. (2020) Hsu et al. (2020) Lv et al. (2018) Matsumoto & Sasayama (2018) Daly et al. (2019) Dutta et al. (2020) S. Ehrlich et al. (2017) Goyal et al. (2016) Marimpis et al. (2020) Mo & Niu (2019) Saha et al. (2016) Bakhtiyari et al. (2019) Lopes et al. (2019) Bo et al. (2017) Nemati &Naghsh-Nilchi (2017) Kumar et al. (2016) Panda et al. (2020) Rahman et al. (2019) Rahman et al. (2020) Bo et al. (2017) Chapaneri & Jayaswal (2018) Bortz et al. (2019) X. Xu et al. (2019) Tarvainen et al. (2017) Chin et al. (2018) Ramírez et al. (2020) Nawa et al. (2018) Marimpis et al. (2020)

Source: Authors.

Table 5: Studies related to non-invasive assistive solutions based on human-machine musical interfaces.

Group	Description	Amount of Studies	Studies
G4	Non-invasive assistive solutions based on human-machine musical interfaces	12	LingHu & Shu (2018) Daly et al. (2016) S. K. Ehrlich et al. (2019) Yi-Hsiang et al. (2018) Kobayashi & Fujishiro (2016) Desai et al. (2018) Boumpa et al. (2018) Daly et al. (2020) Shan et al. (2018) Syed et al. (2018) Mideska et al. (2016) Hasan et al. (2020)

Source: Authors.

Regarding the use of AI, six works did not use or did not make explicit the use of these techniques. Among the studies involving AI, we noticed a preference for supervised learning methods, especially the SVMs. In addition, there is also a predominance in the use of technologies using ARM-type microcontrollers, and sensors with remote communication.

Most of the works of this group propose the development of interfaces for monitoring and modulating mental and

affective states through music (LingHu & Shu, 2018; Mideska et al., 2016; Shan et al., 2018; Daly et al., 2016; S. K. Ehrlich et al., 2019; Kobayashi & Fujishiro, 2016; Daly et al., 2020). In Desai et al. (2018)'s work they present an approach in the opposite direction.

In their study, they propose a musical brain machine interface (BCMI) which modulates musical parameters - such as harmony, rhythm, and melody - from the biofeedback captured by the user's EEG signal. The main goal is to adapt the music according to their emotional state. Thus, instead of using music to modulate the emotional state, the authors propose to alter the music according to the user's brain activity.

Some other studies searches for technologies to improve the development of activities of daily living for people with disabilities. An interface based on sound signals was proposed by Boumpa et al. (2018) to help people with dementia in the development of their daily activities (Boumpa et al., 2018). Hasan et al. (2020)'s study uses the emotional response to musical stimuli to identify movement intention (Hasan et al., 2020). Such a solution can be incorporated in assistive BCIs. Other works seek to apply these musical interfaces to improve communication and interaction between people (Syed et al., 2018; Yi-Hsiang et al., 2018). Syed et al. (2018)'s study proposes a tool to help people with hearing impairments and hearing people communicate. In Yi-Hsiang et al. (2018)'s work, in turn, musical synchronization is studied to encourage interaction between individuals with Autism Spectrum Disorder and robots.

In general, there is a growth in the interest of researchers in human-machine interfaces, including those that make use of sound and musical aspects to achieve their goals. Despite this, there is a concentration of studies of this nature in regions with greater economic development, as they generally involve the use of high-tech and high-cost devices for capturing and processing signals. It was also observed that the use of AI in this context is still shy, possibly due to the difficulty associated with the integration of this technology in hardware with remote communication systems and real-time responses.

3.5 Musical composition, recommendation and customization

This group has 18 of the 144 studies returned by the search string. They deal with approaches to composition, personalization or music recommendation. These works are presented in Table 6. The vast majority of these works were elaborated by groups located in Eastern countries, mainly in China. Works in this area were most published in 2016 and 2020, with no work made available in 2019.

Table 6: Studies related to musical composition, recommendation and customization.

Group	Description	Amount of Studies	Studies
G5	Musical composition, recommendation and customization	18	Savery et al. (2019) Deshmukh et al. (2017) Subramaniam et al. (2018) Y.-C. Wu & Chen (2016) Chang et al. (2017) Vinayagasundaram et al. (2016) Su et al. (2020) Chen et al. (2020) González & McMullen (2020) Chen et al. (2020) Tiple et al. (2016) Seanglidet et al. (2016) Hossan & Chowdhury (2016) Gu et al. (2018) Sawata et al. (2019) Amali et al. (2018) Gilda et al. (2017) Fernandes et al. (2021)

Source: Authors.

Considering the origin of the data used in the studies, the majority of the groups used their own database. Three of them used the Facial Expression Recognition 2013 (FER-2013) Dataset, to identify emotions through facial expression. Furthermore, two of them used the DEAP Dataset, which is a dataset for emotion analysis using EEG, physiological and video signals.

From these works, we observed that the composition, personalization and musical recommendation are generally carried out using multimodal signals. Most works combine physiological signals such as galvanic skin response, heart and respiratory rates, and facial expressions. Some of these works also use electroencephalographic signals or functional images to analyze brain activity. As they are multimodal signals and often difficult to interpret, most studies use AI methods to analyze the data. In general, deep learning techniques are more frequently used, especially LSTMs and CNNs configurations. Some works also use shallow learning approaches, such as SVM and K-means. Works that do not use AI techniques commonly use computational methods of signal filtering and threshold identification to process the data.

Some of the works propose systems and approaches still in the state of computational testing phase. Others are in the validation and implementation phase with real users. From these studies, there are two works that present proposals for systems for automatic emotion detection (González & McMullen, 2020; Vinayagasundaram et al., 2016). Three other studies use techniques for generating affective or personalized music (Savery et al., 2019; Y.-C. Wu & Chen, 2016; Chen et al., 2020). These studies of musical generation combine several aspects related to music composition and affectivity, and musical taste factors. Generally, these studies focus on the therapeutic context, to favor user engagement.

The vast majority of works in this group deal with approaches for recommending music or audiovisual content in a personalized way (Deshmukh et al., 2017; Subramaniam et al., 2018; Chang et al., 2017; Chen et al., 2020; Tiple et al., 2016; Seanglidet et al., 2016; Gu et al., 2018; Amali et al., 2018; Gilda et al., 2017; Su et al., 2020). In these works, several aspects are used to identify users' opinions. Most of these studies analyze facial expressions (Chang et al., 2017), EEG (Chang et al., 2017) and text recognition (Subramaniam et al., 2018).

Furthermore, Hossan & Chowdhury (2016) propose an approach to emotional state control based on affective music (Hossan & Chowdhury, 2016). Sawata et al. (2019), on the other hand, presents a strategy for identifying favorite music (Sawata et al., 2019). Fernandes et al. (2021)'s study proposes the conversion of EEG signals into music as a way to promote musical personalization (Fernandes et al., 2021).

3.6 Other approaches

The approach used in the works presented here, in most cases, determines the conduct of the research. In this section, however, 19 articles could not be grouped since they do not have a common field among them (Table 7). It can be said that these studies are somewhat unique, from their own perspective or application. Among them, 12 works were published in the years 2019 and 2020. This number represents 63% of the total articles found since 2016. This increase may also indicate a growing trend towards new approaches involving emotion recognition (Zhang et al., 2018; Mehta et al., 2019; Al-Qazzaz et al., 2020; Aydın, 2020; Rizos & Schuller, 2019; Thammasan et al., 2017), and the use of music combined with affective computing.

Geographically, 58% of the works were published in the Asian continent. Most of them were placed in China and Japan, with two publications each. As for the databases used, almost 42% were created exclusively for the study itself. It may indicate a scarcity of material for approaches that generally require very detailed and sufficiently diverse databases. In this context, studies that focus on the creation of databases under stimuli (Suhaimi et al., 2018), and even an alternative perspective are crucial. Regarding the study phases, eight of the 19 articles are classified as experimental phase. However there is also a well-balanced distribution, and often mixed with modeling, testing and concept design.

An important observation is that in seven studies the authors did not use any Artificial Intelligence (AI) techniques. Apparently, this fact may have some coherence because it deals with other approaches and alternatives. However, very similarly to the works already presented in this section 4, these publications also involve brain signal capture Barnstaple et al. (2020), database creation (Suhaimi et al., 2018), music generation, and physiological (Lui e Grunberg, 2019; Nalepa et al., 2019) or affective response to acoustic stimuli. In the studies which present AI methods, the algorithm widely used is the SVM. The SVM is present in six publication. Nonetheless, along with classical approaches, other algorithms were investigated -in most cases- such as LSTM, KNN, and Deep Neural Network.

It is interesting to bring attention that some of these studies - considering the emotion prediction perspective - investigate the similarities between the images' emotional content and music (Xing et al., 2019; Verma et al., 2019; Parra et al., 2019). In these articles, the identification of visual and sound attributes, and their relationship with emotional stimulation help to understand two striking points. The first one is the emotion recognition itself, whereas the second one is the characteristics present in these media. These characteristics can be used in both acoustic (Saha et al., 2016) and visual (Suhaimi et al., 2018) stimulation to obtain an affective response.

In contrast, other works use well-known attributes to create affective sentences from images (Konno et al., 2018) and generate synthetic music from an original database (Herremans e Chew, 2019). Although the articles present promising approaches, it is important to draw attention to some aspects. For example, the need of higher datasets, better description of the evaluation metrics, and meticulous detailed methodology. A significant improvement on these aspects may possibly bring more precise answers and substantially enhance the proposed models.

Table 7: Studies related to other approaches.

Group	Description	Amount of Studies	Studies
G6	Other approaches	19	Konno et al. (2018) Zhang et al. (2018) Mehta et al. (2019) X. Li et al. (2020) Al-Qazzaz et al. (2020) Barnstaple et al. (2020) Aydın (2020) Nalepa et al. (2019) Xing et al. (2019) Verma et al. (2019) Navarro-Tuch et al. (2018) Lui & Grunberg (2017) Saha et al. (2016) Suhaimi et al. (2018) Rizos & Schuller (2019) Herremans & Chew (2019) Thammasan et al. (2017) Kyong et al. (2019) Parra et al. (2019)

Source: Authors.

4. Conclusion

This review aimed to identify AI methods to perform automatic emotion recognition applied to human-machine musical interfaces. We were interested in mapping studies related to emotions perceived through facial expressions, speech, EEG or physiological signals. Innovative therapeutic approaches were also of interest, especially those that use musical stimuli and computational methods.

The review protocol was systematically defined and followed throughout the process. The search string resulted in

290 publications found in the five main scientific databases on the fields of intelligent computing, engineering, and medicine. From these publications, 144 remained in this review after the qualitative and content analysis.

These studies were grouped into six groups based on their content. Most of them had a high interdisciplinary content. Associations of therapy issues with human-computer interfaces, music, and robotics were common. The studies in this review showed that the areas of affective computing and music therapy are being more combined in the state-of-the-art. While interdisciplinarity proved to be a feature of this field of study, it was also a challenge during the grouping process. Thus, making it often hard to categorize a study in only one of the groups.

The review also demonstrated that most databases used in the included studies are not public. This makes it difficult for other groups to access the data collected and used in many of the works. The lack of dissemination of the databases compromises the reproducibility of the studies. Moreover, it serves as an obstacle to the accessibility of the proposed solutions. Therefore, we encourage the dissemination of more databases in the area. This dissemination can not only provide greater reach of their authors works but can also benefit researchers from places with unsatisfying structure to carry out data acquisition.

This study also found that the use of AI in the context of affective computing and music therapy is growing with the years. The papers that did not use AI techniques commonly performed interviews, forms, and cognitive-behavioral tests to extract and assess information. The use of Support Vector Machines and decision trees architectures is still vast in the literature. However, in many cases, deep learning methods are also incorporated, mainly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). In the case of CNNs, the authors have been investing in networks with more and more layers, as data has become increasingly complex and multimodal. As for RNNs, the most explored in the studies has been the Long ShortTerm Memory (LSTM), which may provide information in the temporal dimension. Temporality is an essential factor in speech processing and musical composition. Overall, we find that studies using AI and other computational techniques often do not provide a well-described methodology and/or parameters. This compromises the quality and reproducibility of these studies.

Many of the included studies that used AI applied it to music processing and emotion recognition. In music processing, AI was used to recognize patterns and use them to recommend personalized content or even compose music. Regarding emotion recognition, most works use deep networks and multimodality data. Most of these studies succeed in recognizing primary emotions (i.e. anger, fear, surprise, disgust, contempt, happiness, and sadness) but are not efficient in identifying secondary emotions such as love, frustration, shame, and relief. Another challenge for emotion recognition is how to differentiate felt from expressed emotions. Some studies seek to overcome this challenge by assessing multimodality data (e.g. combine EEG, facial expression and physiological signals).

The importance of multimodality data was recurrent in the works. This also points to the relevance of these studies being carried out by multidisciplinary teams. The combination of professionals from different areas is beneficial for the development of robust and useful solutions.

Regarding the development of Human-Machine Musical Interfaces (HMMIs), the importance of using music in an appropriate, thought-out, programmed, and personalized way to achieve non-musical goals was perceived. Over the years, there has been an increase in the demand for innovations, including therapeutic ones. In this review, many works already propose HMMIs to provide personalized feedback based on the users' biomedical data. Unfortunately, there are still relatively few studies in this field, which are mostly concentrated in more developed countries.

From this systematic literature review, we hope to provide theoretical foundations to encourage the development of research in affective computing combined to music therapy. This is a very promising area where there is still a lot to be explored. Artificial Intelligence tools applied to emotion recognition can optimize music therapy processes and, thus, enhance

its effects. Finally, we believe that the popularization of therapeutic approaches such as music therapy has great potential to improve the quality of life of people with cognitive, motor and behavioral disorders.

Acknowledgments

The authors thank the partial financial support from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) Finance Code 001; the Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE), Brazil, under the code IBPG-0013-1.03/20; and the Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq-DT2-2018.

References

- Agres, K. R., Schaefer, R. S., Volk, A., Hooren, S. v., Holzapfel, A., Bella, S. D., ... Magee, W. L. (2021). Music, Computing, and Health: A Roadmap for the Current and Future Roles of Music Technology for Health Care and Well-Being. *Music & Science*, 4, 1–32. doi: 10.1177/2059204321997709
- Al-Qazzaz, N. K., Sabir, M. K., Ali, S. H. B. M., Ahmad, S. A., & Grammer, K. (2020). Electroencephalogram Profiles for Emotion Identification over the Brain Regions Using Spectral, Entropy and Temporal Biomarkers. *Sensors(Basel)*, 20. doi:10.3390/s20010059
- Amali, D. N., Barakbah, A. R., Anom Besari, A. R., & Agata, D. (2018). Semantic Video Recommendation System Based on Video Viewers Impression from Emotion Detection. In *2018 international electronics symposium on knowledge creation and intelligent computing (ies-kcic)* (p. 176-183). doi: 10.1109/KCIC.2018.8628592
- Aranha, R. V., Silva, L. S., Chaim, M. L., & Nunes, F. d. L. d. S. (2017). Using Affective Computing to Automatically Adapt Serious Games for Rehabilitation. In *2017 IEEE 30th international symposium on computer-based medical systems (cbms)* (p. 55-60). doi: 10.1109/CBMS.2017.89
- Arroyo-Palacios, J., & Slater, M. (2016). Dancing with Physio: A Mobile Game with Physiologically Aware Virtual Humans. *IEEE Transactions on Affective Computing*, 7(4), 326-336. doi: 10.1109/TAFFC.2015.2472013
- Aydın, S. (2020). Deep Learning Classification of Neuro-Emotional Phase Domain Complexity Levels Induced by Affective Video Film Clips. *IEEE Journal of Biomedical and Health Informatics*, 24(6), 1695-1702. doi: 10.1109/JBHI.2019.2959843
- Bakhtiyari, K., Taghavi, M., Taghavi, M., & Bentahar, J. (2019). Ambiance Signal Processing: A Study on Collaborative Affective Computing. In *2019 5th international conference on web research (icwr)* (p. 35-40). doi: 10.1109/ICWR.2019.8765251
- Bankar, C., Bhide, A., Kulkarni, A., Ghube, C., & Bedekar, M. (2018). Driving Control Using Emotion Analysis Via EEG. In *2018 IEEE Punecon* (p. 1-7). doi: 10.1109/PUNECON.2018.8745412
- Barnstaple, R., Protzak, J., DeSouza, J. F., & Gramann, K. (2020). Mobile brain/body imaging in dance: A dynamic transdisciplinary field for applied research. *European Journal of Neuroscience*. doi: <https://doi.org/10.1111/ejn.14866>
- Bermúdez i Badia, S., Quintero, L. V., Cameirão, M. S., Chirico, A., Triberti, S., Cipresso, P., & Gaggioli, A. (2019). Toward emotionally adaptive virtual reality for mental health applications. *IEEE Journal of Biomedical and Health Informatics*, 23(5), 1877-1887. doi: 10.1109/JBHI.2018.2878846
- Bertazone, T. M. A., Ducatti, M., Camargo, H. P. M., Batista, J. M. F., Kusumota, L., & Marques, S. (2016). Ações multidisciplinares/interdisciplinares no cuidado ao idoso com Doença de Alzheimer. *Rev Rene*.
- Bhargava, A., O'Shaughnessy, K., & Mann, S. (2020). A novel approach to eeg neurofeedback via reinforcement learning. In *2020 IEEE Sensors* (p. 1-4). doi: 10.1109/SENSORS47125.2020.9278871
- Bo, H., Ma, L., & Li, H. (2017). Music-evoked emotion classification using EEG correlation-based information. In *2017 39th annual international conference of the IEEE engineering in medicine and biology society (EMBC)* (p. 3348-3351). doi: 10.1109/EMBC.2017.8037573
- Bortz, B., Jaimovich, J., & Knapp, R. B. (2019). Cross-Cultural Comparisons of Affect and Electrodermal Measures While Listening to Music. In *2019 8th international conference on affective computing and intelligent interaction (ACII)* (p. 55-61). doi: 10.1109/ACII.2019.8925476
- Boumpa, E., Charalampou, I., Gkogkidis, A., Ntaliani, A., Kokkinou, E., & Kakarountas, A. (2018). Assistive System for Elders Suffering of Dementia. In *2018 IEEE 8th international conference on consumer electronics - berlin (icce-berlin)* (p. 1-4). doi: 10.1109/ICCE-Berlin.2018.8576216
- Brosch, T., Scherer, K. R., Grandjean, D. M., & Sander, D. (2013). The impact of emotion on perception, attention, memory, and decision-making. *Swiss Medical Weekly*, 143. doi: 10.4414/sm.w.2013.13786
- Bulagang, A. F., Mountstephens, J., & Teo, J. (2021). Multiclass emotion prediction using heart rate and virtual reality stimuli. *Journal of Big Data*, 12. doi: 10.1186/s40537-020-00401-x
- Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *1st Conference on Fairness, Accountability and Transparency*, 77-91.
- Caetano, L. A. O., Silva, F. S., & Silveira, C. A. B. (2017). Alzheimer, sintomas e Grupos: Uma Revisão Integrativa. *Revista do NESME*.

- Cambria, E. (2016). Affective computing and sentiment analysis. *IEEE Intelligent Systems*, 31(2), 102–107.
- Cameirão, M. S., Pereira, F., & i. Badia, S. B. (2017). Virtual reality with customized positive stimuli in a cognitive-motor rehabilitation task. In *2017 international conference on virtual rehabilitation (icvr)* (p. 1-7). doi: 10.1109/ICVR.2017.8007543
- Cavallo, F., Rovini, E., Dolciotti, C., Radi, L., Ragione, R. D., Bongioanni, P., & Fiorini, L. (2020). Physiological response to Vibro-Acoustic stimulation in healthy subjects: a preliminary study*. In *2020 42nd annual international conference of the ieee engineering in medicine biology society (embc)* (p. 5921-5924). doi: 10.1109/EMBC44109.2020.9175848
- Chang, H.-Y., Huang, S.-C., & Wu, J.-H. (2017). A personalized music recommendation system based on electroencephalography feedback. *Multimedia Tools and Applications*, 76, 19523—19542. doi: 10.1007/s11042-015-3202-4
- Chapaneri, S., & Jayaswal, D. (2018). Deep Gaussian Processes for Estimating Music Mood. In *2018 15th ieee india council international conference (indicon)* (p. 1-5). doi: 10.1109/INDICON45594.2018.8987036
- Chatziagapi, A., Paraskevopoulos, G., Sgouropoulos, D., Pantazopoulos, G., Nikandrou, M., Giannakopoulos, T., ... Narayanan, S. (2019). Data Augmentation using GANs for Speech Emotion Recognition. *INTERSPEECH 2019*.
- Chavan, D. R., Kumbhar, M. S., & Chavan, R. R. (2016). The human stress recognition of brain, using music therapy. In *2016 international conference on computation of power, energy information and commuincation (iccepic)* (p. 200-203). doi: 10.1109/ICCPEIC.2016.7557197
- Chen, J., Pan, F., Zhong, P., He, T., Qi, L., Lu, J., ... Zheng, Y. (2020). An Automatic Method to Develop Music With Music Segment and Long Short Term Memory for Tinnitus Music Therapy. *IEEE Access*, 8, 141860-141871. doi: 10.1109/ACCESS.2020.3013339
- Chen, O. T., Chang, S., Ma, Y., Zhang, Y. C., & Lee, Y. L. (2020). Time capsule gift with affective awareness of event memories via near field communication. In *2020 ieee/sice international symposium on system integration (sii)* (p. 585-589). doi: 10.1109/SII46433.2020.9026183
- Chennafi, M., Khan, M. A., Li, G., Lian, Y., & Wang, G. (2018). Study of music effect on mental stress relief based on heart rate variability. In *2018 ieee asia pacific conference on circuits and systems (apccas)* (p. 131-134). doi: 10.1109/APCCAS.2018.8605674
- Chin, Y.-H., Wang, J.-C., Wang, J.-C., & Yang, Y.-H. (2018). Predicting the Probability Density Function of Music Emotion Using Emotion Space Mapping. *IEEE Transactions on Affective Computing*, 9(4), 541-549. doi: 10.1109/TAFFC.2016.2628794
- Colombo, R., Raglio, A., Panigazzi, M., Mazzone, A., Bazzini, G., Imarisio, C., ... Imbriani, M. (2019). The sonichand protocol for rehabilitation of hand motor function: A validation and feasibility study. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(4), 664-672. doi: 10.1109/TNSRE.2019.2905076
- Crespo, A. B., Idrovo, G. G., Rodrigues, N., & Pereira, A. (2016). A virtual reality UAV simulation with body area networks to promote the elders life quality. In *2016 1st international conference on technology and innovation in sports, health and wellbeing (tishw)* (p. 1-7). doi: 10.1109/TISHW.2016.7847780
- Daly, I., Williams, D., Kirke, A., Miranda, E. R., & Nasuto, S. J. (2019). Electroencephalography reflects the activity of subcortical brain regions during approach-withdrawal behaviour while listening to music. *Scientific Reports*, 9. doi: 10.1038/s41598-019-45105-2
- Daly, I., Williams, D., Kirke, A., Weaver, J., Malik, A., Hwang, F., ... Nasuto, S. J. (2016). Affective brain-computer music interfacing. *Journal of Neural Engineering*, 13. doi: 10.1088/1741-2560/13/4/046022
- Daly, I., Williams, D., Malik, A., Weaver, J., Kirke, A., Hwang, F., ... Nasuto, S. J. (2020). Personalised, Multi-Modal, Affective State Detection for Hybrid Brain-Computer Music Interfacing. *IEEE Transactions on Affective Computing*, 11(1), 111-124. doi: 10.1109/TAFFC.2018.2801811
- Dantcheva, A., Bilinski, P., Nguyen, H. T., Broutart, J.-C., & Bremond, F. (2017). Expression recognition for severely demented patients in music reminiscence-therapy. In *2017 25th european signal processing conference (eusipco)* (p. 783-787). doi: 10.23919/EUSIPCO.2017.8081314
- Dechenaud, M., Laidig, D., Seel, T., Gilbert, H. B., & Kuznetsov, N. A. (2019). Development of Adapted Guitar to Improve Motor Function After Stroke: Feasibility Study in Young Adults. In *Annual international conference of the iee engineering in medicine and biology society* (pp. 5488–5493). doi: 10.1109/EMBC.2019.8856651
- Delmastro, F., Martino, F. D., & Dolciotti, C. (2018). Physiological Impact of Vibro-Acoustic Therapy on Stress and Emotions through Wearable Sensors. In *2018 ieee international conference on pervasive computing and communications workshops (percom workshops)* (p. 621-626). doi: 10.1109/PERCOMW.2018.8480170
- Desai, B., Chen, B., Sirocchi, S., & McMullen, K. A. (2018). Mindtrack: Using brain-computer interface to translate emotions into music. In *2018 international conference on digital arts, media and technology (icdamt)* (p. 33-37). doi: 10.1109/ICDAMT.2018.8376491
- Deshmukh, R. S., Jagtap, V., & Paygude, S. (2017). Facial emotion recognition system through machine learning approach. In *2017 international conference on intelligent computing and control systems (iciccs)* (p. 272-277). doi: 10.1109/ICCONS.2017.8250725
- Souza, M. C., da Rocha Alves, A. B., de Lima, D. S., de Oliveira, L. R. F. A., da Silva, J. K. B., de Oliveira Ribeiro, E. C., ... de Oliveira, e. a., I. D. (2017). The treatment of Alzheimer in the context of musicotherapy. *International Archives of Medicine*, 10.
- Dorneles, S. O., Barbosa, D. N. F., & Barbosa, J. L. V. (2020). Sensibilidade ao contexto na identificação de estados afetivos aplicados à educação: um mapeamento sistemático. *Revista Novas Tecnologias na Educação - RNOTE*, 18.
- Dragulin, S., Constantin, F. A., & Rucsanda, I. (2019). The Use of Music Therapy for Adults' Anxiety Relief. In *2019 5th experiment international conference (exp.at'19)* (p. 490-493). doi: 10.1109/EXPAT.2019.8876558

- Dutta, E., Bothra, A., Chaspari, T., Ioerger, T., & Mortazavi, B. J. (2020). Reinforcement learning using eeg signals for therapeutic use of music in emotion management. In *2020 42nd annual international conference of the ieee engineering in medicine biology society (embc)* (p. 5553-5556). doi: 10.1109/EMBC44109.2020.9175586
- EC. (2020). The 2021 Ageing Report: Underlying Assumptions and Projection Methodologies [Computer software manual]. Retrieved from https://ec.europa.eu/info/sites/default/files/economy-finance/ip142_en.pdf (Last accessed: 2021 Jul. 05)
- Ehrlich, S., Guan, C., & Cheng, G. (2017). A closed-loop brain-computer music interface for continuous affective interaction. In *2017 international conference on orange technologies (icot)* (p. 176-179). doi: 10.1109/ICOT.2017.8336116
- Ehrlich, S. K., Agres, K. R., Guan, C., & Cheng, G. (2019, 03). A closed-loop, music-based brain-computer interface for emotion mediation. *PLOS ONE*, *14*(3), 1-24. Retrieved from <https://doi.org/10.1371/journal.pone.0213516> doi:10.1371/journal.pone.0213516
- Ekman, P., & Friesen, V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 124–129.
- English, B. A., & Howard, A. (2017a). The effects of adjusting task difficulty on learning motor and cognitive aspects of a multitasking task. In *2017 ieee symposium series on computational intelligence (ssci)* (p. 1-7). doi: 10.1109/SSCI.2017.8285396
- English, B. A., & Howard, A. M. (2017b). The effects of musical cues on motor learning using a robotic wrist rehabilitation system — a healthy pilot study. In *2017 ieee workshop on advanced robotics and its social impacts (arso)* (p. 1-6). doi: 10.1109/ARSO.2017.8025208
- Fang, R., Ye, S., Huangfu, J., & Calimag, D. P. (2017). Music therapy is a potential intervention for cognition of Alzheimer’s disease: a mini-review. *Translational Neurodegeneration*, *6*(1), 2.
- Fernandes, C. M., Migotina, D., & Rosa, A. C. (2021). Brain’s Night Symphony (BrainSy): A Methodology for EEG Sonification. *IEEE Transactions on Affective Computing*, *12*(1), 103-112. doi: 10.1109/TAFFC.2018.2850008
- Fonteles, J. H., Serpa, Y. R., Barbosa, R. G., Rodrigues, M. A. F., & Alves, M. S. P. L. (2018). Gesture-controlled interactive musical game to practice hand therapy exercises and learn rhythm and melodic structures. In *2018 ieee 6th international conference on serious games and applications for health (segah)* (p. 1-8). doi: 10.1109/SeGAH.2018.8401367
- Gallego, M. G., & Garcia, J. G. (2017). Music therapy and Alzheimer’s disease: Cognitive, psychological, and behavioural effects. *Neurología (English Edition)*, *32*(5), 300–308. doi: 10.1155/2014/908915
- Geethanjali, B., Adalarasu, K., Jagannath, M., & Guhan Seshadri, N. P. (2019). Music-Induced Brain Functional Connectivity Using EEG Sensors: A Study on Indian Music. *IEEE Sensors Journal*, *19*(4), 1499-1507. doi: 10.1109/JSEN.2018.2873402
- Geraets, C. N., Stouwe, E. C. d., Pot-Kolder, R., & Veling, W. (2021). Advances in immersive virtual reality interventions for mental disorders: A new reality? *Current Opinion in Psychology*, *41*, 40–45. doi: 10.1016/j.copsyc.2021.02.004
- Gilda, S., Zafar, H., Soni, C., & Waghurdekar, K. (2017). Smart music player integrating facial emotion recognition and music mood recommendation. In *2017 international conference on wireless communications, signal processing and networking (wisnet)* (p. 154-158). doi: 10.1109/WiSPNET.2017.8299738
- González, E. J. S., & McMullen, K. (2020). The design of an algorithmic modal music platform for eliciting and detecting emotion. In *2020 8th international winter conference on brain-computer interface (bci)* (p. 1-3). doi: 10.1109/BCI48061.2020.9061664
- Goudbeek, M., Goldman, J. P., & Scherer, K. R. (2009). Emotion dimensions and formant position. *INTERSPEECH 2009 Annual Conference of the International Speech Communication Association*.
- Goyal, A., Kumar, N., Guha, T., & Narayanan, S. S. (2016). A multimodal mixture-of-experts model for dynamic emotion prediction in movies. In *2016 ieee international conference on acoustics, speech and signal processing (icassp)* (p. 28222826). doi: 10.1109/ICASSP.2016.7472192
- Greer, T., Mundnich, K., Sachs, M., & Narayanan, S. (2020). The Role of Annotation Fusion Methods in the Study of Human Reported Emotion Experience During Music Listening. In *icassp 2020 - 2020 ieee international conference on acoustics, speech and signal processing (icassp)* (p. 776-780). doi: 10.1109/ICASSP40776.2020.9054329
- Gu, F., Niu, J., Das, S. K., & He, Z. (2018). RunnerPal: A Runner Monitoring and Advisory System Based on Smart Devices. *IEEE Transactions on Services Computing*, *11*(2), 262-276. doi: 10.1109/TSC.2016.2626372
- Hamdan, A. C. (2008). Avaliação neuropsicológica na doença de alzheimer e no comprometimento cognitivo leve. *Psicol. argum*, 183–192.
- Han, W., & Chan, C. (2006). An efficient MFCC extraction method in speech recognition. *International Symposium on Circuits and Systems*.
- Han, Y., Nishio, Y., Yi-Hsiang, M., Oshiyama, C., Lin, J.-Y., Takanishi, A., & Cosentino, S. (2018). A humanrobot interface to improve facial expression recognition in subjects with autism spectrum disorder. In *2018 9th international conference on awareness science and technology (icast)* (p. 179-184). doi: 10.1109/ICAwST.2018.8517228
- Hasan, S. M. S., Siddiquee, M. R., Marquez, J. S., & Bai, O. (2020). Enhancement of Movement Intention Detection Using EEG Signals Responsive to Emotional Music Stimulus. *IEEE Transactions on Affective Computing*, 1-1. doi: 10.1109/TAFFC.2020.3025004
- Herremans, D., & Chew, E. (2019). MorpheuS: Generating Structured Music with Constrained Patterns and Tension. *IEEE Transactions on Affective Computing*, *10*(4), 510-523. doi: 10.1109/TAFFC.2017.2737984
- Hossan, A., & Chowdhury, A. M. M. (2016). Real time EEG based automatic brainwave regulation by music. In *2016 5th international conference on informatics, electronics and vision (iciev)* (p. 780-784). doi: 10.1109/ICIEV.2016.7760107

- Hsu, Y.-L., Wang, J.-S., Chiang, W.-C., & Hung, C.-H. (2020). Automatic ECG-Based Emotion Recognition in Music Listening. *IEEE Transactions on Affective Computing, 11*(1), 85-99. doi: 10.1109/TAFFC.2017.2781732
- Huang, W., & Benjamin Knapp, R. (2017). An exploratory study of population differences based on massive database of physiological responses to music. In *2017 seventh international conference on affective computing and intelligent interaction (acii)* (p. 524-530). doi: 10.1109/ACII.2017.8273649
- Ibrahim, I. A., Ting, H.-N., & Moghavvemi, M. (2019). Formulation of a Novel Classification Indices for Classification of Human Hearing Abilities According to Cortical Auditory Event Potential signals. *Arabian Journal for Science and Engineering, 44*, 7133-7147. doi: 10.1007/s13369-019-03835-5
- Ingale, A. B., & Chaudhari, D. S. (2012). Speech Emotion Recognition. *International Journal of Soft Computing and Engineering, 2*.
- Izard, C. E. (1977). *Human Emotions*. New York: Springer.
- Jagiello, R., Pomper, U., Yoneya, M., Zhao, S., & Chait, M. (2019). Rapid Brain Responses to Familiar vs. Unfamiliar Music – an EEG and Pupillometry study. *Scientific Reports, 9*. doi: 10.1038/s41598-019-51759-9
- Jeong, M., & Ko, B. C. (2018). Driver's Facial Expression Recognition in Real-Time for Safe Driving. *Sensors*. doi: 10.3390/s18124270
- Kanehira, R., Ito, Y., Suzuki, M., & Hideo, F. (2018). Enhanced relaxation effect of music therapy with VR. In *2018 14th international conference on natural computation, fuzzy systems and knowledge discovery (icnc-fskd)* (p. 1374-1378). doi: 10.1109/FSKD.2018.8686951
- Kikuchi, T., Nagata, T., Sato, C., Abe, I., Inoue, A., Kugimiya, S., ... Hatabe, S. (2018). Sensibility Assessment For User Interface and Training Program an Upper-Limb Rehabilitation Robot, D-SEMUL. In *Annual international conference of the IEEE engineering in medicine and biology society* (pp. 3028-3031). doi: 10.1109/EMBC.2018.8513074
- King, J., Jones, K., Goldberg, E., Rollins, M., MacNamee, K., Moffit, C., ... Amaro, e. a., J. (2019). Increased functional connectivity after listening to favored music in adults with Alzheimer dementia. *The Journal of Prevention of Alzheimer's Disease, 6*(1), 56-62.
- Kirana, M. C., Lubis, M. Z., & Amalia, E. D. (2018). The effect of sound manipulation to know response rate in autism children using fft. In *2018 international conference on applied engineering (icae)* (p. 1-5). doi: 10.1109/INCAE.2018.8579418
- Kobayashi, A., & Fujishiro, I. (2016). An Affective Video Generation System Supporting Impromptu Musical Performance. In *2016 international conference on cyberworlds (cw)* (p. 17-24). doi: 10.1109/CW.2016.11
- Konno, M., Suzuki, K., & Sakamoto, M. (2018). Sentence Generation System Using Affective Image. In *2018 joint 10th international conference on soft computing and intelligent systems (scis) and 19th international symposium on advanced intelligent systems (isis)* (p. 678-682). doi: 10.1109/SCIS-ISIS.2018.00114
- Krüger, C., Kojic, T., Meier, L., Möller, S., & Voigt-Antons, J.-N. (2020). Development and validation of pictographic scales for rapid assessment of affective states in virtual reality. In *2020 twelfth international conference on quality of multimedia experience (qomex)* (p. 1-6). doi: 10.1109/QoMEX48832.2020.9123100
- Kumar, N., Guha, T., Huang, C., Vaz, C., & Narayanan, S. S. (2016). Novel affective features for multiscale prediction of emotion in music. In *2016 IEEE 18th international workshop on multimedia signal processing (mmSP)* (p. 1-5). doi: 10.1109/MMSP.2016.7813377
- Kyong, J.-S., Noh, T.-S., Park, M., Oh, S.-H., Lee, J., & Suh, M.-W. (2019, 06). Phantom Perception of Sound and the Abnormal Cortical Inhibition System: An Electroencephalography (EEG) Study. *Annals of Otolaryngology, Rhinology & Laryngology, 128*, 84S-95S. doi: 10.1177/0003489419837990
- Le, D., & Provost, E. M. (2013). Emotion recognition from spontaneous speech using hidden markov models with deep belief networks. *IEEE Workshop on Automatic Speech Recognition and Understanding, 216-221*.
- Leslie, G., Ghandeharioum, A., Zhou, D., & Picard, R. W. (2019). Engineering Music to Slow Breathing and Invite Relaxd Physiology. In *2019 8th international conference on affective computing and intelligent interaction (acii)* (Vol. abs/1907.08844, pp. 276-282).
- Li, Q., Wang, X., Wang, S., Xie, Y., Xie, Y., & Li, S. (2020). More flexible integration of functional systems after musical training in young adults. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 28*.
- Li, X., Zhao, Z., Song, D., Zhang, Y., Pan, J., Wu, L., ... Wang, D. (2020). Latent factor decoding of multi-channel eeg for emotion recognition through autoencoder-like neural networks. *Frontiers in Neuroscience, 14*, 87. Retrieved from <https://www.frontiersin.org/article/10.3389/fnins.2020.00087> doi: 10.3389/fnins.2020.00087
- Liao, C.-Y., Chen, R.-C., Tai, S.-K., & Hendry. (2017). Using single point brain wave instrument to explore and verification of music frequency. In *2017 international conference on innovative and creative information technology (icitech)* (p. 1-6). doi: 10.1109/INNOCIT.2017.8319142
- Lin, X., Mahmud, S., Jones, E., Shaker, A., Miskinis, A., Kanan, S., & Kim, J.-H. (2020). Virtual Reality-Based Musical Therapy for Mental Health Management. In *2020 10th annual computing and communication workshop and conference (ccwc)* (p. 0948-0952). doi: 10.1109/CCWC47524.2020.9031157
- LingHu, Y.-f., & Shu, H. (2018). ARM-Based Feedback System For Athletes' Psychological Adjustment. In *2018 17th international symposium on distributed computing and applications for business engineering and science (dcabes)* (p. 72-75). doi: 10.1109/DCABES.2018.00028
- Liu, W., Zhang, C., Wang, X., Xu, J., Chang, Y., Ristaniemi, T., & Cong, F. (2020). Functional connectivity of major depression disorder using ongoing eeg during music perception. *Clinical Neurophysiology, 131*(10), 2413-2422. doi: <https://doi.org/10.1016/j.clinph.2020.06.031>
- Lopes, P., Liapis, A., & Yannakakis, G. N. (2019). Modelling Affect for Horror Soundscapes. *IEEE Transactions on Affective Computing, 10*(2), 209-222. doi: 10.1109/TAFFC.2017.2695460

- Lourinho, B. B. A. S., & Ramos, W. F. (2019). O envelhecimento, o cuidado com o idoso e a doença de alzheimer. *enciclopédia biosfera*, 723.
- Lubetzky, A. V., Kelly, J., Wang, Z., TaghaviDilamani, M., Gospodarek, M., Fu, G., ... Hujsak, B. (2019). Head Mounted Display Application for Contextual Sensory Integration Training: Design, Implementation, Challenges and Patient Outcomes. In *2019 international conference on virtual rehabilitation (icvr)* (p. 1-7). doi: 10.1109/ICVR46560.2019.8994437
- Lui, J. H., Samani, H., & Tien, K.-Y. (2017). An affective mood booster robot based on emotional processing unit. In *2017 international automatic control conference (cacs)* (p. 1-6). doi: 10.1109/CACS.2017.8284239
- Lui, S., & Grunberg, D. (2017). Using skin conductance to evaluate the effect of music silence to relieve and intensify arousal. In *2017 international conference on orange technologies (icot)* (p. 91-94). doi: 10.1109/ICOT.2017.8336096
- Lv, C., Li, S., & Huang, L. (2018). Music Emotions Recognition Based on Feature Analysis. In *2018 11th international congress on image and signal processing, biomedical engineering and informatics (cisp-bmei)* (p. 1-5). doi: 10.1109/CISP-BMEI.2018.8633223
- Lyu, M.-J., & Yuan, S.-M. (2020). Cloud-Based Smart Dog Music Therapy and Pneumonia Detection System for Reducing the Difficulty of Caring for Patients With Dementia. *IEEE Access*, 8, 20977-20990. doi: 10.1109/ACCESS.2020.2969482
- Malheiro, R., Panda, R., Gomes, P., & Paiva, R. P. (2018). Emotionally-relevant features for classification and regression of music lyrics. *IEEE Transactions on Affective Computing*, 9(2), 240-254. doi: 10.1109/TAFFC.2016.2598569
- Marimpis, A. D., Dimitriadis, S. I., & Goebel, R. (2020). A multiplex connectivity map of valence-arousal emotional model. *IEEE Access*, 8, 170928-170938. doi: 10.1109/ACCESS.2020.3025370
- Marosi-Holzberger, E., Prieto-Corona, D. M. B., Yáñez-Téllez, M. G., Rodríguez-Camacho, M. A., Rodríguez-Camacho, H., & Guerrero-Juarez, V. (2013). Quantitative Spectral EEG Assessments During Affective States Evoked By The Presentation Of The International Affective Pictures. *Journal of Behavior, Health & Social Issues*, 14.
- Matsumoto, K., & Sasayama, M. (2018). Lyric Emotion Estimation Using Word Embedding Learned from Lyric Corpus. In *2018 IEEE 4th international conference on computer and communications (iccc)* (p. 2295-2301). doi: 10.1109/CompComm.2018.8780811
- Mehta, D., Siddiqui, M. F. H., & Javaid, A. (2019, 04). Recognition of emotion intensities using machine learning algorithms: A comparative study. *Sensors*, 19. doi: 10.3390/s19081897
- Meska, M. H. G., Mano, L. Y., Silva, J. P., Junior, G. A. P., & Mazzo, A. (2020). Reconhecimento de emoções para ambiente clínico simulado com uso de odores desagradáveis: estudo quase experimental. *Revista Latino-A / Enfermagem*, 8.
- Mideska, K. G., Singh, A., Hoogenboom, N., Hellriegel, H., Krause, H., Schnitzler, A., ... Muthuraman, M. (2016). Comparison of imaging modalities and source-localization algorithms in locating the induced activity during deep brain stimulation of the STN. In *Annual international conference of the IEEE engineering in medicine and biology society* (pp. 105-108). doi: 10.1109/EMBC.2016.7590651
- Mitrpanont, J., Phandhu-fung, J., Klubdee, N., Ratanalaor, S., Pratiphakorn, P., Damrongvanakul, K., ... Mitrpanont, T. (2017). iCare-Stress: Caring system for stress. In *2017 6th ict international student project conference (ict-ispc)* (p. 1-4). doi: 10.1109/ICT-ISPC.2017.8075319
- Mo, S., & Niu, J. (2019). A novel method based on ompgw method for feature extraction in automatic music mood classification. *IEEE Transactions on Affective Computing*, 10(3), 313-324. doi: 10.1109/TAFFC.2017.2724515
- Mulla, F., Eya, E., Ibrahim, E., Alhaddad, A., Qahwaji, R., & Abd-Alhameed, R. (2017). Neurological assessment of music therapy on the brain using Emotiv Epoc. In *2017 internet technologies and applications (ita)* (p. 259-263). doi: 10.1109/ITECHA.2017.8101950
- Murad, D., Ye, F., Barone, M., & Wang, Y. (2017). Motion initiated music ensemble with sensors for motor rehabilitation. In *2017 international conference on orange technologies (icot)* (p. 87-90). doi: 10.1109/ICOT.2017.8336095
- Nalepa, G. J., Kutt, K., & Bobek, S. (2019). Mobile platform for affective context-aware systems. *Future Generation Computer Systems*, 92, 490-503. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0167739X17312207> doi: <https://doi.org/10.1016/j.future.2018.02.033>
- Navarathna, R., Carr, P., Lucey, P., & Matthews, I. (2019). Estimating audience engagement to predict movie ratings. *IEEE Transactions on Affective Computing*, 10(1), 48-59. doi: 10.1109/TAFFC.2017.2723011
- Navarro-Tuch, S. A., Solís-Torres, R., Bustamante-Bello, R., López-Aguilar, A. A., González-Archundia, G., & HernándezGonzález, O. (2018). Variation of facial expression produced by acoustic stimuli. In *2018 international conference on mechatronics, electronics and automotive engineering (icmeae)* (p. 60-64). doi: 10.1109/ICMEAE.2018.00018
- Nawa, N. E., Callan, D. E., Mokhtari, P., Ando, H., & Iversen, J. (2018). Decoding music-induced experienced emotions using functional magnetic resonance imaging - Preliminary results. In *2018 international joint conference on neural networks (ijcnn)* (p. 1-7). doi: 10.1109/IJCNN.2018.8489752
- Nemati, S., & Naghsh-Nilchi, A. R. (2017). Exploiting evidential theory in the fusion of textual, audio, and visual modalities for affective music video retrieval. In *2017 3rd international conference on pattern recognition and image analysis (ipria)* (p. 222-228). doi: 10.1109/PRIA.2017.7983051
- Nichols, E., Szoeko, C. E., Vollset, S., Abbasi, N., Abd-Allah, F., Abdela, J., ... Barker-Collo, S. e. a. (2019). Global, regional, and national burden of Alzheimer's disease and other dementias, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology*, 18, P88-106. doi: 10.1016/S1474-4422(18)30403-4
- NRC. (2001). Preparing for an Aging World: The Case for Cross-National Research [Computer software manual]. Retrieved from <https://www.ncbi.nlm.nih.gov/books/NBK98375/> (Last accessed: 2021 Jul. 05)

- Oliveira, E., & Jaques, P. A. (2013). Classificação de emoções básicas através de imagens capturadas em vídeos de baixa resolução. *Revista Brasileira de Computação Aplicada*, 5, 40–54.
- Ortegon-Sarmiento, T., Penuela, L., & Uribe-Quevedo, A. (2020). Low Back Pain Attenuation Employing Virtual Reality Physiotherapy. In *2020 22nd symposium on virtual and augmented reality (svr)* (p. 169-173). doi: 10.1109/SVR51698.2020.00037
- O'Toole, P., Glowinski, D., & Mancini, M. (2019). Understanding chromaesthesia by strengthening auditory -visual-emotional associations. In *2019 8th international conference on affective computing and intelligent interaction (acii)* (p. 1-7). doi: 10.1109/ACII.2019.8925465
- Pais, M., Martinez, L., Ribeiro, O., Loureiro, J., Fernandez, R., Valiengo, L., ... Forlenza, O. (2020). Early diagnosis and treatment of Alzheimer's disease: new definitions and challenges. *Braz J Psychiatry*, 431–441.
- Panda, R., Malheiro, R., & Paiva, R. P. (2020). Novel audio features for music emotion recognition. *IEEE Transactions on Affective Computing*, 11(4), 614-626. doi: 10.1109/TAFFC.2018.2820691
- Parra, F., Scherer, S., Benezeth, Y., Tsvetanova, P., & Tereno, S. (2019). (revised may 2019) development and cross-cultural evaluation of a scoring algorithm for the biometric attachment test: Overcoming the challenges of multimodal fusion with "small data". *IEEE Transactions on Affective Computing*, 1-1. doi: 10.1109/TAFFC.2019.2921311
- Paxiuba, C. M., & Lima, C. P. (2020). An Experimental Methodological Approach Working Emotions and Learning Using Facial Expressions Recognition. *Brazilian Journal of Computers in Education*, 28, 92–114.
- Picard, R. W. (1997). *Affective Computing*. Massachusetts, USA: MIT Press, 2.
- Plut, C., & Pasquier, P. (2019). Music Matters: An empirical study on the effects of adaptive music on experienced and perceived player affect. In *2019 IEEE conference on games (cog)* (p. 1-8). doi: 10.1109/CIG.2019.8847951
- Poria, S., Cambria, E., Bajpai, R., & Hussain, A. (2017). A review of affective computing: From unimodal analysis to multimodal fusion. *Information Fusion*, 37, 98–125.
- Qin, Y., Zhang, H., Wang, Y., Mao, M., & Chen, F. (2020). 3d music impact on autonomic nervous system response and its therapeutic potential. In *2020 IEEE conference on multimedia information processing and retrieval (mipr)* (p. 364-369). doi: 10.1109/MIPR49039.2020.00080
- Rahman, J. S., Gedeon, T., Caldwell, S., & Jones, R. (2020). Brain Melody Informatics: Analysing Effects of Music on Brainwave Patterns. In *2020 international joint conference on neural networks (ijcnn)* (p. 1-8). doi: 10.1109/IJCNN48605.2020.9207392
- Rahman, J. S., Gedeon, T., Caldwell, S., Jones, R., Hossain, M. Z., & Zhu, X. (2019). Melodious Micro-frissons: Detecting Music Genres From Skin Response. In *2019 international joint conference on neural networks (ijcnn)* (p. 1-8). doi: 10.1109/IJCNN.2019.8852318
- Ramdinwawii, E., & Mittal, V. K. (2017). Effect of Different Music Genre: Attention vs. Meditation. In *2017 Seventh International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)* (pp. 135–140). doi: 10.1109/ACIIW.2017.8272603
- Ramírez, A. V., Hornero, G., Royo, D., Aguilar, A., & Casas, O. (2020). Assessment of Emotional States Through Physiological Signals and Its Application in Music Therapy for Disabled People. *IEEE Access*, 8, 127659-127671. doi: 10.1109/ACCESS.2020.3008269
- Ranjkar, E., Rafatnejad, R., Nobaveh, A. A., Meghdari, A., & Alemi, M. (2019). Design, Fabrication, and Evaluation of the “Maya” Social Robot. In *2019 7th international conference on robotics and mechatronics (icrom)* (p. 52-62). doi: 10.1109/ICRoMb48714.2019.9071795
- Ricci, G. (2019). Social Aspects of Dementia Prevention from a Worldwide to National Perspective: A Review on the International Situation and the Example of Italy. *Behavioural neurology*, 2019, 8720904. doi: 10.1155/2019/8720904
- Rizos, G., & Schuller, B. (2019). Modelling Sample Informativeness for Deep Affective Computing. In *Icassp 2019 - 2019 IEEE international conference on acoustics, speech and signal processing (icassp)* (p. 3482-3486). doi: 10.1109/ICASSP.2019.8683729
- Rizzi, L., Rosset, I., & Roriz-Cruz, M. (2014). Global epidemiology of dementia: Alzheimer's and vascular types. *BioMed research international*, 2014, 908915. doi: 10.1155/2014/908915
- Rushambwa, M. C., & Mythili, A. (2017). Impact assessment of mental subliminal activities on the human brain through neuro feedback analysis. In *2017 third international conference on biosignals, images and instrumentation (icbsii)* (p. 1-6). doi: 10.1109/ICBSII.2017.8082299
- Saha, D. P., Bortz, B. C., Huang, W., Martin, T. L., & Knapp, R. B. (2016). Affect-Aware Intelligent Environment Using Musical Cues as an Emotion Learning Framework. In *2016 12th international conference on intelligent environments (ie)* (p. 178-181). doi: 10.1109/IE.2016.39
- Sanders, Q., Chan, V., Augsburg, R., Cramer, S. C., Reinkensmeyer, D. J., & Do, A. H. (2020). Feasibility of Wearable Sensing for In-Home Finger Rehabilitation Early After Stroke. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(6), 1363-1372. doi: 10.1109/TNSRE.2020.2988177
- Santhosh, A. K., Sangilirajan, M., Nizar, N., Radhamani, R., Kumar, D., Bodda, S., & Diwakar, S. (2020). Computational exploration of neural dynamics underlying music cues among trained and amateur subjects. *Procedia Computer Science*, 171, 1839-
1847. Retrieved from <https://www.sciencedirect.com/science/article/pii/S1877050920311789> (Third International Conference on Computing and Network Communications (CoCoNet'19)) doi: <https://doi.org/10.1016/j.procs.2020.04.197>
- Savery, R., Rose, R., & Weinberg, G. (2019). Establishing Human-Robot Trust through Music-Driven Robotic Emotion Prosody and Gesture. In *2019 28th IEEE international conference on robot and human interactive communication (ro-man)* (p. 1-7). doi: 10.1109/RO-MAN46459.2019.8956386

- Sawata, R., Ogawa, T., & Haseyama, M. (2019). Novel Audio Feature Projection Using KDLPCA-Based Correlation with EEG Features for Favorite Music Classification. *IEEE Transactions on Affective Computing*, 10(3), 430-444. doi: 10.1109/TAFFC.2017.2729540
- Scardua, D. A., & Marques, K. (2018). Estudo da Identificação de Emoções Através da Inteligência Artificial. *Multivix, Edu.*
- Scott, L. (2017). Creating Opera for Mobile Media: Artistic Opportunities and Technical Limitations. In *2017 14th international symposium on pervasive systems, algorithms and networks 2017 11th international conference on frontier of computer science and technology 2017 third international symposium of creative computing (ispan-fcst-isc)* (p. 477-484). doi: 10.1109/ISPAN-FCST-ISCC.2017.86
- Seanglidet, Y., Lee, B. S., & Yeo, C. K. (2016). Mood prediction from facial video with music “therapy” on a smartphone. In *2016 wireless telecommunications symposium (wts)* (p. 1-5). doi: 10.1109/WTS.2016.7482034
- Shan, Y., Chen, T., Yao, L., Wu, Z., Wen, W., & Liu, G. (2018). Remote detection and classification of human stress using a depth sensing technique. In *2018 first asian conference on affective computing and intelligent interaction (acii asia)* (p. 1-6). doi: 10.1109/ACIIAsia.2018.8470364
- Sheffield, J., Karcher, N., & Barch, D. (2018). Cognitive deficits in psychotic disorders: a lifespan perspective. *Neuropsychol Rev*, 509—533.
- Shen, X., Hu, X., Liu, S., Song, S., & Zhang, D. (2020). Exploring EEG microstates for affective computing: decoding valence and arousal experiences during video watching*. In *2020 42nd annual international conference of the ieee engineering in medicine biology society (embc)* (p. 841-846). doi: 10.1109/EMBC44109.2020.9175482
- Shukla, J., Cristiano, J., Oliver, J., & Puig, D. (2019). Robot Assisted Interventions for Individuals with Intellectual Disabilities: Impact on Users and Caregivers. *International Journal of Social Robotics*, 11, 631–649. doi: 10.1007/s12369-019-00527-w
- Sonawane, A., Inamdar, M. U., & Bhargale, K. B. (2017). Sound based human emotion recognition using MFCC and multiple SVM. *International Conference on Information, Communication, Instrumentation and Control (ICICIC)*.
- Sondhi, M. (1968). New Methods of Pitch Extraction. *IEEE Trans. Audio and Electroacoustics*, 262–266.
- Song, M., Yang, Z., Baird, A., Parada-Cabaleiro, E., Zhang, Z., Zhao, Z., & Schuller, B. (2019). Audiovisual Analysis for Recognising Frustration during Game-Play: Introducing the Multimodal Game Frustration Database. In *2019 8th international conference on affective computing and intelligent interaction (acii)* (p. 517-523). doi: 10.1109/ACII.2019.8925464
- Soysal, O. M., Kiran, F., & Chen, J. (2020). Quantifying Brain Activity State: EEG analysis of Background Music in A Serious Game on Attention of Children. In *2020 4th international symposium on multidisciplinary studies and innovative technologies (ismsit)* (p. 1-7). doi: 10.1109/ISMSIT50672.2020.9255308
- Sra, M., Vijayaraghavan, P., Rudovic, O., Maes, P., & Roy, D. (2017). DeepSpace: Mood-Based Image Texture Generation for Virtual Reality from Music. In *2017 ieee conference on computer vision and pattern recognition workshops (cvprw)* (p. 2289-2298). doi: 10.1109/CVPRW.2017.283
- Stappen, L., Karas, V., Cummins, N., Ringeval, F., Scherer, K., & Schuller, B. (2019). From Speech to Facial Activity: Towards Cross-modal Sequence-to-Sequence Attention Networks. In *2019 ieee 21st international workshop on multimedia signal processing (mmsp)* (p. 1-6). doi: 10.1109/MMSP.2019.8901779
- Su, J. H., Liao, Y. W., Wu, H. Y., & Zhao, Y. W. (2020). Ubiquitous music retrieval by context-brain awareness techniques. In *2020 ieee international conference on systems, man, and cybernetics (smc)* (p. 4140-4145). doi: 10.1109/SMC42975.2020.9282963
- Subramaniam, G., Verma, J., Chandrasekhar, N., C., N. K., & George, K. (2018). Generating Playlists on the Basis of Emotion. In *2018 ieee symposium series on computational intelligence (ssci)* (p. 366-373). doi: 10.1109/SSCI.2018.8628673
- Suhaimi, N. S., Yuan, C. T. B., Teo, J., & Mountstephens, J. (2018). Modeling the affective space of 360 virtual reality videos based on arousal and valence for wearable EEG-based VR emotion classification. In *2018 ieee 14th international colloquium on signal processing its applications (cspa)* (p. 167-172). doi: 10.1109/CSPA.2018.8368706
- Syed, S., Chagani, S., Hafeez, M., Timothy, S., & Zahid, H. (2018, 10). Sign recognition system for differently abled people. In *Proceedings of tencon 2018* (p. 1148-1153). doi: 10.1109/TENCON.2018.8650365
- Tarvainen, J., Laaksonen, J., & Takala, T. (2017). Computational and Perceptual Determinants of Film Mood in Different Types of Scenes. In *2017 ieee international symposium on multimedia (ism)* (p. 185-192). doi: 10.1109/ISM.2017.10
- Teo, J., & Chia, J. T. (2018). Deep Neural Classifiers For Eeg-Based Emotion Recognition In Immersive Environments. In *2018 international conference on smart computing and electronic enterprise (icscee)* (p. 1-6). doi: 10.1109/ICSCEE.2018.8538382
- Thammasan, N., Hagad, J. L., Fukui, K.-i., & Numao, M. (2017). Multimodal stability-sensitive emotion recognition based on brainwave and physiological signals. In *2017 seventh international conference on affective computing and intelligent interaction workshops and demos (aciw)* (p. 44-49). doi: 10.1109/ACIIW.2017.8272584
- Tiple, B. S., Joshi, P. P., & Patwardhan, M. (2016). An efficient framework for recommendation of Hindustani Art Music. In *2016 international conference on computing communication control and automation (iccubea)* (p. 1-5). doi: 10.1109/ICCUBEA.2016.7860008
- Tiwari, A., & Tiwari, R. (2017). Design of a brain computer interface for stress removal using yoga a smartphone application. In *2017 international conference on computing, communication and automation (iccca)* (p. 992-996). doi: 10.1109/CCAA.2017.8229939

- UN. (2020). World Population Ageing 2020: Highlights [Computer software manual]. Retrieved from https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/undesa_pd-2020_world_population_ageing_highlights.pdf (Last accessed: 2021 Jul. 05)
- Verma, G., Dhekane, E. G., & Guha, T. (2019). Learning affective correspondence between music and image. In *Icassp 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (p. 3975-3979). doi: 10.1109/ICASSP.2019.8683133
- Vicencio-Martínez, B., A. A. and Tovar-Corona, & Garay-Jiménez, L. I. (2019). Emotion Recognition System Based on Electroencephalography. *International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*, 11–13.
- Vinayagasundaram, B., Mallik, R., Aravind, M., Aarthi, R. J., & Senthilraj, S. (2016). Building a generative model for affective content of music. In *2016 International Conference on Recent Trends in Information Technology (ICRTIT)* (p. 1-6). doi: 10.1109/ICRTIT.2016.7569588
- Wang, K., Wen, W., & Liu, G. (2016). The autonomic nervous mechanism of music therapy for dental anxiety. In *2016 13th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)* (p. 289-292). doi: 10.1109/ICCWAMTIP.2016.8079858
- Wang, X., Wenya, L., Toivainen, P., Tapani, R., & Cong, F. (2020). Group analysis of ongoing EEG data based on fast doublecoupled nonnegative tensor decomposition. *Journal of Neuroscience Methods*, 330, 108502. doi: <https://doi.org/10.1016/j.jneumeth.2019.108502>
- Wang, Y., & Kosinski, M. (2018). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of Personality and Social Psychology*, 246–257. doi: 10.1037/pspa0000098
- Wardana, A. Y., Ramadijanti, N., & Basuki, A. (2018). Facial Expression Recognition System for Analysis of Facial Expression Changes when Singing. In *2018 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IESKIC)* (p. 98-104). doi: 10.1109/KCIC.2018.8628578
- WHO. (2017). Global action plan on the public health response to dementia (2017-2025) [Computer software manual]. <https://apps.who.int/iris/bitstream/handle/10665/259615/9789241513487-eng.pdf;jsessionid=4DA480FA93471AC53988E52B35F416D8?sequence=1> (Last accessed: 2021 Jul. 05)
- WHO. (2019). Risk reduction of cognitive decline and dementia: WHO Guidelines [Computer software manual]. <https://www.who.int/publications/i/item/risk-reduction-of-cognitive-decline-and-dementia> (Last accessed: 2021 Jul. 05)
- Wingerden, E. v., Barakova, E., Lourens, T., & Sterkenburg, P. S. (2020). Robot-mediated therapy to reduce worrying in persons with visual and intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities*, 39, 229–238. doi: 10.1111/jar.12801
- Wolfe, H., Peljhan, M., & Visell, Y. (2020). Singing robots: How embodiment affects emotional responses to non-linguistic utterances. *IEEE Transactions on Affective Computing*, 11(2), 284-295. doi: 10.1109/TAFFC.2017.2774815
- Wu, C.-H., Huang, Y.-M., & Hwang, J.-P. (2016). Review of affective computing in education/learning: Trends and challenges. *British Journal of Educational Technology*, 47(6), 1304–1323.
- Wu, Y.-C., & Chen, H. H. (2016). Generation of Affective Accompaniment in Accordance With Emotion Flow. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 24(12), 2277-2287. doi: 10.1109/TASLP.2016.2603006
- Xing, B., Zhang, K., Zhang, L., Wu, X., Dou, J., & Sun, S. (2019). Image–music synesthesia-aware learning based on emotional similarity recognition. *IEEE Access*, 7, 136378-136390. doi: 10.1109/ACCESS.2019.2942073
- Xu, T., Yin, R., Shu, L., & Xu, X. (2019). Emotion Recognition Using Frontal EEG in VR Affective Scenes. In *2019 IEEE MTT-S International Microwave Biomedical Conference (IMBIOC)* (Vol. 1, p. 1-4). doi: 10.1109/IMBIOC.2019.8777843
- Xu, X., Deng, J., Coutinho, E., Wu, C., Zhao, L., & Schuller, B. W. (2019). Connecting Subspace Learning and Extreme Learning Machine in Speech Emotion Recognition. *IEEE Transactions on Multimedia*, 21(3), 795-808. doi: 10.1109/TMM.2018.2865834
- Yamada, Y., & Ono, Y. (2019). Detection of Music Preferences using Cerebral Blood Flow Signals. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 490–493). doi: 10.1109/EMBC.2019.8856351
- Yi-Hsiang, M., Han, Y., Lin, J.-Y., Cosentino, S., Nishio, Y., Oshiyama, C., & Takanishi, A. (2018). A Synchronization Feedback System to Improve Interaction Correlation in Subjects With Autism Spectrum Disorder. In *2018 9th International Conference on Awareness Science and Technology (ICAST)* (p. 285-290). doi: 10.1109/ICAwST.2018.8517233
- Zhang, Z., Han, J., Coutinho, E., & Schuller, B. (2019). Dynamic difficulty awareness training for continuous emotion prediction. *IEEE Transactions on Multimedia*, 21(5), 1289-1301. doi: 10.1109/TMM.2018.2871949
- Zhang, Z., Han, J., Deng, J., Xu, X., Ringeval, F., & Schuller, B. (2018). Leveraging Unlabeled Data for Emotion Recognition With Enhanced Collaborative Semi-Supervised Learning. *IEEE Access*, 6, 22196-22209. doi: 10.1109/ACCESS.2018.2821192
- Zhao, J., Mao, X., & Chen, L. (2019). Speech emotion recognition using deep 1D and 2D CNN LSTM networks. *Biomedical Signal Processing and Control*, 47, 312–323.
- Zhu, L., Tian, X., Xu, X., & Shu, L. (2019). Design and Evaluation of the Mental Relaxation VR Scenes Using Forehead EEG Features. In *2019 IEEE MTT-S International Microwave Biomedical Conference (IMBIOC)* (Vol. 1, p. 1-4). doi: 10.1109/IMBIOC.2019.8777812