Revista de Informática Teórica e Aplicada - RITA - ISSN 2175-2745
Vol. 29, Num. 02 (2022) 35-47

RESEARCH ARTICLE

Evaluating the Causal Effect of Multimedia and Affective Temperament in Felt Emotion and Liking

Avaliação do Efeito Causal de Multimídia e Temperamento Afetivo na Emoção Sentida e no Gostar

Rafaela de Amorim Barbosa Silva [1]*, Eanes Torres Pereira[1]

Abstract: In this paper, we propose an approach to evaluate the causal effect of videos on subjects who watched movies from the LIRIS-ACCEDE dataset and from whom the following information was collected: affective temperaments, gender, and electroencephalography (EEG) signals. The affective temperament was obtained by analyzing their answers to the AFECT questionnaire. Evidence was collected from specialized literature to design a Structural Causal Model to be subjected to Do-Calculus Causal Inference. Video concepts were extracted to characterize the major video content after k-means clustering. Information from 15 volunteers was analyzed and the effects of video content, affective temperament, and gender on emotion response and liking were computed. Higher Order Crossings (HOC) were extracted from EEG signals and the features were clustered and used as intermediate evidence of affective influence. This research provides answers for the following questions about the specific watched videos: (i) How does gender affect the felt emotion and liking? (ii) How does the affective temperament of a person affect felt emotion and liking? and (iii) How does the content of a video affect felt emotion and liking? The main contribution of this paper is in the proposed methodology which can be applied to any similar dataset to investigate the causal relationships of video content and affective temperament on the emotion of the audience.
Keywords: Affective Computing — Causality — Multimedia — Affective Temperament

Resumo: Neste artigo, propomos uma abordagem para avaliar o efeito causal de vídeos em sujeitos que assistiram a filmes do conjunto de dados LIRIS-ACCEDE e dos quais foram coletadas as seguintes informações: temperamentos afetivos, gênero e sinais de eletroencefalografia (EEG). O temperamento afetivo foi obtido por meio da análise das respostas ao questionário AFECT. Evidências foram coletadas da literatura especializada para projetar um Modelo Causal Estrutural a ser submetido à Inferência Causal utilizando Do-Calculus. Conceitos de vídeo foram extraídos para caracterizar o conteúdo de vídeo principal após o agrupamento k-means. As informações de 15 voluntários foram analisadas e os efeitos do conteúdo do vídeo, temperamento afetivo e gênero na resposta emocional e gosto foram computados. Cruzamentos de Ordem Superior (HOC - Higher Order Crossings) foram extraídos de sinais de EEG e as características foram agrupadas e usadas como evidência intermediária de influência afetiva. Esta pesquisa fornece respostas para as seguintes questões sobre os vídeos assistidos específicos desta pesquisa: (i) Como o gênero afeta a emoção sentida e o gosto? (ii) Como o temperamento afetivo de uma pessoa afeta as emoções e gostos sentidos? e (iii) como o conteúdo de um vídeo afeta as emoções e gostos sentidos? A principal contribuição deste artigo está na metodologia proposta, que pode ser aplicada a qualquer conjunto de dados semelhante para investigar as relações causais do conteúdo do vídeo e o temperamento afetivo na emoção do público.
Palavras-Chave: Computação Afetiva — Causalidade — Multimídia — Temperamento Afetivo

[1] *UASC/CEEI, Universidade Federal de Campina Grande (UFCG), Campina Grande - Paraíba, Brazil*
***Corresponding author**: eanes@computacao.ufcg.edu.br
DOI: http://dx.doi.org/10.22456/2175-2745.116017 • **Received:** 16/06/2021 • **Accepted:** 13/04/2022

1. Introduction

1.1 Video Affective Content Analysis

With the advances in internet technology which allows the availability of high-speed connections, the world population is producing video content at an accelerating speed. The Cisco Annual Internet Report (2018-2023) [1] shows that video streaming is one of the most demanding technologies in today's home. Due to the increasing amount of available video, it is necessary to create methodologies to analyze, access, and retrieve it fast and considering semantic content. One such content is emotional or affective. There are two main approaches for video affective content analysis: direct (uses audiovisual features), and implicit (uses spontaneous nonverbal responses from subjects) [2].

An important problem with affective video content analysis is related to the causal effect of video content on the audience. To know the causal effect of video content on the audience is important in many contexts, for example:

- Professionals who use biofeedback and neurofeedback could evoke specific emotional states in their clients using videos if they know which video content evokes each emotion;
- The advertising professionals could improve their commercial videos if they know *a priori* which content evoke the desired emotions;
- Streaming platforms could better recommend videos using affective knowledge of the videos.

Wang and Ji [2] affirms that the two main goals of video affective content analysis are: (i) to identify videos that can evoke certain emotions in the users; (ii) to automatically tag each video clip by its affective content. Although Wang and Ji [2] present a comprehensive review on affective content-based analysis, they did not mention studies on the causal effect of video content on audience emotion.

The study of video affective content involves the labeling of videos using dimensional or discrete labels. Discrete labels are based on models similar to the facial expression model proposed by Ekman [3] who states that there are six universal facial expressions associated with specific emotions and a neutral expression. Dimensional labels are based on models similar to the circumplex model of affect [4] which criticizes the theory of basic emotions and proposes that all affective states arise from two dimensions: one related to valence (a pleasure–displeasure continuum) and the other to arousal, or alertness. The Causal Inference Model proposed in this paper uses a valence-arousal model.

1.2 Related Work

During a related work search, no paper was found proposing the analysis of the causal effect of multimedia in human emotion. This subsection is mainly composed of discussion on papers that performed research on the following topics: EEG-based gender classification, temperament classification, EEG-based emotion classification, and EEG-based liking classification. The rationale for discussing only papers of those topics is the fact that the proposed SCM (Structured Causal Model), which can be seen in Figure 8, was created considering domain knowledge.

No paper about Bayesian networks for causal inference was included in this section because Bayesian networks are incapable of adequately capturing and representing causal knowledge [5]. In order to represent causal knowledge, Judea Pearl [6] proposes the use of structured causal models, which are formed by three components: (i) a set of endogenous and a set of exogenous variables; (ii) a directed acyclic graph (DAG), in which causation is represented by the direction of arrows; (iii) causal functions indicating the degree of causation among variables. The details of this theory are given in Subsection 2.1

Lara et al. [7] proposed a model and scale to measure emotional and composite temperament called AFECT. Lara et al. [7] affirm that temperament is conceived as a self-regulated system with six emotional dimensions: volition, anger, inhibition, sensitivity, coping, and control. In the AFECT model, those emotional dimensions are combined to result in 12 affective temperaments, namely: depressive, anxious, apathetic, obsessive, cyclothymic, dysphoric, irritable, volatile, disinhibited, hyperthymic, and euphoric. The questionnaire proposed by Lara [8] was used to collect information from volunteers for the research presented in this paper. The questionnaire is composed of two sections: emotional dimensions (60 questions) and affective types (12 questions). The questions of the emotional dimensions section are grouped into 10 groups. The answers for the emotional dimensions section are in a 7-point Likert scale and the answers for the affective types are in a 5-point Likert scale. The model proposed by Lara [8] explains how to combine the answers to classify the temperaments of the questionnaire respondents. Therefore, the Temperament variable was obtained by applying the AFECT model and collecting answers from volunteers.

The main reason why the EEG is the central variable in the SCM proposed in Figure 8 is the already proven possibility of classification of human gender as well as emotional states using EEG. In this subsection, some papers on EEG-based gender and emotion classification are discussed to demonstrate the feasibility of this type of classification and justify the importance of EEG features in the proposed SCM model.

Kaur et al. [9] proposed to extract Fractal Dimension Features (FDF) from EEG signals and train a Support Vector Machine (SVM) to classify emotions in three categories: calm, anger, and happiness. EEG signals were collected from 10 subjects while watching video clips. The EEG signals were captured from 14 electrodes (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4) and they were labeled according to the Valence-Arousal model. Before FDF extraction, signals were segmented into 3300 1 second segments. The overall accuracy of 60% was achieved, the maximum accuracy was obtained to the happy emotion (80%) and the lowest accuracy was obtained to the anger emotion (40%). Kaur et al. [9] did not mention from which database they obtained the video clips, and they also did not mention the distribution of how many EEG samples per subject was used.

Chakladar and Chakraborty [10] proposed to use Higher Order Statistics features (mean, standard deviation, skewness, and kurtosis) to train a Linear Discriminant Analysis (LDA) classifier for discriminating among four categories of emotions: positive, negative, angry, and harmony. The DEAP dataset was used for training and testing. Chakladar and Chakraborty [10] proposed a correlation-based subset selection dimension reduction approach that considers information from the number of channels, feature class correlation of each channel (using entropy), and inter-channel correlation. The selected channels were FP1, F3, T7, and PO3. The overall accuracy achieved by their model was 82%. There is a tendency in papers that used the DEAP dataset to report their results using accuracy. However, the DEAP dataset emotional classes grouped by quadrant in the Valence-Arousal plane are unbalanced. Therefore, accuracy is not the best way to report and compare results achieved from the DEAP dataset.

The research of Razumnikova [11] draws evidence to the fact that men's and women's brains present different EEG patterns during creative thinking. In her study, Razumnikova [11] examined EEG patterns from 36 males and 27 females subjects. Personality and intelligence scores of all subjects were measured 4-6 weeks before the EEG acquisition. During EEG acquisition, the subject was asked to solve problems that need creative thinking. The EEG differences were computed for two groups for each gender: good performers and non-performers. ANOVA was applied to the data and the task-induced gender differences were computed only for alpha and beta bands. Razumnikova [11] concludes that the results demonstrate the existence of gender differences in attentional processes.

Barry et al. [12] investigated the intra-hemispheric (F3 - O1, F4 - O2, FP1 - F3, FP2 - F4, T3 - T5, T4 - T6, C3 - P3, C4 - P4) and inter-hemispheric (FP1 - FP2, F7 - F8, F3 - F4, C3 - C4, T3 - T4, T5 - T6, P3 - P4, O1 - O2) EEG coherence as a function of age and gender in normal children. The study used EEG data collected from 80 children from 8 to 12 years old. Coherence between an electrode pair for a particular band (Delta, Alpha, and Beta) was computed as the cross-spectral power. The mean coherence for each pair of electrodes was grouped into regions and analyzed using MANOVA (Multivariate Analysis of Variance). Barry et al. [12] say that the results suggest systematic development of EEG coherence as a function of age from 8 to 12 years. Furthermore, the coherences also appeared to develop further in males than age-matched females. Therefore, this study shows the possibility of classifying children's age and gender using EEG signals.

In order to close this subsection, we present some discussion about the most common used features for emotion classification using EEG signals. Although the popularity of Deep Learning (DL) techniques has aroused the interest in using DL to learn features, Alarcão and Fonseca [13] showed that the most popular techniques for feature extraction in the context of emotion classification using EEG are not necessarily related with Deep Learning. Alarcão and Fonseca [13] discussed 99 papers on Emotion Identification and Recognition. Among the most used features for emotion classification in the papers analysed, there are: PSD (Power Spectral Density) and HOC (Higher Order Crossings). Due to the extraction simplicity associated with high capacity of representation, we chose HOC to perform the experiments in this paper.

1.3 Problem Statement

The goal of this study was to elaborate causal questions and analyze the factors that may have caused the emotions induced by the videos. Among the causal questions, social and psychological aspects were taken into consideration, *e.g.*, if the gender of people affect how they feel about a video, how much would it impact the emotion?

In this study, it was assumed that the following factors could affect **the emotion** of the watcher about the video and also **how much they like it**: (i) The gender of the participant; (ii) The affective temperament of the participant; and (iii) The content of the video.

The volunteers answered, on a questionnaire, how they felt watching the video (valence-arousal 9-point scale), represented in this model by Quadrant, and also if they liked the video, did not like it or felt neutral. The causal impact of the variables on each other was calculated based on the following questions: (i) How does gender affect the felt emotion and liking? (ii) How does the affective temperament of a person affect felt emotion and liking? and (iii) How does the content of a video affect felt emotion and liking?

2. Methods

2.1 Causal Inference

Causal Inference deals with the problem of verifying if X caused Y and at which extent X caused Y. There are many approaches for performing causal inference, two well-known are Propensity Scores [14] and Do-Calculus [6]. In this paper, we proposed a Causal Inference procedure for evaluating affective causal effect of watched movies in volunteers using the theory proposed by Pearl [6].

The best way to study a cause and effect relationship would be through a random experimentation. However, causal questions cannot always be addressed through randomized experiments because of physical or ethical impediments, or some other reason. Judea Pearl's work [6] on Causal Inference presents a manner of treating nonexperimental data as if it were from a randomized trial, through the $do(\cdot)$ tool, an intervention.

In Do-Calculus theory, a causal model is composed of three components, namely: (i) a set of endogenous and a set of exogenous variables; (ii) a directed acyclic graph (DAG), in which causation is represented by the direction of arrows;

(iii) causal functions indicating the degree of causation among variables. Due to the usage of a directed acyclic graph, the model used in Do-Calculus is also known as SCM (Structured Causal Model). There are two main approaches to create an SCM: using domain knowledge, or using causal search. In this paper, the SCM was proposed using domain knowledge.

2.1.1 Do-Calculus and Interventions

Rigorously speaking, Do-Calculus is the name given to a set of three rules indicating how and when to exchange observational probabilities by interventional probabilities and vice-versa. To understand the difference between observation and intervention, we may resort to a popular example used in this context, the "smoke causes cancer" example. If a government health foundation collects information for decades about smoking subjects, it will have observational data and would be able to answer questions like the ratios of smoking versus non-smoking people, or if smoking people with diabetes had a higher death rate than non-smoking people. However, to answer a question like "what if those people who have diabetes and are smokers had not smoked?", research would have to develop an experimental protocol asking volunteers who have diabetes and smoke to stop smoking. That type of experiment requires an intervention in the smoking habits of the volunteers, but a protocol like that is unfeasible. Causal inference proposes to use conditioned probabilities to infer information about interventions without actually experimenting.

The theory proposed by Pearl [6] is based on three main components:

- **Back-door adjustment**: Conditions to adjust conditional probabilities equations when there are arrows coming into the causal variable in the SCM, and there is a set of variable Z satisfying other specific conditions;
- **Front-door adjustment**: Criteria applied when the Back-door adjustment is not possible, back-door paths are blocked by the causal variables, and other specific conditions are satisfied;
- **Do-Calculus Rules**: (i) insertion/deletion of observations; (ii) intervention/observation exchange; (iii) insertion/deletion of actions.

The Back-door Adjustment Theorem [6] states that *if a set of variables Z satisfies the back-door criterion relative to (X, Y), then the causal effect of X on Y is identifiable and is given by the Equation* 1.

$$P(y \mid \hat{x}) = \sum_z P(y \mid x,z)P(z) \tag{1}$$

where $\hat{x}$ indicates that the intervention is occurring on the variable x, it is a different notation for $do(x)$ or $do(X=x)$. Another possible written possibility for indicating intervention on x would be $P(Y|X=x_i)$, where x_i is the specific value assumed by X.

The Front-door Adjustment Theorem [6] states that *if Z satisfies the front-door criterion relative to* (X,Y) *and if* $P(x,z) > 0$*, then the causal effect of X on Y is identifiable and is given by the Equation* 2.

$$P(y \mid \hat{x}) = \sum_z P(z \mid x) \sum_{x'} P(y \mid x',z)P(x') \tag{2}$$

where x' indicates the specific values assumed by the variable X. Notice that, in both Equations 1 and 2, the lowercase variables indicate specific values of the set of (uppercase) variables: Y, Z and X.

The following statements of the Rules of Do-Calculus were adapted from the paper [15]. Let X, Y, Z, and W be arbitrary disjoint sets of nodes in a causal DAG G. We denote by $G_{\overline{X}}$ the graph obtained by deleting from G all arrows pointing to nodes in X. Likewise, we denote by $G_{\underline{X}}$ the graph obtained by deleting from G all arrows emerging from nodes in X. To represent the deletion of both incoming and outgoing arrows, we use the notation $G_{\overline{X}\underline{Z}}$. The following three rules are valid for every interventional distribution compatible with G.

1. Insertion/deletion of observations: Equation 3.
2. Action/observation exchange: Equation 4.
3. Insertion/deletion of actions: Equation 5.

$$P(y \mid \hat{x},z,w) = P(y \mid \hat{x},w) \text{ IF } (Y \perp\!\!\!\perp Z \mid X,W)_{G_{\overline{X}}} \tag{3}$$

$$P(y \mid \hat{x},\hat{z},w) = P(y \mid \hat{x},z,w) \text{ IF } (Y \perp\!\!\!\perp Z) \mid X,W)_{G_{\overline{X},\underline{Z}}} \tag{4}$$

$$P(y \mid \hat{x},\hat{z},w) = P(y \mid \hat{x},w) \text{ IF } (Y \perp\!\!\!\perp Z \mid X,W)_{G_{\overline{X},\overline{Z(W)}}} \tag{5}$$

where $Z(W)$ is the set of Z-nodes that are not ancestors of any W-node in $G_{\overline{X}}$.

2.1.2 Graphical Components

The graph (or DAG) of the causal model in question must be examined, there must be certain graphical patterns and causal relationships. There are three main components that can be found in a causal graph: chain, fork and collider. They either "let" causality "flow" from a node X to a node Y, or they "block" it.

Assuming one wants to know the effect of X on Y (measured by the probability $P(Y|do(X))$), in the **chain** structural component, Fig. 1, the node X is an indirect cause of Y, in this structure the causality flows from X to Y passing trough node Z, thus X and Y are dependent. In the **fork** structural component, statistically the same happens, X and Y are dependent, but X is not a cause of Y, there is a correlation between them because the node Z is connecting them trough a common cause, the node Z can also be called a "confounder", the causality flowing from X to Y is spurious. The nodes X and Y are dependent in both cases above mentioned, yet, this dependence can be **blocked** if the probability $P(Y|X)$ is conditioned on Z, $P(Y|X,Z)$.

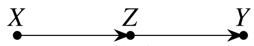

Figure 1. Chain structure

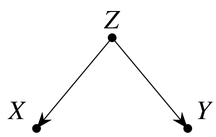

Figure 2. Fork structure

Similarly, one wants the effect of X on Y, in a **collider** structural component, the node Z is a common effect of X and Y, X and Z are dependent, likewise Y and Z, but X and Y are unconditionally independent. Although, if the probability $P(Y|X)$ were conditioned on Z, X and Y would become dependent ($P(Y|X,Z)$. Since Z is caused by both X and Y, when conditioned on a value $Z = z$, the probabilities are limited to cases in which Z takes that specific value, and any changes on X is compensated by Y, and vice-versa.

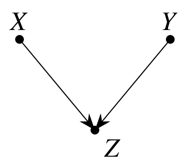

Figure 3. Collider structure

The structural components aforementioned are the building blocks of the graphical analysis required to make interventions with Do-Calculus. For example, on the first rule, Eqn. 3, there's a condition for the equality of the probabilities,

$$P(y \mid \hat{x}, z, w) = P(y \mid \hat{x}, w) \text{ IF } (Y \perp\!\!\!\perp Z \mid X, W)_{G_{\overline{X}}}$$

one cannot remove or add z as a condition if y is not independent from z conditioned on both x, w when we remove the arrows into x, in graph G. In DAG G there must be the graphical structures that make this independence true. Is trough Do-Calculus that a causal question is tested to be identifiable. Notice the symbol $\hat{x}$, it means the same as $do(x)$, it is an intervention, we remove the arrows going into x in the graph G.

In this present study, the proposed method to build the SCM is trough domain knowledge, the DAG is built under a hypothesis of how the variables cause each other, like mentioned on the introduction section, subsection 1.2. And then, the causal effect questions, evaluating the SCM, are tested whether or not they are identifiable by applying the do-calculus rules (or the criteria), with this the causal effect tables can be computed and evaluated.

2.2 Data description

The data used in this study was acquired in a research project conducted in our university [16]. Initially, more than 60 subjects, took part in the experiment. All volunteers answered a questionnaire with personal information, like email, full name, gender and age. They also answered the AFECT questionnaire so that there their affective/emotional temperament could be classified. Also, written informed consent was obtained from all participants. This project was subjected to analysis and approved by the ethics committee of the university.

According to Lara [8] [7], four affective temperaments are more common among the general population: stable, unstable, externalizing, and internalizing. So we used these 4 affective temperaments to classify the participants based on the AFECT [8] questionnaire, it can be found on Appendix C - affective section. The AFECT questionnaire was chosen because both their volunteers, whose information validated the model, and the creators of the model were Brazilian, as well as the participants of our study. The following are examples of behavior for each affective temperament:

- **stable**: regularity and moderation help a lot in adapting; may be overconfident and over-control things; moderate search for stimuli or high search for medium intensity stimuli.
- **unstable**: inconstancy in relationships, which causes trouble in the long run;alternate between seeking and avoiding stimuli; they are reactive individuals.
- **externalizing**: they do something first and then think about the consequences later; high search for stimuli; high reactivity.
- **internalizing**: they are inhibited; harm themselves by lack of attitude; avoid stimulus; they are vulnerable.

At first, it was collected EEG data from over 60 participants, as aforementioned, but after signal quality analysis and dropping problematic samples, the remaining database was from 15 volunteers, the amount actually used. The following tools were used to acquire the EEG signals from the participants: Brain Wave II EEG [1] (BWII EEG) using 25 electrodes, the software BWAnalysis, measuring tape, conductive paste and paper towel. A software was developed for the videos' exhibition. The software shows a video until it ends, then, a simple interface with a 9-point scale for valence, another for arousal and a 3-point scale for liking is displayed to the participant. After measuring the volunteer's head, the electrodes were assembled following the 10-20 system [2]. And then some adjustments were made on the software BW Analysis (calibration, impedance tests).

Higher Order Crossings (HOC) [17] features were computed from the EEG signals. Higher Order Crossings are one

[1] http://www.neurovirtual.com.br/equipamento/brain-wave-ii-eeg/
[2] https://en.wikipedia.org/wiki/10%E2%80%9320_system_(EEG)

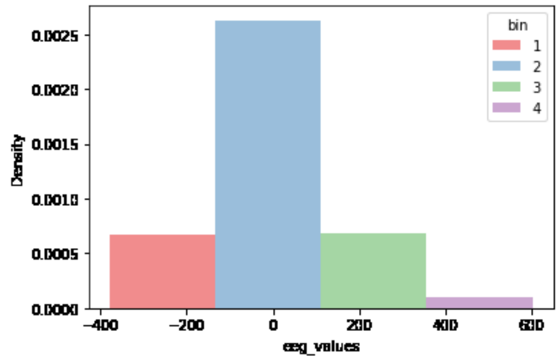

Figure 4. The bins of the histogram are the 4 categories used as the variable called "EEG"

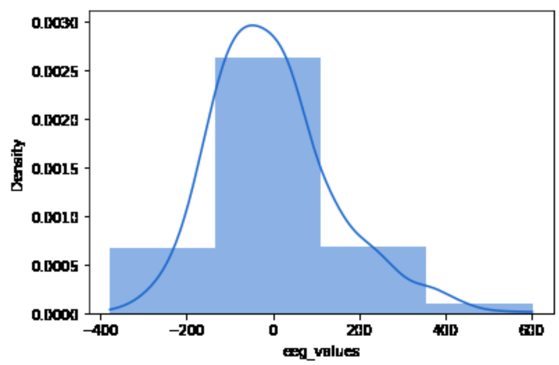

Figure 5. Histogram with KDE for better visualizing the distribution

of the most popular features used for EEG-based emotion classification, which achieves the highest accuracies [18]. Five HOC levels were extracted from the following EEG channels: FP1, FP2, F3, and F4; as proposed by Pereira and Gomes [19], the mean accuracies in experiments using those channels were over 98% using HOC features and SVM classification. Therefore, each movie watched by each subject had 20 HOC features associated with it. Those HOC features were extracted using functions implemented in GNU Octave, then the HOCs of the EEG were used as input to Principal Component Analysis (PCA). The 20 HOC features were reduced to 1 dimension, PCA explained variance ratio is 68,26%, so we lost about 31% of variance explained in exchange of reducing 20 dimensions into 1. Then, the PCA values of the videos were cut into 4 bins with range of values of the same length, it can be seen in Figures 4 and 5.

To induce emotions, video stimuli were presented to the participants, each video had approximately ten seconds of duration. In total, each volunteer watched a noise calibration video and then 60 short videos. Were chosen 15 videos for each quadrant from the arousal-valence plane, the ones with the highest values of valence and arousal for each quadrant in Figure 6. Each participant watched a maximum of 30 minutes, already counting with the intermediate calibration videos that were not used in the analysis. The video stimuli were acquired from LIRIS-ACCEDE [20], which is a large database with many annotated videos, the emotional labels were done trough crowdsourcing by people from different countries with varied cultural and social backgrounds, the authors [20] checked the labeled data for consistency.

Markatopoulou et al. [21] defines the concept annotation process as the task of annotating video fragments with semantic labels referred to as concepts. Examples of concepts are object-related (car, book, pen, etc), general terms (ocean, clouds, forest, etc), and action-related (walking, sleeping, singing). Video Analysis4ALL [3] is an online tool which allows videos to be sent in a variety of formats, and returns a visual analysis of them: scene segmentation and visual concepts detection. This demo was developed based on the following works [21] [22] [23] and [24]. The demo was used to extract the main concepts (or contents) present in each video, it segments the videos into scenes, and returns a list of n concepts ranked from the most probable to the least. Of each video, the first ten scenes (or all of them if the video didn't show ten scenes) were used to get the concepts. From each scene, the 10 first main concepts were taken. The 10 first concepts of the 10 first scenes of a video. Afterwards, k-means clustering was employed to group the videos into 4 clusters according to their similar concepts, since there are 4 quadrants in the arousal-valence plane (Fig. 6).

The valence dimension ranges from unpleasant to pleasant, negative to positive emotions. The arousal dimension represents the intensity of the emotion, or from uninteresting to interesting.

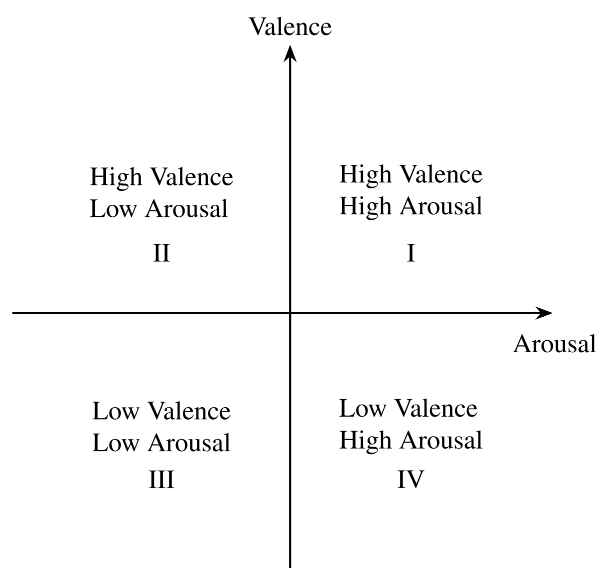

Figure 6. Arousal-Valence plane divided into 4 quadrants.

After watching each video, the participants had to answer a small questionnaire about how they felt during the exhibition. Liking is in a 3-point scale (1: dislike, 2: neutral, 3: like), both Arousal and Valence were in a 9-point scale in the questionnaire (from boring to exciting, unpleasant to pleasant,

[3] Available at http://multimedia2.iti.gr/onlinevideoanalysis/service/start.html

respectively). The answers for these questionnaires were used as ground-truth for the emotions felt by the participants.

The Table 1 contains a small description of each variable in the proposed causal graph from image 8.

Table 1. Variables description

Name	Data Type	Description
Affective Temperament	Integer (1 to 4)	The result of the AFECT questionnaire [7] for each volunteer, the numbers mean the 4 categories mentioned on subsection 2.2.
Gender	Integer (0 or 1)	The gender of the participant, 0 being male and 1 being female
EEG	Integer (1 to 4)	Five HOC features were extracted from 4 EEG channels (FP1, FP2, F3, F4, *i.e.*, 20 levels per video), PCA was computed from the features to reduce them to 1 dimension (explained variance ratio of 68,26%). The PCA values of the videos were discretized by cutting the array in 4 same-length ranges: the bins were named 1, 2, 3 and 4.
Video Concept	Integer (0 to 3)	Each video was attributed a cluster of concepts by the k-means algorithm (4 groups)
AV-Quadrant	Integer (1 to 4)	Arousal-Valence plane quadrant (Fig.6) attributed to each video by each participant right after watching it (the participants informed the emotion felt in a 9-point scale for valence, likewise for arousal)
Liking	Integer (1 to 3)	Each participant informed their liking about each video in a 3-point scale (disliked, neutral, liked)

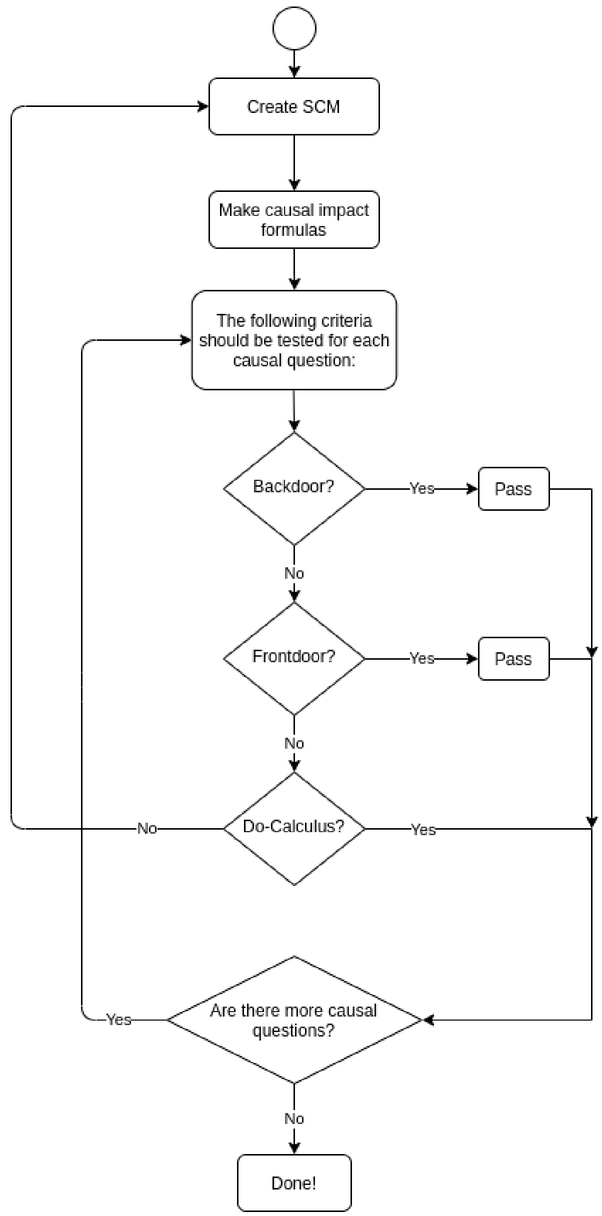

Figure 7. Flowchart with the process used to get the results.

2.3 Proposed Approach

The process followed to achieve the results, described in Figure 7, is going to be explained in this subsection.

2.3.1 Create SCM

SCM means Structured Causal Model, it is a way of describing how the values are assigned to the variables of interest. Formally, it consists of two sets of variables, a graph showing visually how the variables relate to each other, and a set of functions that assign values to each known variable.

In our study, the causal graph was made by using domain knowledge, based on our researches about how EEG signals are affected by gender and age and how they are indicative of the emotions and the liking of a person, see in subsection 1.2.

2.3.2 Make causal questions

The causal questions are represented by causal effect formulas (or equations). The causal effect equation computes the association between two variables. The $do(\cdot)$ operator denotes an **intervention**. The causal questions related to the proposed SCM should be made in the following structure:

$$P(Y = y|do(X = x)) \tag{6}$$

which can be read as *the causal effect of X on Y*. The $do(X = x)$ means that the value of X has been fixed at $X = x$ (the intervention in the value of X), meanwhile, $Y = y$ means that Y took the value y naturally.

2.3.3 The criteria

The *do-calculus* is a powerful tool consisting of three rules that have been proved to be sufficient for deriving all identifiable causal effects [6], but there are simpler tests that cover several scenarios, the **backdoor** criterion and the **front-door** criterion. These criteria were explained in Sub-Section 2.1.

The desired causal effect equation is derived by observing the causal diagram and applying either one of the Criteria or the do-calculus on the joint probabilities of the model's variables.

2.3.4 Test the criteria for all questions

Repeat the process above for all causal questions made, the criteria should be applied to all of them. If *all causal equations* are identifiable, then the SCM is complete.

2.4 Data Analysis

In Table 2 there is information on how all the variables in the proposed causal model are distributed in the dataset. The sum of the counting of the 4 Affective Temperament values results in 15, just like the sum of the counting in Gender, because we used data from 15 participants in total. The counting of the Video Concepts adds up to 60, because Video Concept is the number associated with a group of characteristics present in the video, and there are 60 sample videos.

As for AV-Quadrant and Liking, each person rated valence, arousal, and liking for each video, therefore the sum results in 900 observations (15 $people \times 60\ videos$). For the same reason, there is also 900 observations for the EEG Cluster, for each participant the EEG signals of response to each video were recorded, then the HOC features were extracted, these features were reduced to 1 dimension using PCA and finally, divided into 4 ranges of values with the same length as explained before.

Observing the table, we can see that the sample is quite balanced in terms of gender. We can see that there are many

Table 2. Distributions of all variables of the proposed model

Variable	Variable Value	Abs. Qty.	Rel. Qty.
Affective Temperament	Stable(1)	8	53.3333%
	Unstable(2)	4	26.6667%
	Externalizing(3)	1	6.6667%
	Internalizing(4)	2	13.3333%
Gender	Male(0)	7	46.6667%
	Female(1)	8	53.3333%
AV-Quadrant	1	95	10.5556%
	2	171	19.0000%
	3	490	54.4444%
	4	144	16.0000%
Video Concept	0	23	38.3333%
	1	17	28.3333%
	2	11	18.3333%
	3	9	15.0000%
Liking	Disliked(1)	324	36.0000%
	Neutral(2)	300	33.3333%
	Liked(3)	276	30.6667%
EEG	1	118	13.1111%
	2	598	66.4444%
	3	160	17.7778%
	4	24	2.6667%

more volunteers, around 53%, that fit the stable affective temperament, less than 7% of the sample is of the externalizing temperament. It is eye-catching that most of the videos ended up in the third quadrant, about 54%, and the less marked answer was the first quadrant, around 11%, a difference noticeably big.

Concerning the Video Concept distribution, the majority of the videos were classified in the first cluster (number 0), about 38%, and the minority in the fourth cluster (number 3), about 15%. Not badly distributed. About the distribution of how the volunteers rated the video in the Liking dimensions, one can see that the answers are quite balanced, the percentages of the three options are not so different.

2.5 Proposed hypothetical Causal Model

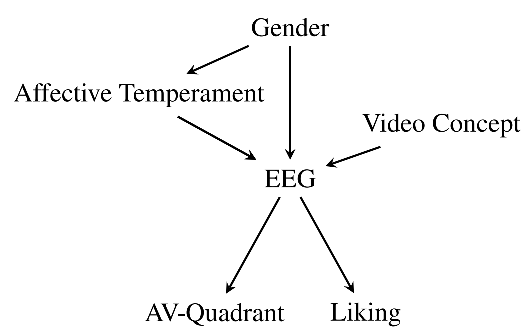

Figure 8. Proposed hypothetical Causal Graph.

The Affective Temperament variable is a number given to each person that represents the affective temperament that best fits them among the four main ones that were used in this research, the mapping is: stable (1), unstable (2), externalizing (3), internalizing (4). The AV-Quadrant ranges from 1 to 4 and it represents the arousal and valence options marked by the participant after watching a video, see Figure 6. An online tool, VideoAnalysis4All, automatically extracted the concepts of each video, the videos were clustered together by a k-means algorithms in 4 groups (since there are 4 quadrants in the arousal-valence plane), each video has a number which is the k-means cluster, the variable Video Concept. Liking is an integer from 1 (disliked) to 3 (liked) that the participant marked in the questionnaire after watching each video. The EEG variable ranges from 1 to 4, is actually the HOCs of the EEG signals reduced with PCA and divided into 4 bins.

3. Results and Discussion

We discuss below the achieved results, the causal effects achieved throughout this study, and the explanation. Each research question is discussed in a specific subsection, and a final discussion about the results is presented in Subsection 3.4. The Figure 8 shows the relationship of all the notable variables, this is the hypothetical model on which causal assumptions in this study were based on.

The Equation 7 is going to be used to calculate the Average Causal Effect (ACE), or simply, the effect of $do(X=x)$. The result is the expected value of Y if the X value was fixed at $X=x$. In the following subsections, we present the results and analysis for our three main causal questions regarding: (i) the causal effect of gender on felt emotion and liking; (ii) the causal effect of affective temperament in felt emotion and liking; and (iii) the causal effect of video concepts on felt emotion and liking.

$$E[Y|do(X=x)] = \sum_y y \cdot P(Y=y|do(X=x)) \tag{7}$$

The causal effect of a variable X on a variable Y is not a simple probability like we are used to in statistics. It actually means that one wants to know what would be the percentage of $Y=y$ if we fixed the value of X of the whole population on $X=x$. All the causal effects shown here were calculated using an implementation of the corresponding causal effect equation using Python3[4], pandas[5] and NumPy[6], as well as the average causal effects.

3.1 How does the Gender affect the felt emotion and the liking?

We have two causal questions, gender affecting emotion and gender affecting liking, but they have similar causal Equations, because the path is similar. The Back-door adjustment cannot be used here because there's no path between Gender and Quadrant with arrows coming into Gender. With $Z=\{EEG\}$ all of the Front-door restrictions were satisfied. So, the causal effect of Gender on Quadrant is given by Equation 8 and the causal effect of Gender on Liking is given by Equation 9:

[4]https://www.python.org
[5]https://pandas.pydata.org
[6]https://numpy.org

3.1.1 Effect of Gender on Quadrant

Table 3. Causal Effect of Gender on Quadrant (Eq. 8)

	Quadrant			
Gender	1	2	3	4
0	0.1276	0.1866	0.5116	0.1742
1	0.1009	0.1956	0.5408	0.1627

$$P(Quadrant|do(Gender)) = \sum_{EEG} P(EEG|Gender) \cdot \sum_{Gender'} P(Quadrant|Gender', EEG) \cdot P(Gender') \tag{8}$$

When calculating the causal impact of Gender on Quadrant, we want to know, e.g., "What's the percentage of videos that would be classified in the Third Quadrant ($Quadrant=3$) if the gender was fixed at the male ($Gender=0$)?". The data indicates that, if gender is fixed at male, the answer would be 0.5116, as in Table 3. The third quadrant, which is low valence/arousal, has much higher values than the other quadrants, for both male and female participants. And the lowest value is for the first quadrant, high valence, and high arousal, also for both males and females.

The average causal effect of both $do(Gender=0)$ and $do(Gender=1)$ on Quadrant is 3, that is, if the gender was male, or gender was female and a random video was selected, the expected quadrant for the video would be the Third Quadrant.

3.1.2 Effect of Gender on Liking

Table 4. Causal Effect of Gender on Liking (Eq. 9)

	Liking		
Gender	1	2	3
0	0.3583	0.3310	0.3107
1	0.3649	0.3336	0.3016

$$P(Liking|do(Gender)) = \sum_{EEG} P(EEG|Gender) \cdot \sum_{Gender'} P(Liking|Gender', EEG) \cdot P(Gender') \tag{9}$$

The impact of Gender on Liking has returned quite balanced results, but we can see the higher results in $Liking=1$, which means that both men and women would have disliked a greater part of the sample videos. And the lower results on that table, for the two genders, are in $Liking=3$, *i.e.*, a smaller part of the videos would have been positively rated by men or by women. Though, overall, the percentages are in balance. The ACE of both $do(Gender = 0)$ and $do(Gender = 1)$ is $Liking = 2$.

The results in Table 4 can be associated with the previous table. The quadrants 3 and 4 have a low valence, they both represent unpleasant emotions, and on Table 3 the third quadrant (low valence/arousal, dull) has the highest values of the whole table for both genders, thus increasing the videos' dislikes ($Liking = 1$) and even neutral liking. The quadrants 1 and 2 (high valence - representing pleasant emotions) have low numbers on Table 3, this way decreasing the values of the liked ($Liking = 3$) videos on Table 4.

3.2 How does the affective temperament affect the felt emotion and liking?

3.2.1 Effect of Temperament on Quadrant

There's an arrow going from Gender to Temperament, with $Z = \{Gender\}$ we block the spurious path $P \leftarrow G \rightarrow E \rightarrow Q$ and the directed path $G \rightarrow P \rightarrow E \rightarrow Q$ is free, the same happens with Liking, thus satisfying all of the Back-door adjustment restrictions. Using $Z = \{EEG\}$ blocks both the spurious and the directed path, and adjusting for $Z = \{Cluster\}$ does not have any effect. We then use the Back-door adjustment with $Z = \{Gender\}$ for both the causal effect of Temperament on Quadrant, in Equation 10, and Temperament on Liking, in Equation 11.

Table 5. Causal Effect of Affective Temperament on Quadrant (Eq. 10)

	Quadrant			
Temperament	1	2	3	4
1	0.1100	0.1800	0.5164	0.1936
2	0.0778	0.3400	0.4578	0.1244
3	0.0444	0.1422	0.3200	0.0267
4	0.0389	0.1856	0.7600	0.0156

$$P(Quadrant|do(Temperament)) = \sum_{Gender} P(Quadrant|Temperament, Gender) \cdot P(Gender) \tag{10}$$

A peculiar characteristic of this dataset is that only one of the volunteers fit the affective temperament 3 (externalizing), so for our two causal questions, when the causal impact of the affective temperament is 3 on every quadrant/liking score it does not sum to 100%, instead, it sums to 53,333% – because that's the probability of a volunteer being of $Gender = 1$ in our dataset – as we do not have any evidence for volunteers of $Gender = 0$ with the externalizing emotional temperament.

"If all volunteers were from the affective temperament t, what would be the percentage of videos rated in Quadrant q?". This time, taking a look in each Temperament separately on Table 5, note that for every affective temperament the third quadrant has the largest percentage. For the affective temperament 1 (stable), the quadrants 1, 2 and 4 have quite similar numbers. Differently from affective temperament 2 (unstable) where the quadrant 1 has much lower values than the rest, quadrants 2 and 3 have the highest results. Now with affective temperament 3 (externalizing), quadrant 4 and 1 have very low percentages. Finally, the internalizing emotional temperament, quadrants 1, 2, and 4 are much lower than the third.

3.2.2 Effect of Temperament on Liking

Table 6. Causal Effect of Affective Temperament on Liking (Eq. 11)

	Liking		
Affective Temperament	1	2	3
1	0.4292	0.3085	0.2623
2	0.2730	0.2704	0.4567
3	0.1244	0.1778	0.2311
4	0.2289	0.5333	0.2378

$$P(Liking|do(Temperament)) = \sum_{Gender} P(Liking|Temperament, Gender) \cdot P(Gender) \tag{11}$$

"If all volunteers were from affective temperament e, then how many videos would be disliked/indifferent/liked by the participants?". Comparing the causal effect of Temperament on Quadrant and Liking, if everyone in our sample had the emotional setting 1, more videos would have been classified in quadrants 3 and 4 – both have low valence (negative emotions) – it is coherent with the larger amount (42.92%) of disliked videos and about 30% of *'indifference'*. If all volunteers had the unstable affective temperament, more videos would have been classified in quadrants 2 and 3, curiously almost 50% of the ratings were positive ($Liking = 3$) even though quadrant 1 has the lowest percentage, the same happens with the externalizing affective temperament, the participants would have classified most of the videos in quadrants 2 and 3, but they would also *like* more videos meanwhile so few videos were classified in quadrant 1, these may be related to personality traits of the affective temperaments, but discussing that is out of the scope of this study. For the internalizing affective temperament, the participants would have classified most of the videos in quadrant 3 and would rate more videos as *indifferent*. The average causal effect of $Temperament = \{1,2,4\}$ on Quadrant is $Quadrant = 3$ and the ACE of $Temperament = 3$ is $Quadrant = 1$. The ACE of $Temperament = \{1,2,4\}$ on Liking is $Liking = 2$, and the ACE of $Temperament = 3$ is $Liking = 1$.

3.3 How does the content of a video affect the felt emotion and liking?

These two causal equations are similar, as well as the others above, because the path in the causal model from Video Concept to Quadrant is similar to the one from Video Concept to Liking. No arrows are going into Video Concept, so we cannot use the Backdoor adjustment in this case. But when using $Z = \{EEG\}$, all the Front-door requirements are satisfied. The causal Equation of the effect of Video Concept on Quadrant is in Equation 12, and of Video Concept on Liking is in Equation 13.

3.3.1 Effect of the Video Concept on Quadrant

Table 7. Causal Effect of Video Concept on Quadrant (Eq. 12)

	Quadrant			
Video Concept	1	2	3	4
0	0.1046	0.1841	0.5566	0.1546
1	0.1035	0.1927	0.5464	0.1574
2	0.1070	0.1835	0.5450	0.1645
3	0.1025	0.1954	0.5497	0.1525

$$P(Quadrant|do(Video\_C)) = \sum_{EEG} P(EEG|Video\_C) \cdot \sum_{Video\_C'} P(Quadrant|Video\_C', EEG) \cdot P(Video\_C') \quad (12)$$

On table 7 we try to answer "If all videos were of Video Concept c, how many of them would be classified in Quadrant q by the participants?". For all Video Concepts, the highest value in each row is at $Quadrant = 3$ with a big gap between the other quadrant values, that means, if all videos had the same video concept more than half of the videos would be classified in Quadrant 3. For any video concept, the average causal effect is $Quadrant = 3$, that is, a randomly selected video is expected to be of quadrant 3.

3.3.2 Effect of the Video Concept on Liking

Table 8. Causal Effect of Video Concept on Liking (Eq. 13)

	Liking		
Video Concept	1	2	3
0	0.3678	0.3299	0.3023
1	0.3599	0.3299	0.3102
2	0.3651	0.3313	0.3036
3	0.3595	0.3258	0.3147

$$P(Liking|do(Video\_C)) = \sum_{EEG} P(EEG|Video\_C) \cdot \sum_{Video\_C'} P(Liking|Video\_C', EEG) \cdot P(Video\_C') \quad (13)$$

By looking at table 8 we try to answer: "If all videos were of Video Concept c, how many of them would be liked, indifferent or disliked by the participants?". For all Video Concepts, the highest value in each row is at $Liking = 1$ (*dislike*) and the lowest in every row is $Liking = 3$, though all values are kind of similar to each other because they all range from 30% to 37%, which is not a very big difference. The ACE of the video concept on liking is $Liking = 2$, or neutral, for $Video\_Custer = \{0, 1, 2, 3\}$, that can be justified by the fact that the ACE of Video Concept on Quadrant is $Quad = 3$, which is low valence/arousal.

3.4 Discussion

Considering the data used and the video stimuli in this study, in general, the effect of Gender, affective Temperament, and Video Concept on Quadrant tend to the third quadrant, and quadrants 1 and 4 always have the lowest values. The effect of the variables is quite balanced within the three Liking options, with a higher tendency to $Liking = 2$.

The LIRIS-ACCEDE videos are not from popular movie studios, it is comprehensible that the elicited emotions were not so intense. Quadrants 2 and 3, which have the highest percentages in all tables, have low arousal , while quadrants 1 and 4 (with the lowest percentages), have high arousal. Also we can see in table 2 that the majority of the videos were rated in quadrant 3, followed by quadrant 2, with a large gap between them.

Considering the causal effect tables shown here, one may take from these results that most of the variables do not causally affect at all neither the elicited emotion nor the liking. Since the causal diagram was made solely based on domain knowledge, some unknown latent variables may have been left of. It is possible to use a technique called "Causal Search" to look up for causal diagrams using statistical dependencies on the input data, there are some causal discovery algorithms available[25], each with a different approach to build the graph. Using Causal Search algorithms to build a causal diagram instead of domain knowledge is an interesting path to be explored to continue with causal analysis on emotional data.

4. Conclusion

In this work, we have presented three causal questions to analyze which factors might cause the emotions induced by the videos on the participants. The proposed questions were if and how much do gender, emotional temperament (affective temperament) and the video's concepts impact the emotion of the person watching and/or the liking of a person about a video. For this study, the EEG signals of 15 participants were used, each participant watched 60 short videos and annotated their emotional response to each video along with the scales of valence, arousal, and their liking.

The Causal Diagram was manually made, based on healthcare and medical studies about affective neuroscience and about how age and gender differences affect EEG signals. The Causal Effect equations were also made manually using the rules of *Do-Calculus* and the Causal Diagram mentioned here. The Causal Effect of an intervention variable X on an interest variable Y, section 3 from Subsection 3.1 to Subsection

3.3, was calculated using an implementation of those causal effect equations using Python3 and auxiliary libraries. It was also calculated the average causal effect of an intervention $X = x$ on Y.

This paper aimed to study the causal relationship between some widely used variables on the emotional computation field. This research can be very valuable and relevant to Computer Science, since that it is one of the first studies to **causally analyze** the impact of video stimulus on emotions, so far. Similar research proposals were found on literature, other researchers have collected body responses (for instance GSR, ECG, EEG, eye-tracking) to a stimulus (videos, images, audios) and then compare them to the participant's ratings, but with purely statistical analysis, accuracy, f1-score and correlation (and *correlation is not causation*).

As mentioned in the Subsection 3.4, a possible study is to investigate the use of causal search algorithms to create the causal diagram instead of doing it manually. Another idea for subsequent studies is to investigate the use of alternative datasets for the creation of new models with, perhaps, different variables, like the effect of a person's age on the arousal/valence quadrant and the liking of a video that was not made clear in this paper; and also to reinforce the causal questions made during this work. At the moment, there are not many publicly available datasets with affective annotations for emotion recognition, the construction of new datasets would allow the creation of new models that could be even more elaborate than the one proposed in this work. Furthermore, the next step in this type of study would be the evaluation of transportability [26] of results between our dataset and some other EEG database for emotional analysis, like DEAP[7], for instance.

Author contributions

Rafaela de Amorim Barbosa Silva implemented software, conducted experiments, wrote the main sections of the paper. **Eanes Torres Pereira** contributed to designing experiments, analyzing results, and in overall text reviewing.

Acknowledgements

The authors would like to thank to ”Programa Institucional de Bolsas de Iniciação Científica (PIBIC) do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) do Brasil.”

This study was partially funded by CNPq (Brazilian National Council for Scientific and Technological Development), Grant N. 459763/2014-8.

References

[1] CISCO. *Cisco Annual Internet Report*. [S.l.], 2020. 1–35 p.

[2] WANG, S.; JI, Q. Video affective content analysis: A survey of state-of-the-art methods. *IEEE Transactions of Affective Computing*, v. 6, n. 4, p. 410–430, 2015.

[3] EKMAN, P. Universal facial expressions of emotions. *California Mental Health Research Digest*, v. 8, n. 4, p. 151–158, 1970.

[4] POSNER, J.; RUSSELL, J. A.; PETERSON, B. S. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, v. 17, n. 3, p. 715–734, 2005.

[5] PEARL, J. Bayesianism and causality, or, why i am only a half-bayesian. In: CORFIELD, D.; WILLIAMSON, J. (Ed.). *Foundations of Bayesianism*. Germany: Springer International Publishing, 2001. p. 19–36.

[6] PEARL, J. *Causality: Models, Reasoning and Inference*. 32 Avenue of the Americas, New York, NY 10013-2473, USA: Cambridge University Press, 2009.

[7] LARA, D. R. et al. The affective and emotional composite temperament (afect) model and scale: A system-based integrative approach. *Journal of Affective Disorders*, v. 140, p. 14–37, 2012.

[8] LARA, D. *Temperamento e Humor: uma Abordagem Integrada da Mente*. Brazil: Observatório Gráfico, 2011.

[9] KAUR, B.; SINGH, D.; ROY, P. P. Eeg based emotion classification mechanism in bci. *Procedia Computer Science*, v. 132, p. 752–758, 2018.

[10] CHAKLADAR, D. D.; CHAKRABORTY, S. Eeg based emotion classification using correlation based subset selection. *Biollogically Inspired Cognitive Architectures*, v. 24, p. 98–106, 2018.

[11] RAZUMNIKOVA, O. M. Gender differences in hemispheric organization during divergent thinking: an eeg investigation in human subjects. *Neuroscience Letters*, v. 362, n. 3, p. 193–195, 2004.

[12] BARRY, R. J. et al. Age and gender effects in eeg coherence: I. developmental trends in normal children. *Clinical Neurophysiology*, v. 115, n. 10, p. 2252–2258, 2004.

[13] ALARCAO, S. M.; FONSECA, M. J. Emotions recognition using eeg signals: A survey. *IEEE TRANSACTIONS ON AFFECTIVE COMPUTING*, v. 10, n. 3, p. 374–393, 2019.

[14] LANZA, S. T.; MOORE, J. E.; BUTERA, N. M. Drawing causal inferences using propensity scores: A practical guide for community psychologists. *American Journal of Community Psychology*, v. 52, n. 0, p. 380–392, 2013.

[15] PEARL, J. *do*-calculus revisited. In: FREITAS, N. de; MURPHY, K. (Ed.). *Proceedings of the Twenty-Eighth Conference on Uncertainty in Artificial Intelligence*. Corvallis, OR: AUAI Press, 2012. p. 4–11.

[7] https://www.eecs.qmul.ac.uk/mmv/datasets/deap/

[16] PEREIRA, E. T. et al. Empirical evidence relating eeg signal duration to emotion classification performance. *IEEE Transactions on Affective Computing*, v. 12, n. 1, p. 154–164, 2021.

[17] PETRANTONAKIS, P. C.; HADJILEONTIADIS, L. J. Emotion recognition from eeg using higher order crossings. *IEEE Transactions on Information Technology in Biomedicine*, v. 14, n. 2, p. 186 – 197, 2010.

[18] JENKE, R.; PEER, A.; BUSS, M. Feature extraction and selection for emotion recognition from eeg. *IEEE Transactions of Affective Computing*, v. 5, n. 3, p. 327 – 339, 2014.

[19] PEREIRA, E. T.; GOMES, H. M. The role of data balancing for emotion classification using eeg signals. In: *2016 IEEE International Conference on Digital Signal Processing (DSP)*. [S.l.: s.n.], 2016. p. 555–559.

[20] BAVEYE, Y. et al. Liris-accede: A video database for affective content analysis. *IEEE Transactions on Affective Computing*, v. 6, n. 1, p. 43–55, 2015.

[21] MARKATOPOULOU, F.; MEZARIS, V.; PATRAS, I. Implicit and explicit concept relations in deep neural networks for multi-label video/image annotation. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 29, n. 6, p. 1631–1644, 2019.

[22] MARKATOPOULOU, F.; MEZARIS, V.; PATRAS, I. Deep multi-task learning with label correlation constraint for video concept detection. In: *Proceedings of the 24th ACM International Conference on Multimedia*. New York, NY, USA: Association for Computing Machinery, 2016. (MM ’16), p. 501–505.

[23] APOSTOLIDIS, E.; MEZARIS, V. Fast shot segmentation combining global and local visual descriptors. In: . [S.l.: s.n.], 2014.

[24] SIDIROPOULOS, P. et al. Temporal video segmentation to scenes using high-level audiovisual features. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 21, p. 1163–1177, 08 2011.

[25] GLYMOUR, C.; ZHANG, K.; SPIRTES, P. Review of causal discovery methods based on graphical models. *Frontiers in Genetics*, v. 10, p. 524, 2019.

[26] BAREINBOIM, E.; PEARL, J. Causal inference and the data-fusion problem. *Proceedings of the National Academy of Sciences*, v. 113, n. 27, p. 7345 – 7352, 2016.