Cite as:
Tzafilkou, K., Economides, A. A., & Protogeros, N. (2022). Mobile sensing for emotion recognition in smartphones: a literature review on non-intrusive methodologies, *International Journal of Human-Computer Interaction*, 38(11), pp.1037–1051. https://doi.org/10.1080/10447318.2021.1979290

Mobile sensing for emotion recognition in Smartphones: A literature review on non-intrusive methodologies

Tzafilkou Katerina*[1], Economides, A. Anastasios[2], Protogeros Nicolaos[3]

[1] Department of Economics, University of Macedonia, Thessalonki, Greece
[2] SMILE lab (smile.uom.gr), Information Systems IPPS, University of Macedonia, Thessaloniki, Greece
[3] Department of Accounting and Finance, University of Macedonia, Thessalonki, Greece

*Corresponding author: Tzafilkou Katerina (katerinatzaf@gmail.com, tzafilkou@uom.edu.gr)

Abstract
This paper aims to provide the reader with a comprehensive background for understanding current knowledge on the use of non-intrusive Mobile Sensing methodologies for emotion recognition in Smartphone devices. We examined the literature on experimental case studies conducted in the domain during the past six years (2015-2020). Search terms identified 95 candidate articles, but inclusion criteria limited the key studies to 30. We analyzed the research objectives (in terms of targeted emotions), the methodology (in terms of input modalities and prediction models) and the findings (in terms of model performance) of these published papers and categorized them accordingly. We used qualitative methods to evaluate and interpret the findings of the collected studies. The results reveal the main research trends and gaps in the field. The study also discusses the research challenges and considers some practical implications for the design of emotion-aware systems within the context of Distance Education.

Keywords: Affective computing; Mobile emotion sensing; Mobile learning emotion recognition; Multimodal signals; Smartphone sensing

1. Introduction

Smartphones have now become ubiquitous instruments not only for communications but also for information retrieval, since they contain a variety of sensors that can determine their location, orientation, meteorological conditions and more. Over recent years, smartphones have also been turned into powerful measurement platforms, since hundreds of 'apps' have been designed to access the sensor-collected information (Odenwald, 2019).

In parallel, the fields of Affective Computing and Machine Learning have introduced a set of tools and methodologies to diagnose the user's affective states from various behavioral and physiological data. Stepping on the association between mobile sensor data and emotional status (Wang et al., 2020), the field of Mobile Sensing brings this 'privilege' to smart mobile devices by allowing them to recognize 'hidden' affective states within users' interaction behaviors. Mobile devices and in particular smartphones are considered suitable platforms for collecting emotional responses in a non-intrusive manner, i.e., without interrupting users (Shu et al., 2019).

To date, most of the research work in Affective Computing has focused on unimodal or multimodal emotion recognition mainly using text, image, and speech data (Poria et al., 2017). Implementing machine learning and deep learning algorithms for supervised learning on visual, textual, and aural information has shown interesting results, while Convolutional Neural Networks (CNNs) are proved highly effective (showing the highest accuracy scores) for visual inputs e.g., in face-tracking research (Hossain and Muhammad, 2019; Levi and Hassner, 2015; Jain et al., 2018; Kwak and Kim, 2018; Ruiz-Garcia et al., 2018). However, the use of un-obtrusive modalities like accelerometer, gyroscope, or touchscreen interaction is quite limited. Among other models, researchers have frequently applied Bayes Networks and Decision Trees to detect stress and relax through accelerometer (unimodal or multimodal) input (e.g., Ciman and Wac, 2016; Maxhuni et al., 2017). Decision trees have also been used for the detection of the basic emotions of happiness, sadness, and anger (Zualkernan et al., 2017). Deep learning methodologies have been mainly used in the recognition of the basic emotions achieving quite high scores of accuracies like for instance in Alshamis et al. (2017) and in Hussian and Muhammand (2017) where the authors used a unimodal approach of camera-based input. According to recent findings of Kanjo et al. (2019), the adoption of deep learning approaches in human emotion classification is more effective than traditional machine learning when large number of sensors input is utilized.

Unfortunately, most of the personalized e-learning systems neglect emotions and the ones that detect them tend to use obtrusive methods (Wahid and Rasheed, 2019). Research on students and teachers has shown strong correlations between learning and emotion (Kort et al., 2001) which is equally reflected in mobile learning environments (Benta et al., 2005; 2015). In the context of online learning and distance education, it is found that the case of the covid-19 pandemic crisis has affected some basic learning related emotions like boredom and anxiety and caused several mood changes that might affect the learning performance (Irawan et al., 2020). The transition to Emergency Distance Education also has a strong impact on students' emotions and psychological states (Besser et al., 2020). For this reason, predicting or diagnosing learning-related emotions in online learning activities is useful to provide with affective-based recommendation, adaptation, or personalization services. In one recent example, Ashwin and Guddeti (2020) designed a teaching framework to inquiry intervention using students' affective states. In their framework, the authors use facial expressions, hand gestures and body postures to predict the unobtrusive and multi-modal students' affective states. Similarly, Ez-zaouia et al., (2020) designed an emotion aware system to allow teachers to provide useful feedback when strong emotions are detected during the online learning process. Unfortunately, their work might be difficult or impractical for mobile learning (m-learning) frameworks.

Since learning-related emotions (LREs) play a crucial role in learners' engagement, overall behavior and learning achievement, m-Learning can significantly benefit from mobile sensing. The emerging field of Mobile Sensing promises new intelligent achievements in emotion recognition since modern mobile devices are equipped with a range of powerful embedded sensors, such as gyroscopes, GPSs, front-face cameras, accelerometers, and magnetometers, offering the opportunity to develop efficient multimodal mobile sensing methodologies. Recent mobile sensing research includes environmental modalities or location data (Bogomolov et al., 2014; Kanjo et al., 2018; 2019) as well. Smart mobile devices bring intrinsic features such as portability, connectivity, built-in sensors, and context-sensitive characteristics, facilitating multimodal affective analysis and rendering m-learning one of the most challenging technological innovations in education.

Mobile learning researchers (e.g., Lee et al., 2012; Shu et al., 2019) agree on the importance of unobtrusively recognize learner emotions to provide with better learning experiences and enhanced acceptance of the mobile learning applications.

Although mobile sensing-based technologies and machine learning models have been broadly studied across a wide variety of domains like transportation, healthcare, safety, and more (Khan et al., 2013; Laport-López et al., 2020; Zhang et al., 2016), emotion detection in app interaction and learning (m-learning) remains unexplored to some extent. To reveal this gap and assist researchers in the field, this study reviews the latest mobile sensing methodologies for emotion recognition, focusing on non-intrusive sensing methodologies. Specifically, the present study poses the following three research questions:

RQ1: What types of non-intrusive mobile modalities have been used to detect affective states including learning related emotions?

RQ2: Which prediction/classification models provide the highest accuracy scores in non-intrusive emotion detection tasks?

RQ3: What are the main emotions, including learning related emotions that have been recognized by mobile sensing algorithms?

Throughout the manuscript we answer on each one of the research objectives proving with separate information in tables and figures. In the end of the study, we discuss the main findings and the practical implementations, following a combined approach to interpret ate the results.

1.1. Useful terms and definitions

- Mobile Sensing refers to the “use of sensors of a mobile device to acquire data from the environment” (Liu and Koc, 2015).
- Non-intrusive data collection refers to methods of obtaining data without engaging or disturbing the activity of the people/users (Emden and Smith, 2004). In this study the term refers to the usage of embedded sensors and devices in smartphones that allow the collection of data in a non-intrusive manner during user-system interaction.
- Unimodal refers to the methodologies that used one single mobile modality of either a built-in sensor or internal device.
- Multimodal refers to the methodologies that used several mobile modalities combining built in sensors and/or internal devices.

1.2. Related work and contribution

Mobile Sensing is a popular technology that has attracted research in recent years. According to Gartner’s Hype Cycle of “Emotion Detection/Recognition” (Gartner, 2018), the productivity plateau is expected to be reached in about two to five years. The interest towards research in emotion detection on mobile devices is continuously growing; however, the integration of non-intrusive detection methodologies for emotion aware mobile interfaces is still quite an unexplored field of study.

Systematic reviews of research and applications have found that Mobile Sensing has an efficient potential for emotion detection. However, the studies also point out the difficulties of measuring mobile emotion due to the lack of standards, the limited processing power of the mobile devices and the numerous possibilities of the measurement (Kołakowska et al., 2020; Meyer et al., 2019; Geven et al., 2009). Most of these reviews are usually focused on the use of mobile sensing within a broad range of data collection methods (e.g., Kołakowska et al., 2020), examining both obtrusive (e.g., using wearables) and unobtrusive methods (e.g., Meyer et al., 2019; Geven et al., 2009). Other reviews (e.g., in Laport-López et al., 2020) concern all the mobile sensing methodologies that were used in several research fields, and not particularly in the emotion detection research field. Other review works (e.g. Pham and Wang, 2016; 2018; Radhamani and Dallin, 2018) focus solely on camera-based emotion recognition since it is a rich and well-promising non-intrusive modality.

Non-intrusive methods are far different from intrusive methodologies since the latter has access to much richer feedback (including biometrics) using wearables or external devices. Collecting data from mobile embedded sensor is also considered a highly non-obtrusive way of detecting emotions, since it requires no extra equipment or interaction with the user (Olsen and Torresen, 2016). To this end, non-intrusive mobile detection should be considered a standalone category due to the main limitations it faces to collect rich data feedback as well as to the unique advantages that it offers to the user experience (by not destructing the user from the natural flow of his/her activity). Hence, this review presents the non-intrusive methodologies for mobile emotion recognition, trying focus on the educational context as well (but without excluding the rest). By providing answers to the afore-mentioned research questions, the research community can have a clear vision on the progress of mobile sensing efforts and the next steps towards the design of effective LRE recognition models in m-learning environments.

The contribution of this work is essential and might be useful in the context of Distance Education (DE) and Emergency Remote Education (ERE) that is fully adopted worldwide due to the covid-19 pandemic crisis. User engagement is one of the greatest challenges in DE and researchers claim that emotion awareness is the key to designing online courses engaging for the students (Ez-zaouia et al., 2020; Irawan et al., 2020).

The ability to effectively detect student emotions in mobile-based learning can bring useful practical implications in the field of emotion-aware m-learning systems in synchronous and asynchronous education.

To sum up, the main contributions of this review are as follows:

- To provide a complete documentation of the applied research so far on non-intrusive mobile sensing approaches for emotion detection tasks, highlighting the progress in the educational community as well.
- To overview the classification and prediction models that lead to the highest accuracy scores, given the limited power capacity of the mobile devices.
- To compare the key aspects between unimodal and multimodal non-intrusive sensing methodologies in terms of targeted emotions and input modalities. Also, researchers are

provided with an overview of the most frequently detected emotions (including LREs) and the various detection methodologies traced over the recent years.

2. Basic emotions and learning related emotions (LREs)

Emotions have been categorized by Ekman (2007) into six basic categories including 'anger', 'disgust', 'fear', 'happiness', 'sadness' and 'surprise'. Ekman's model which is partially based on Darwin's (1965) ideas of emotions describes the attachment of characteristics to each of these emotions, allowing them to be expressed in varying degrees and as a discrete category. Ekman (2002) also argues that "these basic emotions differ not only in expression but probably in other important aspects, such as appraisal, antecedent events, probable behavioral response, physiology, etc.". A more recent affective representation, the Hourglass of Emotions (Cambria et al., 2012), represents affective states through a set of labels and affective dimensions, aiming to describe the full range of emotional experiences. Another popular emotion model is the Pleasure-Arousal-Dominance (PAD) model (Mehrabian and Russell, 1974). The PAD model describes emotions through three dimensions: pleasure, arousal, and dominance. Pleasure (or valence) determines between positive and negative emotions with arousal differentiates between active and passive states. The dimension of dominance reveals the controlling or dominant nature of an emotion. Each emotion can be represented as a linear combination of the three PAD dimensions while the zero-point values represent a neutral emotional state (Kołakowska et al., 2020).

Simplified versions of Ekman's and PAD emotion models have been universally applied in educational research since emotions are considered crucial determinants to the learning processes and have an impact on academic achievement (Hulleman et al., 2010; Goetz and Hall, 2013; Shun, et al., 2015). Learning-Related emotions (LREs) also called 'achievement emotions' have been emerged in the educational research context mainly after the development of the Control-Value Theory (CVT) (Perkun et al., 2007) that categorized them according to valence (positive, negative) and the nature of response (activating or deactivating). While positive LREs (e.g., enjoyment) have shown to bring positive outcomes to achievement (Dettmers et al., 2011), unpleasant emotions like anxiety or boredom can shape disengagement and hence lower learning performance (Dettmers et al., 2011). Perceived difficulty is also conspired a learning related affect since it can significantly affect engagement and performance (Afergan et al, 2014; Lynch et al, 2013; Pham and Wang, 2016). Several research findings (e.g. Gilleade et al., 2005; Kapoor et al., 2007; Van Lankveld et al., 2010) show that 'difficult' tasks that lead to repeated failure tend to result in negative emotions. In contrast, when students receive materials or challenges at just the right level of difficulty, they tend to develop positive emotions (D'Mello and Graesser, 2011; Van Lankveld et al., 2010).

Fear and embarrassment have also been examined in the context of negative LREs, where the authors proved their negative effects on student motivation and performance (Pekrun and Stephens, 2010; Rowe and Fitness, 2018). Overall, the emotions of enjoyment, frustration, anxiety, boredom, and confusion are outlined as the most frequently experienced emotions during technology-based learning (e.g., Calvo & D'Mello, 2011; D'Mello & Graesser, 2012; Schrader & Nett, 2018).

3. Review methodology

To carry out this research we adopted parts of the widely used guidelines on conducting systematic literature reviews presented by Kitchenham and Charters (2007), Nightingale (2009), and Okoli and Schabram (2010). The review guidelines mainly include three main steps: i) search strategy; ii) study selection and quality assessment criteria; and iii) data extraction and synthesis.

3.1. *Search strategy*

A literature review of studies mainly indexed in ACM, Google Scholar and Google semantics, IEEE, Web of Science, Elsevier, Science Direct and Springer Link was conducted. The keywords were *mobile emotion, Smartphone emotion, mobile sensing, mobile affective computing, mobile affective states, detecting learning emotions, non-intrusive affect recognition and unobtrusive affect recognition*. The keyword '*mobile emotion'* was the one that retrieved the maximum value of related results.

The search of literature on mobile sensing for emotion recognition was established within a period of the last six years (2015-2020). This is mainly because this review aims to provide us with the most recent trends and advances in the field. Furthermore, only a limited number of studies mainly concerning face-tracking methodologies are found in the years before 2015. Most of the reviewed studies are found within the year 2016; while all 2020 upcoming publications have not appeared yet.

Also, most of the reviewed studies were detected in the IEEE (16 studies) and ACM (7 studies) databases, whereas 3 studies were detected in Springer Link, 2 in Elsevier and 2 in Google Semantic/Scholar. Most of the studies that were conducted in the learning context were indexed in ACM (4 out of 5) and only one in IEEE.

3.2. Study selection and quality assessment criteria

A total of ninety-five candidate articles were initially reviewed, however, only thirty of them met the study requirements. The retrieved studies were analyzed to be included (or excluded) following the criteria described in Table 1.

Table 1 Inclusion and exclusion criteria

Inclusion	Exclusion
Studies conducted only for/in mobile smartphone devices.	Studies conducted for/in desktop or tablet environments.
Studies published between 2015 and 2020.	Studies published earlier than 2015.
The data extraction process includes solely non-intrusive methods.	Studies including only self-assessment feedback (and not as the model's output data), wearables, or external devices to collect physiological signal.
The detection mechanism applies on emotional states.	Studies focused on mental health, social behavior, etc.
Studies that present empirical data	Studies that do not present empirical data (e.g. theoretical and conceptual articles, tool demonstration, etc.)
Studies that used a machine learning methodology to detect emotions	Studies that used qualitative or correlation-based analyses to explain links in behavior and emotions

3.3. *Data extraction and synthesis*

In this step of the review methodology, each one of the selected key studies selected was fully read to extract the following attributes: title, author, source, study focus, research context (learning or other), modality-based methodology (unimodal/multimodal), targeted emotions, used modalities and prediction models.

Furthermore, we categorized the reviewed mobile data extraction sources/modalities into built-in (or embedded) mobile sensors and internal devices (or apps/services) as depicted in Table 2. A detailed taxonomy of mobile sensing systems used in all research fields can be found in Laport-López et al. (2020). Table 3 describes the most common built-in mobile sensors used in non-intrusive methodologies for emotion detection tasks. In this study, front camera is included in the list of non-intrusive modalities since it is regarded as a built-in optical sensor (Odenwald, 2019) which does not require any external device or equipment. Recent advances in online face tracing tools (FaceReader, 2021) are quite flexible in head position and manage to sufficiently track the basic face expressions of the users without distracting them.

Table 2 Examples of reviewed built in (non-intrusive) sensors and internal devices

Built-in mobile sensors	**Mobile internal devices (or services)**
accelerometer, gravity sensor, gyroscope, light sensor, GPS (Global Positioning System) Sensor	front camera, microphone, touch screen, pedometer app, network/Wi-Fi, keyboard, app log, SMS, PPG (implicit photoplethysmography sensing)

Table 3 Most commonly used built-in sensors in non-intrusive mobile emotion detection studies

Reviewed built-in sensor	**Sensor type**	**Description**
Accelerometer sensor	Motion sensor	Used as user interface controller measuring the acceleration of smart phone in three different axes: X, Y and Z (Android developers, 2020; Ali et al., 2014).
Gyroscope sensor	Motion sensor	Measures the rate of rotation around the three axes: X Y and Z (Android developers, 2020). Gyroscope when combined with accelerometer allow to measure motion along six axes: left, right, up, down, forward, and backward, providing more accurate motion measuring capabilities (Ali et al., 2014).
Light sensor	Environment sensor	Measures the luminance of the surrounding environments and adjusts brightness of mobile phone accordingly to optimize screen visibility (Android developers, 2020; Ali et al., 2014).
GPS (Global Positioning System) sensor	Position sensor	A navigation tracking system received information from the GPS satellites and calculates a user's exact location using triangulation (Ali et al., 2014).

Gravity sensor	Motion sensor	Usually a virtual sensor or app (Ali et al., 2014) measuring the force of gravity along the x, y and z axes (Android developers, 2020).

4. Results and discussion

4.1. Overview and classification of the reviewed studies

Of the 30 reviewed studies, the majority (16 articles) were published in 2016 and 2017. The number of participants in the studies ranged from 3 (Shapsough et al., 2016) to 127 (Mottelson and Hornbsk, 2016). Most of the studies were carried out with less than 55 participants, while most of the camera-inclusive studies (e.g., Alshamsi et al., 2017; Shu et al., 2019; Suchitra et al., 2016) used existing image datasets, of much larger volumes of records (thousands of images). Also, most of the reviewed studies (17 articles) followed a unimodal sensing methodology including the modalities of camera, keyboard, accelerometer, and touch screen, while most of the LREs-focused studies (4 articles) that were conducted in educational contexts measured signals from photoplethysmography (PPG) sensor to implicitly gather heart rate features from students. An overview of the research studies is provided in Appendix A sorted by modality category and date of publication.

In the following Tables 4, 5 and 6 we classify the reviewed articles according to the main classification criteria (input modalities, prediction models, targeted emotions,). Then, in the next sections we synthesize and discuss the main findings of every classification case, providing answers to the stated research questions.

Based on the review of the input modalities used in the selected studies, most non-intrusive mobile sensing studies utilized the accelerometer sensor and the embedded front camera

Table 4. Classification of case studies according to the input modalities

Modality	Authors and Year (Paper Ref)
accelerometer	Barron-Estrada et al., 2018; Ceja et al., 2015; Ciman and Wac, 2016; Dong et al., 2019; Gjoreski et al., 2015; Maxhuni et al., 2016;2017; Mottelson and Hornbsk, 2016; Olsen and Torresen, 2016; Shapsough et al., 2016; Wang et al., 2020; Zualkernan et al., 2017
camera	Alshamsi et al., 2017; Jeong and Lynn, 2016; Hussian and Muhammand, 2017; Pham and Wang, 2018; Shu et al, 2019; Siddiqi et al., 2017; Suchitra et al., 2016; Suk and Prabhakara, 2015
keyboard	Ghosh et al., 2019a; b;c; Kanjo et al., 2017; Shapsough et al., 2016; Zualkernan et al., 2017
touch screen	Bhattacharya, 2017; Ciman and Wac, 2016; Mottelson and Hornbsk, 2016; Pham and Wang, 2018; Shah et al., 2015
gyroscope	Barron-Estrada et al., 2018; Ciman and Wac, 2016; Dong et al., 2019; Mottelson and Hornbsk, 2016; Wang et al., 2020
GPS	Ciman and Wac, 2016; Dong et al., 2019; Gjoreski et al., 2015; Wang et al., 2020
network	Maxhuni et al., 2016; Ciman and Wac, 2016; Gjoreski et al., 2015
light sensor	Dong et al., 2019; Ciman and Wac, 2016; Gjoreski et al., 2015; Wang et al., 2020

PPG	Pham and Wang, 2016;2018; Xiao and Wang, 2015;2016
microphone	Barron-Estrada et al., 2018; Gjoreski et al., 2015; Maxhuni et al., 2016
gravity	Dong et al., 2019; Wang et al., 2020
pedometer	Dong et al., 2019; Wang et al., 2020
call log/app log	Gjoreski et al., 2015; Maxhuni et al., 2016

The preferred techniques of researchers' emotion recognition are provided in Table 5. This study identified the following prediction models that were mainly used in the selected reviewed studies: Linear Regression (LR), Bagging/Boosting (BB), K-nearest neighbor (KNN), Neural Networks (NN), Random Forests (RF), Decision Trees (DT), Logistic Regression (LogR), Bayes Network /Naive Bayes (BN), Support Vector Machine (SVM), Fuzzy Classification (FC), Multiple Classifier/Fusion (MC), and k-Means clustering (k-means). As depicted, DT, SVM and NN are the most used models in emotion detection tasks.

Table 5. Classification of case studies according to the recognition/prediction model

Recognition/Prediction Model	**Authors and Year (Paper Ref)**
Linear Regression (LR)	Bhattacharya, 2017; Shah et al., 2015
Bagging/Boosting (BB)	Maxhuni et al., 2017; Shah et al., 2015
K-nearest neighbor (KNN)	Alshamsi et al., 2017; Ciman and Wang, 2016; Xiao and Wang, 2015
Neural Networks (NN)	Ceja et al., 2019; Ciman and Wang, 2016; Hussian and Muhammand, 2017; Kanjo et al., 2017; Olsen, 2016; Shu et al., 2019;
Random Forests (RF)	Ghosh et al., 2019b; c; Maxhuni et al., 2017
Decision Trees (DT)	Ceja et al., 2015; Ciman and Wang, 2016; Maxhuni et al., 2016;2017; Shapsough et al., 2016; Xiao and Wang, 2015; Zualkernan et al., 2017
Logistic Regression (LogR)	Kanjo et al., 2017
Bayes Network /Naive Bayes (BN)	Ceja et al., 2015; Ciman and Wang, 2016; Xiao and Wang, 2015
Support Vector Machine (SVM)	Ciman and Wac, 2016; Kanjo et al., 2017; Maxhuni et al., 2017; Mottelson and Hornbsk, 2016; Olsen, 2016; Pham and Wang, 2016; Shu et al, 2019; Suk and Prabhakara, 2015; Xiao and Wang, 2015
Fuzzy Classification (FC)	Barron-Estrada et al., 2018
Multiple Classifier/Fusion (MC)	Dong et al., 2019; Wang et al., 2020
k-Means clustering (k-means)	Ghosh et al., 2019a; Shah et al., 2015

Table 6 illustrates the targeted emotions that were defined in the research objectives of the key studies. Most of the studies targeted more than two emotions, while studies targeting one single emotion usually regarded the emotion of stress. Details of every case study are cited in Appendix A, Table A1.

Table 6. Classification of case studies according to the targeted emotions

Emotion	Authors and Year (Paper Ref)
anger	Alshamsi et al., 2017; Dong et al., 2019; Hussian and Muhammand, 2017; Jeong and Lynn, 2016; Pham and Wang, 2018; Siddiqi et al., 2017; Shah et al., 2015; Shapsough et al., 2016; Suchitra et al., 2016; Suk and Prabhakara, 2015; Wang et al., 2020; Zualkernan et al., 2017
sadness	Alshamsi et al., 2017; Ghosh et al., 2019a; b;c; Jeong and Lynn, 2016; Hussian and Muhammand, 2017; Pham and Wang, 2018; Shah et al., 2015; Siddiqi et al., 2017; Suk and Prabhakara, 2015; Zualkernan et al., 2017;
happiness	Alshamsi et al., 2017; Barron-Estrada et al., Ghosh et al., 2019a; b;c; Hussian and Muhammand, 2017; Siddiqi et al., 2017; Shu et al, 2019; Suchitra et al., 2016; Suk and Prabhakara, 2015; Shapsough et al., 2016; Zualkernan et al., 2017
stress	Ceja et al., 2015; Ciman and Wac, 2016; Dong et al., 2019; Gjoreski et al., 2015; Ghosh et al., 2019a; b;c; Maxhuni et al., 2016; 2017; Wang et al., 2020
relax	Ciman and Wac, 2016; Dong et al., 2019; Ghosh et al., 2019a; b;c΄Wang et al., 2020
disgust	Alshamsi et al., 2017; Pham and Wang, 2018; Shah et al., 2015; Suk and Prabhakara Jeong and Lynn, 2016; Hussian and Muhammand, 2017; Suchitra et al., 2016
surprise	Alshamsi et al., 2017; Suk and Prabhakara Jeong and Lynn, 2016; Hussian and Muhammand, 2017; Pham and Wang, 2018; Siddiqi et al., 2017; Suchitra et al., 2016; Suk and Prabhakara, 2015
fear	Alshamsi et al., 2017; Jeong and Lynn, 2016; Pham and Wang, 2018; Shah et al., 2015; Suk and Prabhakara, 2015
engagement	Barron-Estrada et al., 2018; Jeong and Lynn, 2016; Pham and Wang, 2018
boredom/disengagement	Barron-Estrada et al., 2018; Xiao and Wang, 2015; 2016
contempt	Jeong and Lynn, 2016; Pham and Wang, 2018
pleasantness/arousal/joy	Jeong and Lynn, 2016; Olsen and Torresen, 2016; Pham and Wang, 2018
attention	Pham and Wang, 2018
confusion/difficulty	Pham and Wang, 2016; Xiao and Wang, 2015
positive, negative, neutral	Bhattacharya, 2017; Kanjo et al., 2017; Mottelson and Hornbsk, 2016

4.2. Key studies analysis

About modalities

In this section we present the findings of the review process and answer on the initially set research question RQ1.

> RQ1: *What types of non-intrusive mobile modalities have been used to detect affective states including learning related emotions?*

The commonly used modalities for emotion detection through mobile sensing belong in the category of 'personal sensing' that Laport-López et al. (2020) and Khane et al. (2013) defined as the type of sensing focused on personal and individual monitoring mainly to collect information about users' everyday life and activities. Furthermore, the built-in sensors used in the reviewed studies belong in the categories of motion, location, and environment, as defined by Android developers' taxonomy (2020).

In Table 7 we provide an overview of the mobile built-in sensors used for emotion detection in the cited studies. This review study found that the most frequently used built-in sensors in non-intrusive mobile sensing tasks is the accelerator, while the most used embedded device is the camera. PPG seems to be the most popular means of data collection in m-learning environments, targeting the detection of LREs like boredom, engagement, and attention. Table 7 illustrates the frequency of every modality detected in the reviewed studies and depicts whether a modality is used alone (unimodal only), both alone and in combination with other modalities (unimodal and multimodal) or only in combination with other modalities (multimodal only). As depicted, accelerometer, camera, keyboard, and touch screen are the modalities that have been used as single data inputs in unimodal methodologies, but also in combination with other modalities. Th modalities of GPS, network, light sensor, PPG, microphone, gravity, pedometer, and call/app/log have been used only in multimodal methodologies, possibly implying their difficulty/weakness to compose a single source of data in emotion detection models. Details on the (multi)modality combination and study references are presented in Table A1 in Appendix A.

Table 7 Mobile emotion sensing modalities for non-intrusive data collection used in the reviewed studies

Modality	**Num of studies**	**Modality methodology**	**Recognized emotions**
accelerometer	12	unimodal and multimodal	pleasantness, arousal, anger, stress, relax, depression, engagement
camera	8	unimodal and multimodal	happiness, anger, sadness, disgust, surprise
keyboard	6	unimodal and multimodal	happiness, sadness, stress, relax
touch screen	5	unimodal and multimodal	positive, negative, neutral, sadness, anger, fear, disgust
gyroscope	5	multimodal only	anger, stress, relax, depression, boredom
GPS	5	multimodal only	anger, stress, relax, depression
network	3	multimodal only	anger, stress, relax, depression
light sensor	4	multimodal only	anger, stress, relax, depression
PPG	4	multimodal only	anger, fear, sadness, surprise, joy, disgust, contempt, engagement, attention, boredom, confusion, difficulty, disengagement
microphone	3	multimodal only	engagement, stress
gravity	2	multimodal only	anger, stress, relax, depression
pedometer	2	multimodal only	anger, stress, relax, depression
call log/app log	2	multimodal only	stress

About detection/prediction models and accuracy scores

In this section we provide the findings of the review analysis we conducted to reply to RQ2.

> RQ2: *Which prediction/classification models provide the highest accuracy scores in non-intrusive emotion detection tasks?*

Figure 1 depicts the overall prediction models and accuracy scores of the reviewed studies. In cased where several models were applied, we demonstrate the ones achieving the highest scores (e.g., Ciman and Wac, 2016, Maxhuni et al., 2017 and Xiao and Wamg, 2015). As depicted, the Neural Network based GMM classifier to investigate Gaussian mixtures has brought the highest accuracy score providing above 99% of accuracy. This finding is in accordance with Radhamani and Dallin (2018) who made a comparative review on all emotion detection models (both in mobile and non-mobile devices) that used the modality of camera to retrieve their input.

The most recent emotion detection attempts (2019-2020) seem to prefer NN, SVM and MC models but we do not see any significant progress towards achieving high accuracy scores. What is interesting is that NN achieves a high score of 85 % in the work of Ghosh et al (2019a), but only a 60% accuracy score in the work of Shu et al. (2019). To explain such differentiations in accuracy scores from similar classification methods we need to deeper examine the characteristics of each study and the architectures of each deployed model. For instance, Ghosh et al (2019a) deployed a multi-task learning (MTL) Neural Network to process text-input from keyboard, while Shu et al. (2019) architecture a convolution neural network (CNN) to process image-data from camera. Also, Ghosh et al (2019a) conducted their research on a small sample of 22 students/participants, while Shu et al. (2019) retrieved their data from the CK dataset (Lucey et al., 2010) containing 27,000 images. Drawing from the afore described, we cannot conclude in accurate performance results from the comparison of similar detection/recognition models due to several potential differentiations that might lie in input data and the preprocessing data processes, or even the sample size and demographic characteristics of the participants.

The multimodal work of Gjoreski et al. (2015) is excluded from the diagram since the detection accuracy score (43%) was much below the other studies and can be considered as an outlier. The authors used a Decision Tree (DT) C4.5 (J48) model to classify stress levels in a population of 48 participants.

Furthermore, as depicted in Figure 1, there is not any stable increase on accuracy scores for non-intrusive methodologies over the years. Most of the recent non-intrusive methodologies have brought an average score of around 75%, while the highest performance has been achieved by methodologies developed during 2016 and 2017. This conclusion is in contrast with the remark of Mayer et al. (2019) who have observed that over time, both the speed as well as the accuracy of the measurement has improved considerably. As stated, the review study of Mayer et al. (2019) includes the non-intrusive methodologies as well, and this might be the main cause of this different view. Although deep-learning-based emotion-recognition generally produces better accuracy scores, it also requires a large amount of processing capacity, such as a graphic processing unit (GPU) and central processing unit (CPU) which is limited in mobile device (Gu et al., 2017; Ko et al., 2018).

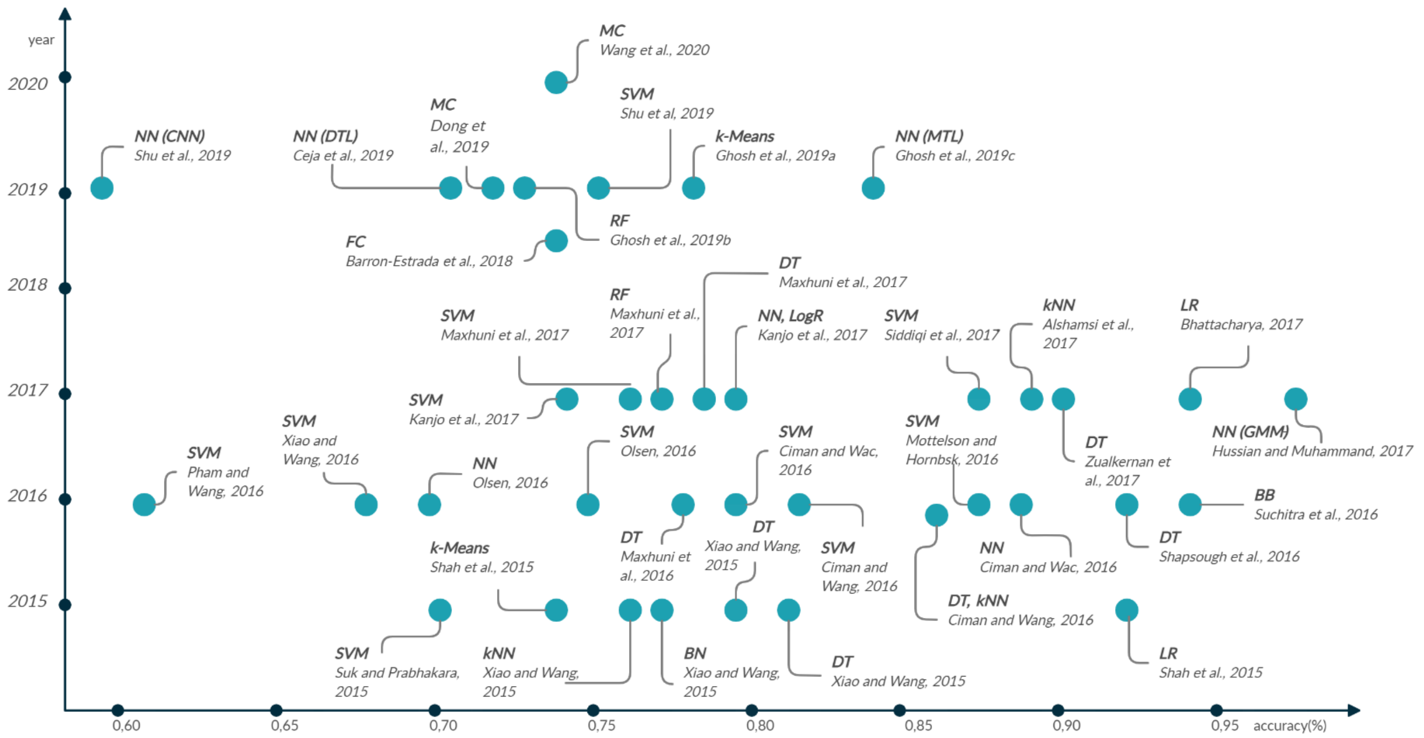

Figure 1 Classification of case studies according to the prediction models and accuracy scores over the years (2015-2020)

Following, Figure 2 depicts all the reviewed recognition/prediction models, the targeted emotions and the achieved accuracy scores that were implemented through the two most popular mobile modalities: camera and accelerometer. As depicted, all six camera-based research works were unimodal using no other source of data to detect user emotions. Only one study (Pham and Wang, 2018) used a multimodal camera inclusive approach, but the accuracy score (achieved by commercial software) is not provided in their article.

It was a camera-based study (Hussian and Muhammand, 2017) that achieved the highest score of 99% at classifying the basic emotions of anger, happiness, sadness, disgust, surprise, and neutral state. In the more recent attempts (Shu et al., 2019) camera face recognition was used to detect happiness, via a convolutional neural network (CNN) and the SVM model approach.

Accelerometer-based studies followed both a unimodal and a multimodal approach. The reviewed multimodal accelerometer-inclusive approaches brought in the majority higher performance results, mainly elaborating DT methods, BNs, NNs, kNNs and SVM models. This finding is in accordance with D'Mello (2015) who supported multimodal systems are consistently more accurate than unmoral ones. Also, stress was the main emotion being targeted by most accelerometer-based studies.

The study of Barron-Estrada (2018) was the only one that used the accelerometer sensor (in a multimodal approach) to detect LRES of engagement and boredom. The author worked on a Fuzzy Classification (FC) model to achieve an accuracy score of 73%.

What is also interesting is that we do not meet any LREs detected via camera-based input, revealing the research gap in the educational context. However, camera-based input has been used in several computer-based tasks to predict students' emotions, for instance Terzis et al. (2013) achieved 87% accuracy for predicting students' emotions of anger, sadness, and neutral state during a computer-based assessment (CBA) task, solely using camera input. In Xiao and Wang (2015) the authors applied

an DT approach and the PPG modality to detect the emotions of boredom and confusion. The authors used PPG signals to detect detected confusion and boredom levels in watching video activities. DT achieved the highest accuracy scores, where the accuracy of concussion detection was 82%, and 79% of boredom. The second highest score was achieved through DT for confusion (77.3%), and through BN for boredom (78%), as depicted in Figure 1.
Since today students tend to prefer mobile environments mainly because of their attractiveness and novelty (Nikou and Economides, 2019), this review suggests the integration of camera inclusive non-intrusive emotion detection in mobile learning or CBA environments.

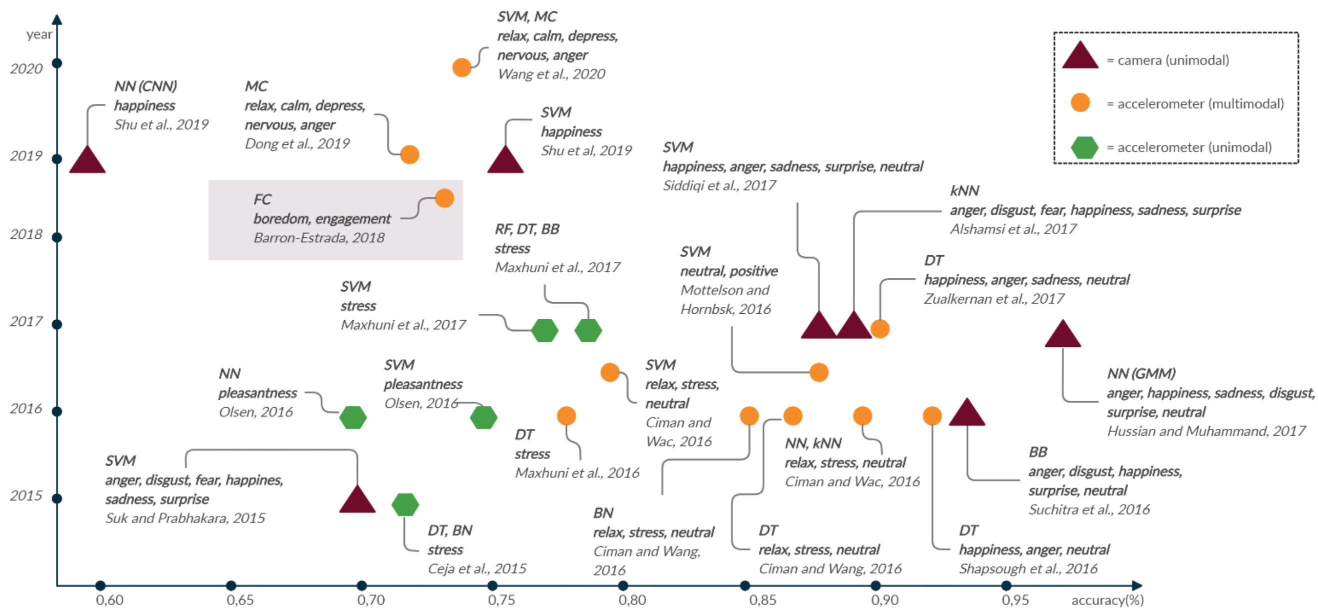

Figure 2 Accuracy scores, targeted emotions and used methods in camera and accelerometer-based non-intrusive emotion detection tasks over the years 2015-2020

About *targeted emotions (including LREs)*

In this section we provide the findings of the review analysis we conducted to reply to RQ3.

> RQ3: *What are the main emotions, including learning related emotions that have been recognized by mobile sensing algorithms?*

Table 8 presents the frequency of the emotions that were detected in the reviewed studies, determining the LREs among them. Many mobile emotion sensing studies (about 40%) are targeted towards the recognition of the six basic Ekman's (Ekman, 2007) emotions. Several studies also include the examination of a neutral state (Hussian et al., 2017; Maxhuni et al., 2017; Suchitra et al., 2016; Zualkernan et al., 2017) and stress (Ceja et al., 2015; Cimac and Wac, 2016; Ghosh et al., 2019a; b;c; Gjoreski et al., 2015; Mahuni et al., 2016;2017).
Regarding studies on LREs, these are quite limited (only five articles) and they mainly investigate the emotions of boredom, confusion, engagement, attention, and perceived difficulty. However,

some LREs like happiness, joy, contempt, stress, nervousness, and anger that have been examined outside the educational context, included university students as participants. Unfortunately, several important LRES like for instance curiosity which tends to be a determinant learning emotion for students' learning achievement in digital environments (D'Mello and Graesser, 2014, Lodge et al., 2018) are totally missing.

On the other hand, there are some relevant studies that were excluded from this review because they did not use a data prediction/machine learning model, but rather some qualitative or correlation analysis. A representative example is the work of Schrader and Kalyuga (2020) that identified that students apply more pressure on a pen-based tablet when feel frustrated. They also found an indirect relationship between pen pressure and performance, mediated by student engagement.

The study of Bahreini et al., (2016) designed a multimodal emotion recognition methodology to detect students' feedback through web camera in microphone in online learning. The study was excluded from this review because it is applied on computer devices and not mobile phones. However, the remarkably high accuracy scores (<96%) imply its contribution in the context of synchronous distance learning using computer desktop and laptop devices.

Table 8 Recognised emotion frequency

Emotion	LRE*	Num of studies	Emotion	LRE*	Num of studies
anger	YES	12	boredom	YES	3
sadness	NO	11	positive-negative	NO	3
happiness	NO	10	contempt	YES	2
stress	YES	10	depress	NO	2
relax	NO	6	pleasantness/joy	YES	2
disgust	NO	6	attention	YES	1
surprise	NO	6	confusion/difficulty	YES	2
fear	YES	4	engagement	YES	3

*LRE stands for Learning Related Emotion

Table 9 depicts the study characteristics of the reviewed works that focused on students' emotion detection in learning environments. As presented, Massive Open Online Courses (MOOCs) was the main learning environment that was used, and most studies conducted their research on small sample sizes (less than 50 students). Finally, they were all conducted on students of higher education, either undergraduates or postgraduates.

According to the above findings and the strong potential of mobile sensing methodologies to detect LRES, this reviewed work makes it clear that more research needs to be conducted in the educational context to unobtrusively detect students' emotions and provide emotion-based personalized services or empathetic agents. Empathetic agents have been shown to be effective in persisting students' positive emotions and altering an emotional state of fear to a neutral one (Moridis and Economides, 2012).

Overall, this literature review strongly encourages the research community towards the design of non-intrusive frameworks for the detection of all basic LREs in m-learning environments, without the need to either interrupt or disrupt the students' natural learning flow. This implementation will be useful in the covid-19 distance education where students' emotions play a determinant role in their engagement and performance (Besser et al., 2020; Ez-zaouia et al., 2020, Irawan, 2020) and online communication with professor or interaction with peers might be difficult. Since many students use their mobile devices to attend DE courses, the design of emotion aware m-learning tools and platforms will significantly contribute towards the enhancement of their performance since teachers, or the systems will be able to provide them with appropriate feedback.

Table 9 Student populations and learning context in LREs studies

Author (Year)	Sample size	Learning Environment	Educational level
Pham and Wang (2018)	28	Massive Open Online Courses (MOOCs)	Higher education-Undergraduate students
Barron-Estrada et al. (2018)	10	Educational Android app	Higher education-undergraduate students
Pham and Wang (2016)	32	Massive Open Online Courses (MOOCs)	Higher education-Undergraduate students
Xiao and Wang (2016)	48	Massive Open Online Courses (MOOCs)	Higher education-Graduate students
Xiao and Wang (2015)	18	Massive Open Online Courses (MOOCs)	Higher education-Undergraduate and Graduate students

4.3. Challenges and limitations

Given that different sensors have different data features that can be used to capture the emotional state of users (Wang et al., 2020), categories of sensor modalities can be defined according to specific context, activity, device, or other criteria (Shu et al., 2019). However, not all these interactions allow for unobtrusive data collection process. Researchers agree that the use of unobtrusive methods for emotion detection is important since users do not want to feel any burden associated with the recognition process (Lee et al., 2012; Wahid and Rasheed, 2019).

According to D'Mello (2015) "multimodal systems have been shown to be consistently more accurate than their best unimodal counterparts, with an average improvement of 9.83% (median of 6.60%)". A representative example is the study of Kanjo et al. (2019) where the authors achieved a significant 95% learning accuracy score, while using both build-in sensors (accelerometer, gyroscope, GPS) and an external device, the Microsoft wrist Band 2 to capture heart rate (HR) data, galvanic skin response (SGR) and body temperature as well as environmental data. Although such hybrid approaches including physiological inputs tend to increase the accuracy levels of algorithmic emotional classification, they undoubtedly disrupt the user's natural flow of mobile interaction for learning and hence they are not considered useful and practical to be implemented in m-learning activities. The researchers also agree that monitoring physiological signals relies on expensive dedicated devices, which is impossible for pervasive device-free detection (Zhang et al., 2018). For this reason, several attempts have been made to non-invasively sense mobile emotion, revealing very promising results and significantly activating the social media and the interest of the mobile industry.

Unfortunately, there are a limited number of studies on LREs recognition in mobile-learning environments (including non-intrusive physiological signal collection approaches). A relevant study on mobile-learning that uses the multimodal process affect recognition is the work of Benta et al. (2015) where the authors receive data from built-in mobile sensors (accelerometer), front camera, using the FaceReader module (Nodus Information Technology, 2019) and external sensor (heart-rate device chest belt). Their research is focused on the detection of Ekman's basic emotions using a mobile-learning application in a foreign language. However, their multimodal approach has serious limitations due to the small sample size (4 participants) and the low accuracy score the emotion detection model.

Although mobile emotion sensing has achieved great success in experimental evaluations, several issues remain that deserve further investigation.

One main limitation of the presented studies is that, although all of them used non-intrusive methods they needed a long period of time to collect their input data. For instance, it was a five weeks' period in Maxhuni et al. (2017), 8 weeks in Maxhuni et al. (2016), 12 weeks in Ceja et al. (2015) and 4 weeks in Ciman and Wac (2016). Also, most of the reviewed works used small datasets, rendering weak their findings for broader and generalized applications. Recently though there are efforts (e.g., Ceja et al., 2019) that demonstrate how deep learning methodologies and data augmentation can be used for mobile emotion recognition on limited and shortly in time collected training data sets.

Second, most of the reviewed studies are mainly focused on the detection of the Ekman's basic emotions or positive/negative/neutral affective states, and not on specific LREs, revealing the research gap in m-learning environments. Unfortunately, there is a limited amount of mobile emotion sensing studies that use solely non-intrusive methodologies to recognize learning related emotions.

Third, whereas an emotion aware system for educational/ learning contexts should work in real time and achieve high accuracy (Hossain and Muhammad, 2017) most of the cited studies have conducted their research on small sample sizes and do not provide any real-time detection or feedback response. Indicatively, Xiao and Wang (2016) used a sample only of 10 students and a sample of 18 students in their earlier work (Xiao and Wang, 2015). Shah et al. (2015) used a sample of 32 participants, while in a more encouraging attempt Mottelson and Hornbsk (2016) gathered feedback from 127 respondents. The studies that provided a framework for real-time emotion recognition used solely camera-based input. Suk and Prabhakara (2016) used SVM classifier to identify camera-input face emotions achieving an accuracy score of 70.6%, and Suchitra et al. (2016) used the Adaboost classifier to recognize the six basic emotions, achieving an accuracy of 94%. A possible solution for real time emotion detection of high accuracy in mobile devices, due to their limited processing capacities, can be the transportation of the computational part to the Cloud. As suggested by Wang et al. (2020), "the processing of the emotion recognition algorithm is offloaded to the cloud. Mobile phones are just responsible for the collection of sensor data and showing the recognition results".

In general, despite the recently emerged high-resolution front-facing cameras which allow facial expression analysis, biometric applications, and gaze interaction, we notice no significant research outcomes in real-time facial emotion recognition via mobile devices since face detection algorithms tend to perform poorly against front-facing cameras visual inputs. A main cause is the difficulty to efficiently capture the users' face in real time since it is unknown how often users hold devices in a way that allows capturing their full face. In fact, according to a recent study (Khamis et al., 2018) full face is visible about 29% of the time. Although evidence for front-face eye gaze research is more optimistic (eyes have been shown to be visible 75% of the time), most of the learning-related studies use only external or wearable mobile devices to conduct eye-tracking research on tablets and Smartphone devices (Schneider et al., 2016; 2018). Furthermore, tracking eye gaze activity via front camera has proved efficient and useful especially for accessibility inclusiveness issues. Eye-tracking though seems a promising approach in the educational context since recent studies (e.g., Hutt et al., 2017; Sharma, 2020) using eye-tracking data (including intrusive methods to collect signals e.g., with EEG and wristband) revealed promising outcomes in predicting students' effort to complete the upcoming task.

Finally, we see no significant gender-oriented research in mobile sensing studies. Gender is only analyzed in the work of Ceja et al., (2019) to compare accuracy scores of the UAM between gender populations. It is a fact that neuroscientists agree on the existence of sex differences in the neural correlates of emotional reactivity and several regions of activation have been found to differ between men and women during the processing of emotional visual stimuli (Filkowski et al., 2018). To this end, intelligence tutoring systems could take into consideration the affective differences between genders to provide effective services of engagement or learning experience. The grounded gender differences in different areas of HCI and the persistent gender gap in computing fields including education (EC, 2019; SMART 2016/0025, 2018), implies that further actions should be taken to increase the female participation in computer related and m-learning related activities.

Overall, future research on mobile emotion sensing should be conducted on larger populations, be further expanded in the field of m-learning and LRES detection, and trace new unimodal or multimodal methodologies to achieve high accuracy scores in detecting learners' emotional states. Furthermore, recent, and effective research works towards the design of user-adaptive models (UAM) that adapt to users' behavior and emotions, the latter being recognized through unobtrusive means (e.g., accelerometer and speech recognition in the work of Ceja et al., 2019), we believe that mobile emotion sensing can bring new research achievements in sensor-based and ubiquitous emotion recognition for affective UAM in mobile and mobile learning activities.

5. Conclusions

This study presents a review of the state of the art in Mobile Sensing as a promising technology for supporting non-intrusive emotion detection. Focusing on learning related emotions, this review is the first attempt towards reviewing non-intrusive emotion sensing studies using mobile devices. The study attempted to highlight all the recent implementations in the context of m-learning to provide useful insights on the design of emotion-aware m-learning systems. Such an approach seems to be particularly useful in the age of pandemic crisis where many students use their mobile devices to attend distance courses. Moreover, student's emotions are implied by researchers to play a determinant role in emergency remote education and strongly affect their performance.

The generic purpose of this study is to provide the researchers in the field with a comprehensive overview of the used modalities, the algorithmic models and the performance scores that have been achieved within the last years. The first research question aimed to identify the modalities (built-in sensors and internal devices or services) that have been used in emotion recognition tasks. In this regard, camera and accelerometer sensor were found to be the most frequent

modalities, while multimodal accelerometer-inclusive methodologies and PPG were mainly used in the detection of LREs.
The second research question was formulated to examine the detection/prediction models that were used in the reviewed studies as well as the accuracy scores they achieved to detect specific emotions. In this regard, it was found that most studies deployed the SVM and DT classification models, while DT and SVM were mainly deployed to predict LREs. While Linear Regression models showed the average highest accuracy, it was a GMM Neural Network approach that achieved the highest score (almost 99%) to classify the emotions of anger, happiness, sadness, disgust, surprise, and neutral state, based on camera input. Regarding LREs, the highest accuracy score (77%) was achieved in 2015 via a Fuzzy Classification approaching and using the PPG modality, to detect the emotions of boredom and engagement.
Finally, the third research question sought to determine the main emotions and the specific LREs that were most targeted by the reviewed studies. It was found that anger, sadness, and happiness were the most frequently detected emotions, while the LREs of attention, confusion and perceived difficulty were the rarer to me measured.
To sum up, the review results revealed the following:

- Accelerometer sensor and mobile camera are the most frequently used modalities in emotion-detection tasks in Smartphone devices.
- The main emotions to be recognized are anger, happiness, sadness, disgust, and surprise.
- Engagement and boredom are the most frequently detected learning related emotions in the more limited context of education.
- Several machine learning models have been used achieving different scores of accuracies, generally higher than 70%.

Overall, the results have highlighted the main trends and directions of mobile emotion sensing in Smartphone devices. The findings of this review work reveal a research gap in the field of mobile emotion sensing for m-learning environments, and especially on the detection of learners' LREs via non-intrusive and real time methodologies. To this end, this study encourages for more extended research on LREs detection via non-intrusive methods in m-learning environments.

5.1. Limitations and future work

This review was limited in that it examined articles from seven databases: ACM, Google Scholar and Google Semantics, IEEE, Web of Science, Elsevier, Science Direct and Springer Link, from 2015 to 2020. The articles in these databases are considered to have a high impact on the field; however, the latest technical reports and business demonstrations of mobile sensing in emotion detection were excluded from this review, which may limit the representation of the state of the art for these applications. Furthermore, this review includes the relevant studies being published within 2020, however, more studies are expected to be published during this year, and hence the year-specific analysis is not representative.
Finally, this review work could be possibly expanded in the future by collecting mobile intrusive studies as well and conducting comparative examinations between types of modalities. It also is advisable to conduct further meta-analysis or review studies on the field to deeply evaluate the classification techniques and design enhanced version per type of research (e.g., according to the modalities used and the emotions to be detected).

References

1. Afergan, D., Peck, E.M., Solovey, E.T, Jenkins, A., Hincks, S.W., Brown, E.T., Chang, R., and Jacob, R.J.K. (2014). Dynamic difficulty using brain metrics of workload. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14).* Association for Computing Machinery, New York, NY, USA, 3797–3806.
2. Ali, S., Khusro S, Rauf A, Mahfooz S. (2014). Sensors and Mobile Phones: Evolution and State-of-the-Art. *Pakistan Journal of Science.* 66(4), 385-399.
3. Alshamsi, H., Meng, H., and Li, M. (2016). Real time facial expression recognition app development on mobile phones. 12th International Conference on Natural Computation. *Fuzzy Systems and Knowledge Discovery (ICNC-FSKD),* Changsha, 1750-1755.
4. Android Developers (2020). Sensor Overview. Online: https://developer.android.com/guide/topics/sensors/sensors_overview
5. Ashwin, T.S., Guddeti, R.M.R. (2020). Impact of inquiry interventions on students in e-learning and classroom environments using affective computing framework. User Model User-Adap Inter.
6. Barron-Estrada, M.L., Zatarain-Cabada, R., and. Aispuro-Gallegos, C.G. (2018). Multimodal Recognition of Emotions with Application to Mobile Learning, 2018 *IEEE 18th International Conference on Advanced Learning Technologies* (ICALT), Mumbai, 416-418.
7. Benta, K-I. (2015). Affective aware museum guide. *IEEE International Workshop on Wireless and Mobile Technologies in Education* (WMTE'05), 53-55.
8. Besser, A., Lotem, S., & Zeigler-Hill, V. (2020). Psychological stress and vocal symptoms among uUniversity pProfessors in Israel: Implications of the shift to online synchronous tTeaching dDuring the COVID-19 pPandemic. Journal of Voice.
9. Bhattacharya, S. (2015). A linear regression model to detect user emotion for touch input interactive systems, International Conference on Affective Computing and Intelligent Interaction (ACII), Xi'an, 970-975.
10. Bogomolov, A., Lepri, B., Ferron, M., Pianesi, F., and Pentland, A.S. (2014). Daily Stress Recognition from Mobile Phone Data, Weather Conditions and Individual Traits. In Proceedings of the 22nd ACM international conference on Multimedia (MM '14). Association for Computing Machinery, New York, NY, USA, 477–486.
11. Bosch, N., D'Mello, S., Baker, R., Ocumpaugh, J., Shute, V., Ventura, M., Wang, L., and Zhao, W. (2015). Automatic Detection of Learning-Centered Affective States in the Wild. *In Proceedings of the 20th International Conference on Intelligent User Interfaces (IUI '15*). ACM, New York, NY, USA, 379-388.
12. Calvo, R.A., and D'Mello, S.K. (2011). New Perspectives on Affect and Learning Technologies. *Springer Science & Business Media*, 3, Heidelberg.
13. Ceja, E., Osmani, V., Mayora, O. (2015). Automatic stress detection in working environments from smartphones' accelerometer data: a first step, IEEE J. Biomed. Health Inform. Jul; 20(4), 1053-60.
14. Ceja, E., Riegler, M., Kvernberg, A.K. et al. (2019). User-adaptive models for activity and emotion recognition using deep transfer learning and data augmentation. *User Model User-Adap Inter.*
15. Ciman, M., and Wac, K. (2016). Individuals' stress assessment using human-smartphone interaction analysis. *IEEE Trans. Affective Comput*. 99, 1–15.

16. D’Mello, S. (2013). A selective meta-analysis on the relative incidence of discrete affective states during learning with technology. *Journal of Educational Psychology* 105, 4, 1082–1099.
17. Dettmers, S., Trautwein, U., Lüdtke, O., Goetz, T., Frenzel, A. C., and Pekrun, R. (2011). Students’ emotions during homework in mathematics: Testing a theoretical model of antecedents and achievement outcomes. *Students’ Emotions and Academic Engagement*, 36(1), 25–35.
18. D'Mello, S., and Graesser, A. (2014). Confusion, *in Handbook of Emotions and Education* eds R. Pekrun and L. Linnenbrink-Garcia (New York, NY: Routledge), 289–310.
19. D'Mello, S., Graesser, A. (2012). Dynamics of affective states during complex learning. Learn. Instr. 22 (2), 145–157.
20. Dong, L., Xu, Y., Wang, P., and He, S. (2019). Classifier fusion method based emotion recognition for mobile phone users, i*n Proc. Int. Conf. Broadband Commun., Netw. Syst.* Cham, Switzerland: Springer, 216–226
21. EC (2019) Online: https://ec.europa.eu/digital-single-market/en/women-ict, online accessed 18/2/2019.
22. Ekman, P. (1992). An argument for basic emotions, Cogn. Emotion 6 (3–4), 169–200.
23. Emden, C., & Smith, C. (2004). Non-intrusive research: Ideas and guidelines for expedient thesis completion. Australian Journal of Advanced Nursing, 22(1), 37–43.
24. Ez-zaouia, M., Tabard, A., & Lavoué, E. (2020). EMODASH: A dashboard supporting retrospective awareness of emotions in online learning. *International Journal of Human Computer Studies*, 139 (December 2019), 102411.
25. FaceReader (2021) Loijens, L. and Krips, O. “FaceReader Methodology Note”, Whitepaper Accessed online at https://info.noldus.com/free-white-paper-on-facereader-methodology on June 24, 2021
26. Filkowski, M.M., Olsen, R.M., Duda, B., Wanger, T.J., and Sabatinelli, D. (2018). Sex differences in emotional perception: Meta-analysis of divergent activation. *NeuroImage*, 147, 925-933.
27. Ghosh, S. Goenka,S., Ganguly, N.,. Mitra, B., and De, P. (2019c). Representation Learning for Emotion Recognition from Smartphone Keyboard Interactions, *2019 8th International* Conference on Affective Computing and Intelligent Interaction (ACII), Cambridge, United Kingdom, 704-710.
28. Ghosh, S., Hiware, K., Ganguly, N., Mitra, B., & De, P. (2019a). Emotion detection from touch interactions during text entry on smartphones. *International Journal of Human Computer Studies*, 130(March), 47–57.
29. Ghosh, S., Sahu, S., Ganguly, N., Mitra, B., and De, P. (2019b). EmoKey: An Emotion-aware Smartphone Keyboard for Mental Health Monitoring, *2019 11th International Conference on Communication Systems & Networks (COMSNETS)*, Bengaluru, India, 2019, pp. 496-499.
30. Gilleade, K., Dix, A., Allanson, J. (2005). Affective videogames and modes of affective gaming: assist me, challenge me, emote me. In: DiGRA 2005: Changing Views – Worlds in Play.
31. Gjoreski, M., Gjoreski, H., Luštrek, M., and Gams, M. (2015). Automatic Detection of Perceived Stress in Campus Students Using Smartphones. In Proceedings of the 2015

International Conference on Intelligent Environments (IE ’15*). IEEE Computer Society*, USA, 132–135.

32. Gu J., Wang Z., Kuen J., Ma L., Shahroudy A., Shuai B., Liu T., Wang X., Wang L., Wang G., et al. (2017). Recent advances in convolutional neural networks. *Pattern Recognit*. 1, 1–24.
33. Hossain M.S., and Muhammad, G. (2017). An emotion recognition system for mobile applications, in *IEEE Access*, 5, 2281-2287.
34. Huang, Y., Chang, Dian-Fu and Wu, B. (2017). Mobile Game-Based Learning with a Mobile App: Motivational Effects and Learning Performance. *Journal of Advanced Computational Intelligence and Intelligent Informatics*. 21, 963-970.
35. Hussian, M.S. and Muhammad, G. (2019). Emotion recognition using deep learning approach from audio–visual emotional big data, *Information Fusion*, 49, 69-78.
36. Hussian, M.S., and Mohammed, G. (2017). An Emotion Recognition System for Mobile Applications, IEEE, 5, 2281, 2017.
37. Hutt, S., Mills, C., Bosch, N., Krasich, K., Brockmole, J., and D’Mello, S. (2017). Out of the Fr-Eye-ing Pan”: Towards Gaze-Based Models of Attention during Learning with Technology in the Classroom. *In Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization (UMAP ’17).* Association for Computing Machinery, New York, NY, USA, 94–103.
38. Irawan, A. W., Dwisona, D., & Lestari, M. (2020). Psychological Impacts of Students on Online Learning During the Pandemic COVID-19. KONSELI: *Jurnal Bimbingan Dan Konseling* (E-Journal), 7(1), 53–60.
39. Jain, N., Kumar, S., Kumar, A., Shamsolmoali, P., and Zareapoor, M. (2018). Hybrid deep neural networks for face emotion recognition, *Pattern Recognition Letters*, 115, 101-106.
40. Jeong, S., and Breazeal, C.L. (2016). Improving smartphone users’ affect and wellbeing with personalized positive psychology interventions. *In Proceedings of the Fourth International Conference on Human Agent Interaction (HAI ’16). Association for Computing Machinery*, New York, NY, USA, 131–137.
41. Kanjo, E., Daria, J.K., and Chee, S.A. (2017). NotiMind: Utilizing responses to smart phone notifications as affective sensors. *IEEE Access,* 5, 22023–22035.
42. Kanjo, E., Younis, E., Chee S.A. (2019). Deep learning analysis of mobile physiological, environmental and location sensor data for emotion detection*, Information Fusion,* 49, 46-56.
43. Kapoor, A., Burleson, W., Picard, R.W. (2007). Automatic prediction of frustration. Int. J. Hum. Comput. Stud. 65 (8), 724–736.
44. Khamis, M., Baier, A., Henze, N., Alt, F., and Bulling, A. (2018). Understanding face and eye visibility in front-facing cameras of smartphones used in the wild*. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (CHI '18). ACM, New York, NY, USA, Paper 280, 12 pages.
45. Khan, W.Z., Xiang, Y., Aalsalem, M.Y., and Arshad, Q. (2013). Mobile phone sensing systems: a survey, in *IEEE Communications Surveys & Tutorials*, 15(1), 402-427, First Quarter 2013.
46. Kiavash Bahreini, Rob Nadolski & Wim Westera (2016) Data Fusion for Real-time Multimodal Emotion Recognition through Webcams and Microphones in E-Learning, *International Journal of Human–Computer Interaction*, 32(5), 415-430

47. Kitchenham, B., Charters, S. (2007) Guidelines for performing systematic literature reviews in software engineering. Technical report EBSE 2007-001, Keele University and Durham University Joint Report.
48. Ko, B.C. (2018). A Brief Review of Facial Emotion Recognition Based on Visual Information. *Sensors* (Basel, Switzerland), 18(2), 401.
49. Kołakowska, A., Szwoch, W., & Szwoch, M. (2020). A review of emotion recognition methods based on data acquired via smartphone sensors. *Sensors* (Switzerland), 20(21).
50. Kort, B., Reilly,R., and Picard, R.D. (2001). An affective model of interplay between emotions and learning: reengineering educational pedagogy-building a learning companion. *Proceedings IEEE International Conference on Advanced Learning Technologies*, Madison, WI, USA, 43-46.
51. Kwak, T. and Kim, M. (2018). Mobile App Classification Method Using Machine Learning Based User Emotion Recognition. *In Proceedings of the 8th International Conference on Information Communication and Management (ICICM '18*). ACM, New York, NY, USA, 22-26.
52. Lane, N.D., and Georgiev, P. (2015). Can deep learning revolutionize mobile sensing? In Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications (HotMobile ’15). *Association for Computing Machinery*, New York, NY, USA, 117–122.
53. Laport-López, F., Serrano, E., Bajo, J. et al. (2020). A review of mobile sensing systems, applications, and opportunities. *Knowl Inf Syst* 62, 145–174.
54. Lee, H., Choi, Y.S., Lee, S. and Park, I.P. (2012). Towards unobtrusive emotion recognition for affective social communication. *IEEE Consumer Communications and Networking Conference (CCNC)*, Las Vegas, NV, 260-264.
55. Lehman B., Matthews M., D’Mello S., Person, N. (2008). What are you feeling? Investigating student affective states during expert human tutoring sessions. In: Woolf B.P., Aïmeur E., Nkambou R., Lajoie S. (eds) Intelligent Tutoring Systems. ITS 2008. *Lecture Notes in Computer Science*, vol 5091. Springer, Berlin, Heidelberg.
56. Liu, B., & Koc, A. B. (2015). Mobile Phone Sensing in Scientific Research. Encyclopedia of Mobile Phone Behavior, 410–423.
57. Lodge, J.M., Kennedy, G., Lockyer, L., Arguel, A. and Pachman, M. (2018). Understanding Difficulties and Resulting Confusion in Learning: An Integrative Review. *Front. Educ*. 3(49).
58. Lucey, P., Cohn, J.F., Kanade, T., Saragih, J., Ambadar, Z., and Matthews, I. (2010). The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression*. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, San Francisco, CA, 94-101.
59. Lynch, R., Patten, J.V. and Hennessy, J. (2013). The impact of task difficulty and performance scores on student engagement and progression. *Educational Research*, 55(3), 291-303.
60. Maxhuni, A., Hernandez-Leal, P., Morales, E.F., Sucar, L.E., Osmani, V., Muńoz-Melındez, V., Mayora, O. (2017). Using intermediate models and knowledge learning to improve stress prediction, *Applications for Future Internet*, Springer, 140–151.
61. Maxhuni, A., Hernandez-Leal, P., Morales, E.F., Sucar, L.E., Osmani, V., Muńoz- Melındez, V., Mayora, O. (2016). Stress modelling and prediction in presence of scarce data, J. *Biomed. Inform*. 63, 344–356.

62. Mehrabian, A., and Russell, J.A. (1974). An approach to environmental psychology (1 ed.). Cambridge, Mass.: MIT Press.
63. Meyer, M., Helmholz, P., Rupprecht, M., Seemann, J.D., Toennishoff, T., & Robra-Bissantz, S. (2019). From the Inside Out: A Literature Review on Possibilities of Mobile Emotion Measurement and Recognition. Bled eConference
64. Moridis, C.N., and Economides, A.A. (2012). Affective Learning: Empathetic Agents with Emotional Facial and Tone of Voice Expressions. *IEEE Transactions on Affective Computing*, 3(3), 260-272.
65. Mottelson, A. and Hornbζ, K. (2016). An affect detection technique using mobile commodity sensors in the wild, *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM, 781–792.
66. Nightingale, A. (2009). A guide to systematic literature reviews, *Surgery (Oxford*),27(9), 381-384
67. Nikou, S.A. and Economides, A.A. (2019). A comparative study between a computer-based and a mobile-based assessment: Usability and user experience. *Interactive Technology and Smart Education*, 16(4), 381-391.
68. Odenwald, S. (2019). A Guide to Smartphone Sensors Experimeter ' s Guide To Smartphone Sensors. National Aeronautics and Space Administration. Retrieved from https://spacemath.gsfc.nasa.gov/Sensor/SensorsBook.pdf on November, 13, 2020.
69. Okoli, C., Schabram, K. (2010). A Guide to Conducting a Systematic Literature Review of Information Systems Research. Sprouts: Working Papers on Information Systems, 10(26).
70. Olsen A.F., and Torresen, J. (2016). Smartphone accelerometer data used for detecting human emotions, 3rd International Conference on Systems and Informatics (ICSAI), Shanghai, 410-415.
71. Pekrun, R., and Stephens, S.J. (2010). Achievement emotions in higher education. In Higher Education: Handbook of Theory and Research; Smart, J.C., Ed.; Springer: New York, NY, USA, 25, 257–306
72. Pekrun, R., Goetz, T., Frenzel, A.C., Barchfeld, P., and Perry, R.P. (2011). Measuring emotions in students' learning and performance: the achievement emotions questionnaire (AEQ). *Contemp. Educ.Psychol*. 36, 36–48.
73. Pham, P., and Wang, J. (2016). Adaptive review for mobile MOOC learning via implicit physiological signal sensing. *In Proceedings of the 18th ACM International Conference on Multimodal Interaction (ICMI '16).* ACM, New York, NY, USA, 37-44.
74. Pham, P., and Wang, J. (2018). Adaptive review for mobile MOOC learning via multimodal physiological signal sensing - A longitudinal study. In Proceedings of the 20th ACM International Conference on Multimodal Interaction (ICMI '18). *Association for Computing Machinery*, New York, NY, USA, 63–72.
75. Poria, S., Cambria, E., Bajpai, R., and Hussain, A. (2017). A review of affective computing. Inf. Fusion 37, C (September), 98-125.
76. Radhamani, V., and Dalin, S. (2018). A supporting survey to step into a novel approach for providing automated emotion recognition service in mobile phones, *2nd International Conference on Inventive Systems and Control (ICISC)*, Coimbatore, 35-39.
77. Reychav, I., McHaney, R. (2017). The relationship between gender and mobile technology use in collaborative learning settings: An empirical investigation, *Computers & Education*, 113, 61-74.

78. Rowe, A.D., and Fitness, J. (2018). Understanding the Role of Negative Emotions in Adult Learning and Achievement: A Social Functional Perspective. *Behav Sci (Basel).* 8(2),27.
79. Ruiz-Garcia, A., Elshaw, M., Altahhan, A. et al. (2018). *Neural Comput & Applic*. 29, 359.
80. Schneider, B., Sharma, K., Cuendet, S. et al. (2018). Leveraging mobile eye-trackers to capture joint visual attention in co-located collaborative learning groups. Intern. J. Comput.-Support. Collab. Learn 13, 241–261.
81. Schrader, C., & Kalyuga, S. (2020). Linking students' emotions to engagement and writing performance when learning Japanese letters with a pen-based tablet: An investigation based on individual pen pressure parameters. *International Journal of Human Computer Studies*, 135 (November 2019), 102374.
82. Schrader, C., Nett, U. (2018). The perception of control as a predictor of emotional trends during gameplay. Learn. Instr. 54, 62–72.
83. Shah, S., Teja, J.N. and Bhattacharya, S. (2015). Towards affective touch interaction: predicting mobile user emotion from finger strokes. *J Interact Sci,* 3, 6.
84. Shapsough, S., Hesham, A, Elkhorazaty, Y., Zualkernan, I.A and Aloul, F. (2016). Emotion recognition using mobile phones. *IEEE 18th International Conference on e-Health Networking, Applications and Services* (Healthcom), Munich, 1-6.
85. Sharma, K., Papamitsiou, Z., Olsen, J.K., and Giannakos, M. (2020). Predicting learners' effortful behaviour in adaptive assessment using multimodal data. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (LAK '20). Association for Computing Machinery, New York, NY, USA, 480–489.
86. Shu, J., Chiu, M., & Hui, P. (2019). Emotion Sensing for Mobile Computing. *IEEE Communications Magazine*, 57(11), 84-90.
87. Shun, M. C. Y., Yan, M. C. Bo, A., and Cyril, L. (2015). Modeling Learner's Emotions with PAD, *2015 IEEE 15th International Conference on Advanced Learning Technologies*, 2015, 49-51.
88. Shute, V.J., D'Mello, S., Baker, R., Cho, K., Bosch, N., Ocumpaugh, J., Almeda, V., 2015. Modeling how incoming knowledge, persistence, affective states, and in-game progress influence student learning from an educational game. Comput. Educ. 86, 224–235.
89. Siddiqi, M.H., Alruwaili, M., Bang, J-H., and Lee, S. (2017). Real time human facial expression recognition system using smartphone, IJCSNS *International Journal of Computer Science and Network Security*, 17, 10.
90. SMART 2016/0025 (2018). Women in the Digital Age FINAL REPORT A study prepared for the European Commission DG Communications Networks, Content & Technology by: iClaves. Online: https://ec.europa.eu/digital-single-market/en/news/increase-gender-gap-digital-sector-study-women-digital-age, online accessed: 19/2/2019.
91. Suchitra, S.P., and Tripathi, S. (2016). Real-time emotion recognition from facial images using Raspberry Pi II, *3rd International Conference on Signal Processing and Integrated Networks* (SPIN), Noida, 666-670.
92. Suk, M., and Prabhakaran, B. (2015). Real-Time facial expression recognition on smartphones. *2015 IEEE Winter Conference on Applications of Computer Vision*, Waikoloa, HI, 1054-1059.
93. Terzis, V., Moridis, C.N. & Economides, A.A. (2013). Measuring instant emotions based on facial expressions during computer-based assessment. *Pers Ubiquit Comput,* 17, 43–52 (2013).

94. Van Lankveld, G., Spronck, P., Van Den Herik, H.J., Rauterberg, M., 2010. Incongruitybased adaptive game balancing. In: Van Den Herik, H.J., Spronck, P. (Eds.), Advances in Computer Games. 12th International Conference ACG 2009, Pamplona Spain, May 11-13, 2009. Berlin Heidelberg. Springer, pp. 208–220.
95. Wahid, A. and Rasheed,F. (2019). Emotion Detection Using Unobtrusive methods: An Integrated Approach. 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kannur,Kerala, India, 1166-1171.
96. Wang, P., Dong, L., Xu, Y., Liu, W., and. Jing, N. (2020). Clustering-Based Emotion Recognition Micro-Service Cloud Framework for Mobile Computing*, in IEEE Access*, 8, 49695-49704.
97. Xiao, X., and Wang, J. (2015). Towards Attentive, Bi-directional MOOC Learning on Mobile Phones. In Proc. ICMI 2015.
98. Xiao, X., and Wang, J. (2016). Context and cognitive state triggered interventions for mobile MOOC learning. *In Proceedings of the 18th ACM International Conference on Multimodal Interaction (ICMI '16).* ACM, New York, NY, USA, 378-385
99. Zeng, X. (2017). Mobile Sensing Through Deep Learning. In Proceedings of the 2017 Workshop on MobiSys 2017 Ph.D. Forum (Ph.D. Forum '17). Association for Computing Machinery, New York, NY, USA, 5–6.
100. Zhang, X., Li, W., Chen, X., and Lu, S. (2018). MoodExplorer: towards compound emotion detection via smartphone sensing. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 1(4), Article 176 (January 2018), 30 pages.
101. Zhang, X., Sun, W., Liu, Y., Tang, S., Xing, K., and Mao, X. (2016). Incentives for Mobile Crowd Sensing: A Survey, in *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 54-67, Firstquarter 2016.
102. Zualkernan, I., Aloul, F. Shapsough, S., Hesham, A., El-Khorzaty, Y. (2017). Emotion recognition using mobile phones, Computers & Electrical Engineering, 60, 1-13.

APPENDIX A

Table A1 Background information of the reviewed non-intrusive mobile sensing studies

Built-in Sensor	**Internal Device (Or service)**	**Emotion(s)**	**Learning Environment**	**Primary author (Year of publication)**
Unimodal				
	keyboard	happiness, sadness, stress, relax	No	Ghosh et al., 2019a
	keyboard	happiness, sadness, stress, relax	No	Ghosh et al., 2019b
	keyboard	happiness, sadness, stress, relax	No	Ghosh et al., 2019c

	front camera – static photo	happiness	No	Shu et al, 2019
	embodied camera –video frames	anger, happiness, sadness, disgust, surprise, neutral	No	Hussian and Muhammand, 2017
	touch screen (finger strokes and taps)	positive, negative, neutral	No	Bhattacharya, 2017
accelerometer		stress level (low, moderate, high)	No	Maxhuni et al., 2017
	front camera	anger, disgust, fear, happiness, sadness, surprise	No	Alshamsi et al., 2017
	front camera	happiness, anger, sadness, surprise, neutral	No	Siddiqi et al., 2017
	keyboard	positive, negative, neutral	No	Kanjo et al., 2017
accelerometer		pleasantness, arousal (activation)	No	Olsen and Torresen, 2016
	front camera	anger, disgust, happiness, surprise, neutral	No	Suchitra et al., 2016
	front camera, SMS	joy, anger, disgust, contempt, engagement, fear, sadness, surprise	No	Jeong and Lynn, 2016
	front-camera	anger, disgust, fear, happiness, sadness, surprise	No	Suk and Prabhakara, 2015
	touch screen (finger strokes)	sadness, anger, fear disgust	No	Shah et al., 2015
accelerometer		stress level (low, moderate, high	No	Ceja et al., 2015

Multimodal				
accelerometer, gravity, gyroscope, light sensor, GPS	pedometer, network/wi-fi	relax, depression, stress, anger	No	Wang et al., 2020
accelerometer, gravity, gyroscope, light sensor, GPS	pedometer, network/wi-fi	relax, depression, stress, anger	No	Dong et al., 2019
accelerometer, gyroscope	microphone	engagement, boredom	Yes/Educational Android app	Barron-Estrada et al., 2018
accelerometer	keyboard (soft sensor)	happiness, anger, sadness, neutral	No	Zualkernan et al., 2017
accelerometer	microphone, Google-Maps, network/wi-fi, call/app log	stress level (low, moderate, high)		Maxhuni et al., 2016
accelerometer, gyroscope, GPS, light sensor	touch screen screen on-off, network/wi-fi	relax, stress, neutral	No	Ciman and Wac, 2016
accelerometer	keyboard (soft sensor)	happiness, anger neutral	No	Shapsough et al., 2016
accelerometer, gyroscope	touch screen	neutral, positive	No	Mottelson and Hornbsk, 2016
accelerometer, GPS, light sensor	microphone, network/wi-fi, call log	stress (not stressed, slightly stressed, and stressed**)**	No	Gjoreski et al., 2015
	front- camera, touch screen, photoplethysmography (PPG) sensor for heart rate features	anger, fear, sadness, surprise, joy, disgust, contempt, engagement, attention	Yes /Massive Open Online Courses (MOOCs)	Pham and Wang, 2018
	Photoplethysmography (PPG) sensor for heart rate features	boredom, confusion	Yes /Massive Open Online Courses (MOOCs)	Xiao and Wang, 2015
	Photoplethysmography (PPG) sensor for heart rate	perceived difficulty	Yes /Massive Open Online Courses	Pham and Wang, 2016

	features		(MOOCs)	
	Photoplethysmography (PPG) sensor for heart rate features.	boredom/disengagement	Yes /Massive Open Online Courses (MOOCs)	Xiao and Wang, 2016