

Received November 10, 2021, accepted November 27, 2021, date of publication December 1, 2021,
date of current version December 15, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3132233

A Review of Emotion Recognition Methods From Keystroke, Mouse, and Touchscreen Dynamics

LIYING YANG^{ID} AND SHENG-FENG QIN^{ID}, (Member, IEEE)

School of Design, Northumbria University, Newcastle upon Tyne NE1 8ST, U.K.

Corresponding author: Sheng-Feng Qin (sheng-feng.qin@northumbria.ac.uk)

This work was supported in part by the China Scholarship Council (CSC) under Grant 201906300082, and in part by the Newton Prize 2019 under Award NP2PB_100047.

ABSTRACT Emotion can be defined as a subject's organismic response to an external or internal stimulus event. The responses could be reflected in pattern changes of the subject's facial expression, gesture, gait, eye-movement, physiological signals, speech and voice, keystroke, and mouse dynamics, etc. This suggests that on the one hand emotions can be measured/recognized from the responses, and on the other hand they can be facilitated/regulated by external stimulus events, situation changes or internal motivation changes. It is well-known that emotion has a close relationship with both physical and mental health, usually affecting an individual's and a team's work performance, thus emotion recognition is an important prerequisite for emotion regulation towards better emotional states and work performance. The primary problem in emotion recognition is how to recognize a subject's emotional states easily and accurately. Currently, there are a body of good research on emotion recognition from facial expression, gesture, gait, eye-tracking, and other physiological signals such as speech and voice, but they are all intrusive and obtrusive to some extent. In contrast, keystroke, mouse and touchscreen (KMT) dynamics data can be collected non-intrusively and unobtrusively as secondary data responding to primary physical actions, thus, this paper aims to review the state-of-the-art research on emotion recognition from KMT dynamics and to identify key research challenges, opportunities and a future research roadmap for referencing. In addition, this paper answers the following six research questions (RQs): (1) what are the commonly used emotion elicitation methods and databases for emotion recognition? (2) which emotions could be recognized from KMT dynamics? (3) what key features are most appropriate for recognizing different specific emotions? (4) which classification methods are most effective for specific emotions? (5) what are the application trends of emotion recognition from KMT dynamics? (6) which application contexts are of greatest concern?

INDEX TERMS Affective computing, keystroke dynamics, mouse dynamics, touchscreen dynamics, emotional signal features, emotion elicitation, emotion recognition, machine learning, applications of emotion recognition.

I. INTRODUCTION

The term “emotion” is not uniformly defined, when asked “what is emotion”, answers can be different [1]. Among them, reference [1] gave a relevant comprehensive description in which emotion is defined as “an episode of interrelated, synchronized changes in states of all or most of the five organismic subsystems in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism”. Based on this definition, emotion

The associate editor coordinating the review of this manuscript and approving it for publication was Kah Phooi (Jasmine) Seng^{ID}.

can be described as a multidimensional construct composed of cognitive, motivational, somatic, motoric and subjective elements [2]. This suggests that emotions can be noticed, measured, and facilitated by external stimulus events, situation changes, and internal motivation changes [1].

Measuring the changes in someone's emotional states is the first step towards emotion recognition. And to measure the changes, there is a need to establish a benchmark for classifying different emotional states. Many researchers used some emotion elicitation intermediaries such as short videos and images to stimulate different emotional states and thus to find out subjects' reactions to certain states [3]–[5]. Then

these reaction data together with emotional states were used as labeled data to train emotion classification algorithms [6].

The reason for recognizing emotion is that emotion plays an important role in people's everyday life. People's behaviors are often influenced by their emotional states [7] and the behaviors in turn affect both individual and teamwork performance [7]. It is widely believed that positive emotions on workplace can lead to smoother social interactions and more helping behaviors which thus contribute to positive consequences [8], [9]. Especially, when the whole world is facing the Covid-19 pandemic challenge, the conventional face-to-face teamwork model is affected since most people must work from home. In this case, the web provides an alternative way for "working together" [10], [11]. People work online, use some instant messaging applications to contact their colleagues, and even discuss with them in a virtual meeting room [11]. Nevertheless, isolating from each other geographically may arise some negative emotions or psychological problems which have negative influence on their remote cooperation, such as poor cooperation performance and working quality [12], [13]. Thus, knowing how to recognize emotion at edge and how to interfere emotion to help people keep a positive emotional state and concentrate on their work is important to benefit remote teamwork.

There are various measurement modalities can be used to measure emotions such as changes of facial and vocal expression patterns, body gestures, and complex multi-modalities [14]. Thus, recognizing emotions has attracted many research attentions from different data sources and emotion recognition methods. They can be generally classified into three categories:

- a) Use external body signals for emotion recognition, including facial expression, body gestures, gait, eye-tracking, etc. These signals can be easily noticed by others but not always reflect one's real emotional states [15]–[23].
- b) Use internal physiological signals such as heart rate, sphygmic, skin conductance, blood pressure, Electroencephalography (EEG), etc. These signals can be more precise in reflecting emotions but not be caught as easily as external body signals [24]–[31].
- c) Utilize other contextual signals other than body signal themselves such as voice, text content, KMT dynamics, which can be collected non-intrusively and unobtrusively [32]–[39].

There have been some systematic surveys or reviews in the field of emotion recognition from facial expression [17], body gestures [18], [19], eye-tracking [22], internal physiological signals [25], [27], voice [32]–[34], text [35], [36], so readers interested in the related topics can refer to them. Thus, this paper will not detail them. Instead, it will focus on emotion recognition from KMT dynamics because KMT-related data can be recorded by a tool running in the background without disturbing users' normal work [37]. This unobtrusive characteristic makes it easy to measure more natural emotions in

a normal work environment and suitable for some scenarios such as working environment where data privacy is a main concern.

The dynamics information contained in KMT data can be used as biometrics such as performance indicators [13] and enhanced passwords [38]–[40]. It can also be used for emotion recognition [4], [41], [42], which is the focus of this paper. There have been some review papers [43], [44], [46]–[49] related to emotion recognition from KMT dynamics. Some [43], [44], [46], [47] were published in 2013–2015, reflecting on some earlier research, while the recent paper [48] focused on advances and applications of KMT dynamics, and paper [49] only reviewed the research over the last decade. These studies summarized general information about emotion recognition procedures, features, datasets, and classification methods, but did not provide answers to further questions we concerned (See Table 1). Thus, there still lacks more detailed literature review on the state-of-the-art research in emotion recognition from KMT dynamics.

In order to fill this gap, this paper provides a systematic literature review on emotion recognition from KMT dynamics. Its contributions include:

- a) A systematic literature review of related literature over the past twenty years on emotion recognition from KMT dynamics to find out existing emotion elicitation, recognition methods and their performance profiles.
- b) Raised and answered six research questions for advancing research practice.
- c) Identified key research gaps, challenges, and potential research directions for future research.

The remaining parts of this paper are organized as follows: section II describes the method of how to search and identify related articles; section III provides a critical analysis of existing work from three perspectives: emotion elicitation methods, recognition features, and classification methods; section IV discusses some key applications of KMT dynamics-based emotion recognition; section V summarizes key research gaps and challenges and discusses potential future research directions; the last section concludes this study.

II. METHODOLOGY

We used a systematic literature review method [49] to conduct this study. First, we identified six research questions (see Table 1) and used them as clues to generate literature search terms (See Fig. 1). The search string is built by some related key words or phrases and logical connectors.

Second, in order to answer these RQs, the WEB OF SCIENCE CORE COLLECTION was selected as the main data source to search relevant literatures since it includes the most trusted global citation databases in the world and provides comprehensive citation data for many different academic disciplines. In addition, GOOGLE SCHOLAR is the complementary database since it is the largest academic search

TS=((*"Emotion recognition"* OR *"emotion detection"* OR *"emotion classification"* OR *"emotion prediction"* OR *"emotion sensing"* OR *"affective computing"* OR *"emotion elicitation"*) AND (*"Keystroke dynamics"* OR *"keystroke pattern"* OR *"typing pattern"* OR *"keystroke biometric"* OR *"keyboard"* OR *"mouse"* OR *"mouse dynamic"* OR *"mouse movement"* OR *"touch screen"* OR *"touchscreen"*))

FIGURE 1. Search string for literature research where “TS” means “Topic”, “OR” and “AND” are logical connectors, and others are key words or phrases.

TABLE 1. Research questions.

No.	Questions
RQ 1:	What are the commonly used emotion elicitation methods and databases for emotion recognition?
RQ 2:	Which emotions could be recognized from KMT dynamics?
RQ 3:	What key features are most appropriate for recognizing different specific emotions?
RQ 4:	Which classification methods are most effective for different specific emotions?
RQ 5:	What are the application trends of emotion recognition from KMT dynamics?
RQ 6:	Which application contexts are of greatest concern?

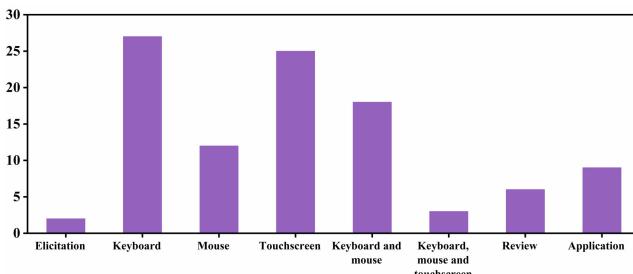


FIGURE 2. Number of articles related to emotion recognition based on their focuses.

engine. And about 20% cited articles in this paper come from GOOGLE SCHOLAR.

Third, we filtered some unrelated literatures by reviewing their abstracts. By using this search strategy, we totally found about 142 articles written in English. We did not set time and discipline restrictions because we would like to include related work as much as possible. Thus, indeed some of these articles are not related with emotion recognition from computer science perspective, so we further filtered out 40 articles by reading their abstracts and conclusions. As a result, we finally got 102 papers that have some relevance with emotion recognition from KMT dynamics. These papers are classified in Fig. 2 based on their research focuses. The research volume with a general increasing trend over the past 20 years is indicated in Fig. 3, and the trend of research about emotion recognition from single or combined KMT dynamics are illustrated in Fig. 4.

III. CRITICAL ANALYSIS

Computers are widely used in work and daily life, and the use of laptops and smartphones is also significantly increasing

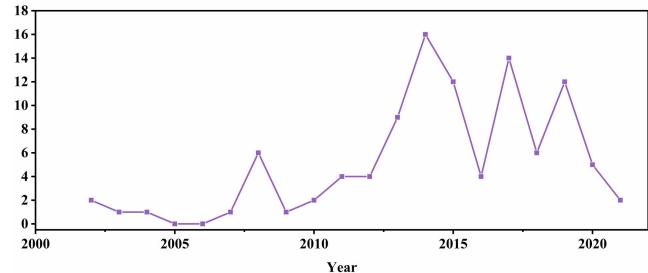


FIGURE 3. Trend of research volume over the past 20 years.

[41]. For example, about two billion people all over the world now are smart phones users [41]. Adult smartphone users frequently type on touchscreen for communicating with others, surfing the Internet or posting on social media every day [49]. These interactions with computers and smartphones provide rich secondary data sources derived from hand motions for emotion recognition [4], making it a hot topic to study the users’ emotional states [49] from KMT dynamics.

In order to recognize emotion, how to model emotions is a primary and essential problem. After a long time evolving of research, there are mainly two approaches to describe emotions: categorical approach and dimensional approach [50]. The categorical approach describes emotions in a discrete way, such as the seven basic emotional states: neutral, happy, sad, angry, surprised, scared, and disgusted [5]. The dimensional approach describes emotions from three dimensions: arousal, valence and dominance [44]. Arousal is defined as the energy of feelings such as low (sleepy) and high (excited); valence describes the level of an emotion from positive to negative; while dominance describes the feelings of control, ranging from lack of control to in control [45]. However, most studies just considered the valence and arousal parameters because integrating dominance is not useful when the accuracy of arousal is low [44].

In this section, we aim to answer our research questions through critically analyzing related articles. The key information for answering these questions is systematically summarized into Tables 6 to 7. Especially, Tables 6 and 7 provide detailed emotion recognition information based on statistical analysis and machine learning respectively. They contain multidimensional information from types of emotions, subject samples, features, recognition techniques, and performance profiles, which is helpful to make comparison from various dimensions. These two comprehensive tables together with other specific tables will answer concerned research questions. They are detailed below.

A. EMOTION ELICITATION METHODS

Emotion elicitation is usually a process that uses some intermediaries to help people arise certain emotional states [42], [52], [53]. It usually happens in the early stage of the whole emotion recognition experiment [52], [53], such as the to stimulate stress [4], [43], [54], [55]. The aim of the elicitation is to make sure that participants are in a certain predefined emotional state before collecting their behavior

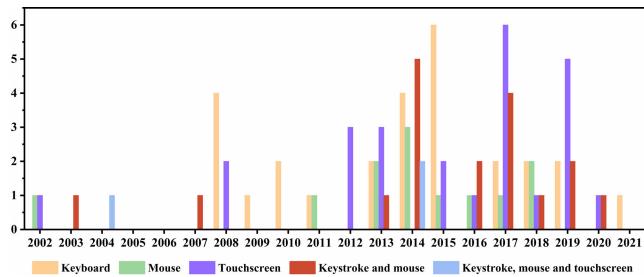


FIGURE 4. The research trend of emotion recognition from single or multiple KMT dynamics.

data [56]. In this case, some KMT dynamics data can be correlated to the certain emotional states and used as labeled data (ground truth) for training classification models [57]. Otherwise, due to the complex nature of human minds, it usually takes a long time to collect enough data corresponding to various emotional states [28]. Therefore, most studies considered stimulating participants' emotions by using pictures [3], [52], sounds [58], [59], and videos [45], [50], [52], [60]. In addition, there are also some other elicitation methods using interactive elicitation methods such as social interactions [42], tasks [50], and games [77] in given environments [55], etc. Table 2 summarized these elicitation methods and some existent databases in detail, which answer RQ 1.

Researchers in [61]–[74] built or introduced respective databases for emotion elicitation which are detailed in Table 2. Others either applied some of the databases or generated their own similar materials for their studies. The common stimuli methods and databases for emotion elicitation are discussed below.

The first commonly used stimuli method is picture-based including still pictures or a short video clip. The picture databases consisting of a set of different kinds of pictures are easy to access for researchers and effective for emotion elicitation. For example, researchers [3] selected 60 pictures from the IAPS database to elicit valence and arousal states, while a set of facial images from the POFA database [53] is utilized to elicit neutral emotional state and Ekman's six basic emotions (happy, sad, fearful, angry, surprised, and disgusted). Here, there is a potential problem that previous stimuli might have a long-lasting impact and then influence the next elicited emotion. In order to minimize the echo from the previous elicitation, some research [60] suggests a countdown (control) mechanism to make sure there is an interval between two adjacent stimuli. For example, some neutral images [60] selected from GAGED and EmoMadrid databases were used to make participants' heart rate return to normal after a positive or negative elicitation.

Short videos are the most used stimuli. It has been proved in both laboratory and field environments that short videos (7- to 11-minute long) can evoke a variety of longer lasting emotions [52]. Some researchers [52] selected short videos to elicit neutral, different level of valence and arousal combinations. Others used short videos [60] to induce discrete

emotional states like happy, sad, fearful, angry, surprised, lovely, hateful, exhausted, and disgusted. It is noticed that some individual video could induce more than one emotion [45], thus it is desired to choose some video clips which are more likely to induce one definite emotion. For example, the researchers [45] found nine 2-minute-long video clips for the purpose.

The second commonly used stimuli method is sound based. There are some sound databases for emotion stimuli. For example, the authors [58] selected 63 sounds from the IADS-2 database to elicit valence and arousal emotions. In the same vein, some researchers [59] used three different styles of music such as relaxation, rock and jazz music while others [75] used heavy metal music and recorded famous funny talks to provoke different levels of arousal.

The above induction approaches are easy to access, but compared with interactive elicitation methods [4], [102], their effects are short-lived because emotions induced by them such as short-videos would fade away along with the experiment.

The interactive elicitation methods like social interactions, tasks, and games were thus used widely to elicit emotions. In research [76], participants were supposed to listen to a set of predefined stories and immerse in these situations for neutral, angry, fearful, happy, sad, and surprised states respectively. The authors [77] believed that widespread touch-based computer games can elicit strong emotions in a short period of time. In addition, research [50] also found that different tasks especially those with clear introductions or longer expressive writings could better elicit specific emotions. The research [42] introduced four text conversations with different topics to trigger angry, happy, sad, surprised, and stressed respectively, during which participants were led to chat about a specific topic, reminisce their most enjoyable holiday experience, or be treated rudely.

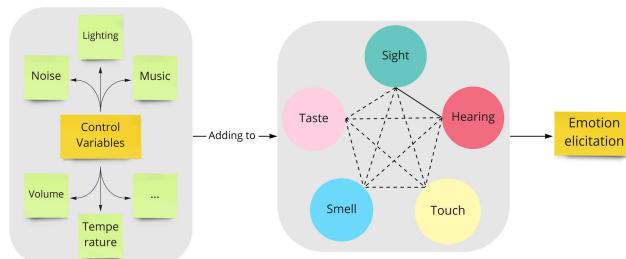
Especially, task-based elicitation method is more practicable to elicit stress emotion. Stress is a common emotion in current society which may impact people negatively in both work and daily life. Some researchers usually described stress as the reaction to exterior stimuli (environment), such as noises, injuries, coldness, and excessive demands [78]. Stress state is supposed to arise when participants were given limited time to complete a task [78]–[80] with various difficulty levels [54], or requested to do tough tasks under a noisy environment such as with loud traffic noise [55].

The Table 2 summarized these emotion elicitation methods, and provided some commonly used picture, sound, and video databases. So, the above information answers RQ1.

However, please note that the elicitation methods in Table 2 are mainly based on two human senses namely sight and hearing. In fact, human could interact with the world through five senses: sight, hearing, smell, touch and taste. Thus, theoretically people could have emotion elicitations from other three senses-based methods, and the combinations of all five senses-based methods. If we regard these five senses as five independent information input variables (channels), each of

TABLE 2. Emotion elicitation methods, databases, and emotion effects.

Stimuli methods	Databases or corresponding stimulus materials	Emotion Effects
Pictures [3, 60-65]	IAPS [61]: Multiple, from furniture to complete scenes, wide range of semantic categories.	Valence, arousal, and dominance
	GAPED [62]: Spiders, snakes, positive and neutral pictures, and scenes that induce emotions related to the violation of moral and legal norms.	Valence, arousal, external norm, and internal norm
	EmoMardrid [63]: Animals, objects, food, landscapes, and humans.	Valence and arousal
	EmoPics [64]: Diverse social situations, animals, and plants.	Valence, excited and complex
	ADFES [65]: Dynamic expressions.	Angry, disgusted, fearful, joyful, sad, and surprised
Sounds [58, 59, 66-70, 75]	POFA [53]: Facial images	Neutral, happy, sad, fearful, angry, surprised, disgusted
	IADS-E [66]: Animals, people, nature, sounds of daily routines, transport, electronic sounds, sound effects, breaking sounds, and music.	Valence, arousal, and dominance
	SEMAINE [67]: Annotated multimodal records of emotionally colored conversations between a person and a limited agent.	Valence, activation, powerful, anticipated/expected, and intensive
	eINTERFACE'05 [68]: Video sequences carrying the desired affective information.	Angry, disgusted, fearful, happy, sad, and surprised
	SEWA [69]: Audio-visual recordings of the subject's reaction to the 4 advertisement stimuli, and video-chat recordings of the subjects discussing the advertisement.	Valence, arousal, and liking/disliking
Videos [45, 52, 60, 70-74]	EMDB [70]: Film clips.	Valence, arousal, and dominance
	ICT-MMMO [71]: YouTube review videos.	Positive, negative, and neutral
	Belfast [72]: Mild to moderately strong emotionally colored naturalistic responses to a series of laboratory-based tasks.	Valence and intensive
	MOUD [73]: Product review videos in Spanish.	Positive, negative and neutral sentimental
	IEMOCAP [74]: Video, speech, motion capture of face, text transcriptions.	Angry, happy, sad, neutral; valence, activated, and dominance
Interactive methods	Tasks [50, 55, 76]	Angry, fearful, happy, sad, surprised, and stressed
	Games [77]	Excited, relaxed, frustrated, and bored
Environment	Social interactions (conversations) [42]	Angry, happy, sad, surprised, and stressed
	Noise [55]	Stressed

**FIGURE 5.** Further possible elicitation methods.

them could have controllable effects by changing its control variables, such as environment noises and lighting conditions, volumes of sounds or brightness of picture, etc. Fig. 5 shows the possible elicitation methods based on information input from the five independent senses and their control variables. It indicates a need for further investigation of effective emotion elicitation methods based on the five senses and their controls, and their combinations.

B. EMOTIONAL STATES THAT CAN BE IDENTIFIED FROM KMT DYNAMICS

Although some emotional states can be elicited using above elicitation methods, it does not mean these emotional states can be recognized from KMT dynamics. Thus, there is still a

need to find out what emotions can be recognized by KMT dynamics. Based on the reviewed literatures, the emotions they have recognized were summarized in the first columns of Tables 6 and 7. It is found that discrete emotional states like happy, sad, angry, surprised, fearful, and disgusted [76] can be recognized qualitatively from KMT dynamics. A few studies distinguished different levels of arousal and valence [78]. Some only classified emotions into positive and negative ones [97], and others even only focuses on one single emotion state such as stress, since it has been a main emotion that threatens people's physical and mental health under the increasing workplace competition [98].

More detailed emotions which can be recognized from KMT dynamics have been shown in first columns of Tables 6 and 7, providing the answer to RQ 2.

C. EXTRACTED KEY FEATURES

Computers and mobile phones are ubiquitous in people's everyday life and work, so it is feasible and low cost to collect adequate data from these devices [41], [49]. It is also unobtrusive and non-invasive to record users' KMT events during their normal use [37], [39]. After collecting these events data, it is necessary to extract suitable features for emotion recognition. The features are simply categorized into **signal features** and **user features** respectively in this paper.

The identified features from the two categories will provide the answer to RQ3.

1) SIGNAL FEATURES

Signal features in this context refer to KMT features in terms of keystroke features, mouse (usage) features and touchscreen (gesture touch) features respectively.

a: KEYSTROKE FEATURES

Keyboard interactions are usually expressed in forms of key press and key release [49]. Based on these two kinds of events and associated time information, keystroke features can be roughly grouped into timing related (K1~K6), frequency related (K7~K17), and other (K18~K22) features [49] (see Table 3). Timing related features mainly include duration, latency and typing speed [43], [49]. The duration refers to a time interval between a key press and release of the same key or some consecutive keys [43], [49]. Latency usually indicates the time between a key up and a key down events [43], [49]. Typing speed is usually defined as the total number of words/characters typed in a certain time unit [43], [49]. The frequency related features usually measure the frequencies of some specific keys, such as delete and backspace which are usually used to reflect typing error rate [54]. For a keyboard equipped with pressure sensors, the key pressure can also be collected and used as a feature to help predict emotions [38], [76]. In addition, some studies [45], [52], [59] also considered normalized features, such as the mean or standard deviation of the times using the enter key. In fact, not all these features are used in every research since not all the features have significant contribution to predicting some certain emotions [43].

The researchers [37] extracted K1, K2, K3, K21 (2D and 3D), K7, K10, K15, and K16 for emotion recognition and calculated mean and standard deviation for each feature. Besides keystroke features, they also used the content features derived from resultant typed text. Considering the happy and stressful states, the researchers [82] first just used four features: K7, K8, K19, and K20. Then, in order to enhance the recognition performance, they tested more features including K20 and the mean and standard deviation for K1 and K2. The study [83] extracted commonly used features like K1, K2, K7, K8, K13, K21 as well as calculated mean and standard deviations based on some most used Polish 2D and 3D words. As a result, it obtained a 36-dimensional feature vector for emotion recognition. The research [51] extracted 19 features from free text and fixed text typing patterns, which include not only the K1, K2, K3, and K6, but also the computed features such as the mode, standard deviation, standard variance, range, min, and max of the above features. In addition, it also collected free and fixed text content. However, not all these features were used to predict emotional states. Further, it iterated all combinations of above features and finally selected seven features (K6, mode and min of K1, K2, and K3) that worked best.

TABLE 3. Normally used keystroke features.

	Keystroke features	Explanation
Timing related	Duration (dwell time) (K1)	Interval between key press and release of the same key (1D) or two consecutive keys (digraphs or 2D) or three consecutive keys (trigraphs or 3D).
	Latency (flight time) (K2)	The time from a key release to a key down behaviors.
	Down-down time (K3)	The time from a key down to another key down.
	Up-up time (K4)	The time from a key release to another key release.
Frequency related	Time since last keypress (K5)	Time since last keypress
	Typing speed (K6)	Total number of words typed in a certain time unit.
	Delete/backspace keys (K7)	
	Enter key (K8)	
	Letter keys (K9)	
	Number keys (K10)	
	Arrow keys (K11)	
	Tab key (K12)	
	Spacebar (K13)	
	Function key (K14)	
Other features	Uppercase characters (K15)	Frequency of these keys. (For example, the delete or backspace keys can reflect error rate/correction rate).
	Punctuation marks (K16)	
	Special characters(K17)	
	Pressure value (K18)	Degree of pressure.
	Input rate (K19)	The ratio of input during the typing procedure.
Other features	Pause rate (K20)	The ratio of pause during the typing procedure.
	Number of total key events (K21)	The number of all key events.
	Typing amplitude (K22)	The maximum value in the range from 10 milliseconds before to 20 milliseconds after the typing instant.

Besides, some researchers [42], [50], [55], [76] thought emotion had impact on typing pressure, so they collected K18 from keyboards with pressure sensors. For example, researchers [76] computed the pressure sequence and other five features: mean value, standard deviation, the difference between max and min, the positive energy center, and the negative energy center from typing data. They also found that two traditional keystroke features: K2 and K3 are useful.

On the contrary, some researchers [3] and [58] intended to investigate the influence of emotion on the typing patterns. In their experiment, K1, K2 and K7 were extracted to find the influence of emotion on these features.

b: MOUSE FEATURES

Mouse is another common input device and mouse features are also effective in emotion recognition whether used separately or in combination with keystroke features [4], [84]. Depending on the types of mouse events, mouse features can be divided into three categories: click related features (M1~M6), movement related (M7~M19) features, and other (M20, M21) features [45], [85]. The most used mouse features are summarized in Table 4.

In [84], [85], researchers also tried out two mouse trajectory features namely attraction (the area contained in the straight line connected trajectory from the start to end positions) and M17. Some [86] extracted more mouse features from game interactions such as M1, M2, M3, M8, and M19. Others [87] utilized even more features including M7, M8, M9, M10, M11, M12, M15, M16, M18, M19, and M21.

Other research using mouse features listed in Table 4 to elicitate emotions are shown in Tables 6 and 7.

c: TOUCHSCREEN (FINGER-STROKE) FEATURES

In recent years when smart phones have become more and more popular among ordinary people, studying users' emotions from touchscreen patterns attracts more research attentions recently. When using touchscreen, people usually move cursor to certain position and then type using finger or stylus as what they do using traditional computer keyboard and mouse, so besides some features related to touchscreen itself (see Table 5), touchscreen features also include similar characteristics with both keystroke and mouse features.

In order to investigate the possibility of emotion recognition from touch-based devices, research [77] extracted T1, T14, and K1 from the mobile gaming interaction data. Then sixteen finger-stroke features were computed, consisting of average, median, max and min values of the T8, T10, T14, and T15. Then the research mapped features onto the specific emotional states through visual inspection. For example, the pressure feature shows a clearer separation between frustration and other emotions.

After using the information gain method, the researchers [96] identified 10 features having a relatively stronger association with emotions from 14 features. These features included traditional keystrokes such as K6~K11, K13, K14, K18 and environment features like location and weather, and other features like device shake count. Similarly, the research [89] used information gain method to rank the importance of each feature, and also used the timestamp of each tap event and the types of key input (alphanumeric keys, delete keys, et.al). In addition, it considered the other two features: working hour indicator and persistent emotion.

The BiAffect research project [41] collected data from participants when they used supplied mobile phones, whose standard keyboard was replaced by a customer keyboard. The collected data were then used to extract K1, K5, K7, K13, M18, T5, T6, and T7 features. In addition, it also investigated the relationship between the typing dynamics and time against each hour of a day and each day of a week.

Different from above research, the research [42] did not consider the duration features by simply regarding them as a part of single input. And as a result, touch events were aggregated and transformed into some two-dimensional heat maps for use.

d: COMBINED FEATURES

Some research used only one KMT dynamic feature, but most research used two or three combined features.

TABLE 4. Normally used mouse features.

	Mouse features	Explanation
Click related	Number of mouse single left clicks (M1)	The number of single left clicks in a defined time interval.
	Number of single right clicks (M2)	The number of single right clicks in a defined time interval.
	Number of double left clicks (M3)	The number of double Left clicks in a defined time interval.
	Duration of mouse clicks (M4)	Time between a mouse button press event and the consecutive mouse button release event.
	Click covered distance (M5)	The distance covered by the cursor while a mouse button was clicked.
	Click Euclidean distance (M6)	Euclidean distance between the coordinates where a mouse button was pressed and released.
	Number of movements (M7)	The number of mouse movements in a defined time interval.
	The distance of mouse movement (M8)	The straight-line distance between the beginning and the end of each movement.
	Duration of mouse movements (M9)	Time of one mouse movement event.
	The length of mouse racing line (M10)	The length of each movement trajectory.
Movement related	Length of pauses in mouse movement (M11)	The length of the pauses between mouse movements (a pause is defined variously by different researchers).
	Number of pauses in mouse movement (M12)	The number of pauses between mouse movements.
	Off-click movement (M13)	Move mouse with no mouse-button is pressed.
	On-click movement (M14)	Move mouse with mouse-button is pressed.
	Mouse inactivity duration (M15)	The total time of mouse pause.
Other features	Movement accumulated angle variation (M16)	The angle variation described by the cursor for every pair of consecutive cursor location compared to the angle described by the previous pair of cursor locations.
	Number of direction changes (M17)	The number of movement direction changes.
	Movement acceleration (M18)	The acceleration in each movement.
	Movement speed (M19)	The mouse speed in each movement.
	Mouse scroll (M20)	The scroll event of the mouse.
	Mouse position (M21)	The cursor position in x and y axis.

Some early research that studied emotions mainly concentrated on the usage of keystrokes and mouse movements. For example, the research [52] used various keystroke and mouse features, such as K1, K20, M1, M2, M3, M8, M11, and M19. The study in [79] defined a keyboard behavior dataset and a mouse behavior dataset respectively. The former consisted of K2, K6, and K7 and the latter contained M1, M2, M12, M15 and M19 features. The research [4] also utilized keystroke features and mouse features in two separate experiments. As a result, the keystroke features K1, K2, and K7 and the mouse features M16 and M19 were used and logged by a constant distance rather than a constant time interval (for example, 10-pixel length straight line). On the contrary, in every 5

TABLE 5. Normally used touchscreen features.

Touchscreen features	Explanation
T1	Touch position represented by x, y, and z coordinates
T2	The send button consisting of coordinates on the screen
T3	Altitude (in radians) of the stylus
T4	Azimuth (in radians) of the stylus
T5	Horizontal distance to last keypress
T6	Vertical distance to last keypress
T7	Distance from the immediate previous key
T8	Touch length
T9	Total angular path of azimuth angle
T10	Touch speed
T11	Speed of the altitude angle (from 25ms windows)
T12	Speed of the azimuth angle (from 25ms windows)
T13	Accelerometer movement
T14	Pressure area
T15	Directness index

seconds interval, [45] logged 17 attributes such as M1, M2, M3, M8, M20, K1, K2, K7, K8, K11, and K14.

As the development of smart mobile devices, a good amount of research combined touchscreen features with keystroke and mouse features into application. The research [90] used K1, K2, and K3 from keystroke dynamics, M8 and M10 from mouse dynamics, and T1, T10 and corresponding time. In addition, other computational features, such as the average, min, max, standard deviation, variance of the first and second derivatives of the above features were also considered.

More used features from literatures are shown in column 3 of Tables 6 and 7. To answer the RQ 3, K1, K2, K7, K18 and M2, M8, M16, M19 and T10, T14 features are found to be the most frequently used features. This may reflect in some degree that these features are more effective than others. Although some research ranked the importance of the features they used, the reason why some features overperform others and what feature groups work well for specific emotions are still not clear. Thus, more work is needed to better answer these questions.

2) USER FEATURES

User features refer to those related to individual differences. Therefore, user features can cause different patterns of using keyboard, mouse and touchscreen for different people [6], [43], [56]. The features like age, gender, culture context, computer proficiency, and some habits like left-handedness, right-handedness and the body postures when users are typing can be regarded as user features [13], [41].

Many applications usually neglect user features especially when they collected data in certain scenarios like in education [92], [94], teamwork [13], [95], and health care [41], [82], [93]. However, in real world settings, users have a variety of culture background, different proficiency of using input devices, and different hand using habits (left-handedness or right-handedness), which can impact the recognition result. For example, a person who uses computer every day usually has a higher typing speed and conversely a person who seldom uses computer types slowly. In this case,

if the user features are ignored, researchers may wrongly think the former is in the positive emotion and the latter is in the negative emotion [56]. Thus, user features are also important to make recognition results more credible.

Only a few research have considered the user features. In [78], participants were requested to provide their information like the age, gender, experience, and education level at first. These features were then used as control variables in different models to reflect their influences on results.

Some researchers [55] did not utilize user features directly in their experiments, but they had some restrictions in choosing participants to avoid the influence of individual user features. The restrictions include right-handedness, no color blindness, no tremors, no medication for hypertension or any other cardiovascular disease, and no history of musculoskeletal disorders. These restrictions make sure that participants have similar profiles and can avoid individual difference in a great degree. The study [53] focused on the elder's emotion detection based on smartphone keystroke dynamics, which has a potential to monitor pathology cases of the elderly people. As a high percent of the older people did not use smartphone, participants were thus required to meet the criteria regarding level of education, experience in using a smartphone, non-existence of any typing difficulties and vision problems, and non-existence of mental disorder. In addition, participants were also requested to hold a smartphone in one hand which aimed to eliminate the influence of posture on keystroke patterns.

User features have been considered as a factor that can influence the recognition result, but most research neither take this kind of features in their study nor investigate what user features and how these features influence the recognition performance. Thus, we cannot draw a conclusion whether user features are contributed to emotion recognition.

For answering our RQ 3, the main signal features are summarized in the third column of Tables 3, 4 and 5. They are most frequently used features. Other minor features are listed in column 3 of Tables 6 and 7.

D. RECOGNITION METHODS

The techniques for emotion recognition from KMT dynamics can be divided into two classes namely **statistical methods** and **machine learning methods** (including neural networks). This subsection aims to answer RQ 4.

1) STATISTICAL METHODS

Statistics is a widely used data analysis method. In the early years, many researchers used statistical methods to recognize emotional states. The difference between them is that they may define different threshold values to classify different emotions.

The SPSS analysis [78] is applied to compute the P-value and F-value between each variable to evaluate the correlation between keystroke dynamics and emotions. Similarly, correlations between learner's behavior B, mouse behavior B(M), and keystroke behavior B(K) are examined using the Pearson

TABLE 6. Statistical analysis-based emotion recognition methods with kmt dynamics.

Emotional states being recognized in the key Ref.	Subject samples	Features	Recognition Techniques	Performance / Result (Quantitative / Qualitative)
Stressed; Arousal, valence [78]	152 (23.68% F, average age 31.4)	K6, K7, K18	(1) Linear regression; (2) ANOVA; (3) Correlation analysis; (4) Coefficient of determination	(1) Excited stated has some negatively correlation with remaining task time. (2) A significant difference of the valence can be observed between the groups. (3) Negative emotion has some negatively correlation with typing speed, and positively correlation with typing error rate and pressure.
Stressed [79]	160	K2 (average), K6 (average), K7. M1, M2, M12, M15, M19 (average).	(1) ANOVA; (2) Spearman correlation; (3) Pearson correlation	(1) There are significant impacts of direct instruction and external stimuli on learner's motivation and affective state. (2) There are significant correlations between direct instruction, external stimuli, affective state, and cognitive state. (3) There are significant correlations between mouse behavior and keystroke behavior.
Stressed [50]	11	K18	(1) Wilcoxon signed rank test; (2) Spearman correlation	(1) Overall typing pressure baseline is negatively correlated with self-reported stress. (2) The stressed condition was associated with significant increases in self-reporting stress.
Positive, negative [97]	15 (10 M and 5 F, mean age 23.4)	K1, K2	T-test	The results prove the significance in the differences in the typing patterns under positive and negative emotions for all subjects.
Stressed [98]	20 students from University of Colombo School of Computing	(1) K1 (Selected 2D for letters: th, he, in, er, an) (2) K1 (Selected 3D for letters: the, and, ing) (3) K7.	Wilcoxon signed rank	90% confidence level for significant differences between stress and non-stress

correlation tests [79], showing that there is a great possibility to develop a cost-effective system to sense learners' emotion by using KMT dynamics. By computing Wilcoxon signed rank and Spearman's Rank-order correlation [50], it is found that the stressful condition is associated with self-reporting stress and there is a positive correlation between the self-reporting stress and typing pressure. The most used effective statistical methods are concluded in the fourth column of Table 6 along with other information like emotions, subject samples, and features, which make them more integrated and easier to make some comparisons.

2) MACHINE LEARNING METHODS

As the computer and AI technologies advance further, more and more research focus on machine learning methods (including Neural Network) in emotion recognition.

Many researchers tested only one machine learning algorithm to classify emotions. For example, researchers [37] aggregated the data and just used Decision Trees (DT) algorithm to identify emotions since DT is a simple and low-cost solution. However, they found that participants' responses were not distributed equally to all levels of each emotion. To solve this data skew, they applied the undersampling method. After classification, a 10-fold cross validation method and Kappa statistic were used to assess the model. The study [96] adopted a machine learning approach to analyze data using Weka. The extracted features were used by a Bayesian Network classifier since this classification method performed best among other machine learning classifiers such as Naïve Bayes, DT, or Neural Network in repetitive 10-fold cross validation experiments. SVM method was also used [90] to recognize emotion with KMT features

separately. After emotion recognition, three evaluation criteria namely recognition accuracy, false positive alarm, and computational time were applied to evaluate the algorithm performance. The study [82] predicted the happy and stressed states using a K-Nearest Neighbors (KNN) classifier in Weka.

Some researchers tested more than one machine learning algorithms. For example, the researchers [77] implemented three learning algorithms: Discriminant Analysis (DA), Artificial Neural Network (ANN) with back propagation, and Support Vector Machine (SVM), to build an automatically discriminating system for classifying four emotional states (excited, relaxed, frustrated, bored), two levels of arousal, and two levels of valence. While others applied different learning methods to different features in a combination way. In [51], both keystroke dynamics and text patterns were used. For keystroke dynamics features, they were inputted into Weka software to classify emotions, while for text patterns, Vector Space Model (VSM) was applied to the ISEAR dataset as text pattern classifier. Similarly, in [76], three learning methods were applied to: global features of pressure sequences, dynamic time warping, and traditional keystroke dynamics respectively.

The study [83] trained different classifiers such as DT, Neural Networks, KNN, Naïve Bayes, AdaBoost, Rotation Forest (RF), and Bayesian Networks for testing emotion recognition, and found that there is no single one classifier good enough for identifying all predefined emotions from all participants. And reversely, building an individual emotion classification model for each participant and for each emotional state was not feasible since data was usually not enough to train a model in this case. In addition, the research observed that keystroke rhythm was not only influenced by emotions,

but also typing devices. These may provide reasons for using the combination approach.

In [89], three models were tested: L2-regularized Logistic Regression (LR), SVM with Radial Basis Functions kernel, and RF using 10-fold cross validation. From the test result, the RF model generated the best classification. The study [4] first labeled the keyboard and mouse features in arousal and valence dimensions, and then tested five machine learning algorithms: LR, SVM, Nearest Neighbors, C4.5, and RF. Similarly, the study [45] exploited Bounded K-means Clustering, KNN method and Weka tools to classify ten discrete emotions.

Different from the above, some research attempted to combine and fuse signal features before feeding them into an emotion classifier. For example, the keypress features and accelerometer features were fused early [41] before feeding them into machine learning models. The reason for early fusion was that extra information may be found in the aligning process of features. Two feature fusion methods: EF-dropna and EF-fillna were applied with the observation that accelerometer values were dropped or filled. Different with other research, this study used CNNs and RNNs in combination since the kernels in CNNs could only capture the local features, and training time of RNNs is longer. So, combining these two methods and taking advantages of each method could give a better result. The researchers [42] proposed a semi-supervised classification pipeline for predicting affective states based on touch data from smartphone. The classification network consisted of fully connected layers and was trained using the labeled data.

The most used machine learning methods are concluded in the fourth column of Table 7. To summarize, statistical methods are usually used to do some qualitative analysis, while machine learning methods are for quantitative analysis by evaluating their classification accuracies. As shown in the column 4 of Table 6, the statistical methods: ANOVA, Spearman correlation, and the Wilcoxon signed rank were more popular. Through these methods, the connection between emotions and KMT dynamics was studied, providing more experimental evidence for future work. From the column 4 of Table 7, the RF, KNN, and SVM algorithms were more commonly used. Most research introduced the accuracy as a criterion to evaluate their recognition performance. The accuracy of these research varied from about 50% to more than 90%. However, the size of subject samples and the emotions intended to be recognized are often different, so it is hard to say which algorithms are more effective than others.

Obviously, there is a trend that more and more research applied machine learning methods. However, there have been some other advanced deep learning techniques performed very well in other field, like Generative Adversarial Network (GAN), Deep Belief Network (DBN), Radial Basis Function Network (RBFN), etc., but whether these models could perform well in emotion recognition still need more future research.

IV. APPLICABILITY EVALUATION

Emotion can influence one's behavior and reflect one's health states to some extent. Thus, an intelligent system capable of recognizing and interfering one's emotions can perform better in the areas where emotion plays a crucial role, such as intelligent tutoring system, healthcare scenarios, design and team works, intelligent toys, lie detection and customer service [3], [51], [76]. To build such affective intelligent systems, awareness of individual's emotion is a basic requirement. In addition, measuring a system user's emotion without causing inconvenience to the user is preferred. With this concern, the researchers [90], [96] proposed to recognize emotions from KMT dynamics. The study [42] also realized that knowing the user's emotional states enables mobile phones to support more intelligent interactions, making the devices react/interact more appropriately, naturally, and friendly with users [76]. There are several application areas requiring emotion recognition research with KMT dynamics (See Table 8). According to Table 8, we will discuss some popular applications in detail and answer the RQ 5 and RQ 6.

A. HEALTHCARE

In healthcare area, emotion recognition from KMT dynamics has been applied in detecting Parkinson's disease and mood disorder. By predicting emotional states based on KMT dynamics, the research [82] showed a potential for exploiting biometric data in the healthcare-oriented domain. In [53], the idea of early identification of Parkinson's disease through behavioral data from smartphone was tested, both emotional and physical states were analyzed for monitoring ageing pathology cases like Parkinson's disease, based on the i-PROGNOSIS project. The research [104] also proved that using keystrokes under emotional stimuli is an effective and intelligent way to detect onset of Parkinson's disease. For predicting mood disorder through emotion prediction based on keyboard data, the research [41] proved that it is feasible and effective to predict presence and severity of mood disturbance, pointing to the potential of medical treatment for mood disorders.

B. EDUCATION

It has been proved that emotion has a significant effect on learning outcomes [105]. Keyboard and mouse are the most common input devices in learning context and collecting data from these devices is widely considered as non-intrusive, so using KMT dynamics features to investigate the learners' emotional states in educational scenario has been a simple and effective way [87], [92], [94]. In this field, the research [54] proposed to recognize emotions of learners who learnt English as a second language for improving learning performance. The study [79] also believed that learners' performance could be affected by their emotions. Thus, it is important to build a modern Intelligent ITS which is capable of recognizing learners' emotional states and adjusting learning content accordingly. While the emotion recognition

TABLE 7. Machine learning and neural network based emotion recognition methods with kmt dynamics.

Emotional states being recognized in the key Ref.	Subject samples	Features	Recognition Techniques	Performance / Result (Quantitative / Qualitative)
Happy, sad, surprised, angry, fearful, neutral [76]	50 (25 M and 25 F)	K2, K3 K18 (mean value, standard deviation, the difference between max and min, the Positive Energy Center (PEC), the Negative Energy Center (NEC)).	Nearest Neighbor classifier	Overall accuracy of 93.4% Average error rate: 6.6%
Angry, confident, excited, hesitant, nervous, relaxed, sad, tired [37]	12 (10 M and 2 F, aged 24-34)	(1) K1, K2, K3, (2D, 3D), K21 (2D, 3D) (3) K7 (mean and standard deviation) (4) K10, K15, K16	DT	Accuracy: 83.9% angry; 86.2% confident; 84.3% excited; 92.3% hesitant; 93.0% nervous; 78.9% relaxed; 93.8% sad; 85.1% tired.
Excited, relaxed, frustrated, bored; Arousal, valence [77]	15	Touchscreen: (1) T1, T14, and K1 (2) T8, T10, T14, and T15 (average, median, max and min values).	(1) Linear SVM (2) Kernel SVM	Accuracy: 77.0% for excited, relaxed, frustrated, bored (Liner SVM); 89.7% arousal (Liner SVM); 86% valence (Kernel SVM).
Frightened, sad, nervous, neutral [90]	50 (various cultural backgrounds)	(1) K1, K2, K3 (1D, 2D, 3D) (2) M8, M10 (Average, max, min, Standard deviation, variance, first order, second derivatives) (3) T1, T10	SVM	Accuracy: 93.20% frightened; 92.1% sad; 65% nervous; 85.1% neutral
Happy, stressed [82]	100	K1, K2, K7, K8, K19, K20	RIPPER	Accuracy: 71.3% stressed; 63% happy
Happy, sad, bored, angry, disgusted, surprised, fearful [83]	9 (7 M, 2 F)	(1) K1, K2, (2D, 3D) (means, standard deviations) (2) K7, K8, K13 (means, standard deviations) (3) K21.	(1) AdaBoost; (2) DT; (3) KNN; (4) Naïve Bayes.	Accuracy: AdaBoost: 81.25% fearful; 63.49% happy; DT: 76.67% angry; 61.11% bored; KNN: 64.58% sad; Naïve Bayes: 55% disgusted; 52.6% surprised. F-scores: RF: 0.79 active; 0.64 negative KNN: 0.95 arousal; 0.93 valence.
Negative, active; arousal, valence [4]	64 (29 M and 35 F, aged 12-51)	(1) K1, K2 (2D, 3D), K7 (2) M16 (standard deviation), M19 (average)	(1) KNN (2) RF	(1) KNN (2) RF
Happy, inspired, sympathetic, disgusted, fearful [45]	35 (28 M, 7 F)	(1) M1, M2, M3, M8, M20 (average, 1st derivative, 2nd derivative, 3rd derivative); (2) K1, K2, K7, K8, K11, K14 (average, 1st derivative, 2nd derivative, 3rd derivative);	RF	Accuracy: 69% (using 3rd derivative)
Depressive, manic symptoms [41]	40	Touchscreen: (1) K1, K5, K7, K13; M18; T5-T7	dpMood (stacking CNN with RNN)	RMSE 2.376 ± 1.605
Valence, arousal, dominance; Stressed; Happy, sad, angry, surprised [42]	70 (35 F, aged 18-31)	(1) Two-dimensional heat maps cover the touch screen keyboard region; (2) T2; (3) The timestamp in milliseconds.	Semi-supervised classification	Accuracy: 67% valence; 63% arousal; 65% dominance; 87% (0.84 AUC) angry; 81% (0.88 AUC) happy; 84% (0.87 AUC) sad; 84% (0.76 AUC) surprised; 92% (0.80 AUC) stressed.
Valence, arousal [56]	8 healthy adult males (aged 23.6 ± 0.32)	(1) K1, K2, K3, K4 (mode, median, mean, first quartile, third quartile, standard deviation, variance, median absolute deviation (MAD), skewness, and kurtosis); (2) K7, K8, K13, K21, K22; (3) Vibration frequency (first, second); (1) K1, K2, K7 (2) Sleep quality, energy, mobility/movement, and heart pulse rates	SVM	Accuracy: valence: 69.8%; arousal: 71.1%
Positive, negative [28]	15 (aged 25-35)	Touchscreen: (1) T3, T4, T9, T11(min), T12(min) (2) M8, M19 (Max, Min, standard deviation) (3) K1, K2, K18 (Standard deviation, average) (4) Number of samples in the stroke (1) K1, K5, K7, K11, K17 (2) Generalizing term rate (3) Average word length (4) Lexical diversity (5) Content diversity (6) Adjective, Noun, Intensity, Positive affect, third person pronoun word rate	DT	Accuracy: average 77.87%
Valence, arousal [81]	13 (5 M and 8 F, aged 19-40)	Touchscreen: (1) T3, T4, T9, T11(min), T12(min) (2) M8, M19 (Max, Min, standard deviation) (3) K1, K2, K18 (Standard deviation, average) (4) Number of samples in the stroke (1) K1, K5, K7, K11, K17 (2) Generalizing term rate (3) Average word length (4) Lexical diversity (5) Content diversity (6) Adjective, Noun, Intensity, Positive affect, third person pronoun word rate	SVC	Accuracy: 70%
Physical stressed, Cognitive stressed [93]	24 (aged 18-56)	(1) K1 (2D, 3D), M4, M9 (Mean, Standard deviation) (mean, Standard Deviation);	(1) ANN (2) KNN	Accuracy: ANN: 62.5% physical stressed; KNN: 75% cognitive stressed.
Engaged, relaxed, distracted,	55 (aged 18-30)	(1) K1 (2D, 3D), M4, M9 (Mean, Standard deviation) (mean, Standard Deviation);	(1) J-48 tree (2) KNN	Accuracy: J-48: 80.48% engaged;

TABLE 7. (Continued.) Machine learning and neural network based emotion recognition methods with kmt dynamics.

delighted, surprised, frustrated, flow, neutral [86]					
Angry, happy, sad, neutral [100]	3	Touchscreen: (1) K2, K7 (2) M18 (average) (1) K7, K17	J-48 tree	Accuracy: 90%	
Happy, sad, stressed, relaxed [101]	120 (aged 18-50)	(2) Mean Session ITD (MSI) (3) Refined Mean Session ITD (RMSI) (4) Session duration (5) Session text length	RF	Average accuracy (AUCROC) of 73% (maximum of 94%)	
Confused, frustrated, shameful, content, flow [102]	282 (aged 12-52)	M8, M19 (average), M16 (Average, standard deviation)	RF	Accuracy: 93.8%	
Valence, arousal [80]	75	(1) K1, K2 (single keystroke, 2D, 3D) (average), K7, K9, K11, K12, K21; (2) M1, M2, M3, M4, M5, M6, M8; (3) the difference between M5 and M6	RF	Accuracy: 59%	
Pleasant, unpleasant [87]	14	(1) M7, M8, M9, M10, M11, M12, M15, M16, M18, M19, (Average, max); M21	KNN	Average correct classification rate (CCR): $68.9\% \pm 0.9\%$	
Anxious [85]	234 (97 M, 137 F) for feature selection; 133 (58M, 75 F) for evaluation experiment	(1) M8, M9, M16, M19 (Mean, standard deviation, skewness, and kurtosis) (2) Inception, end time, decision speed, (3) attraction+, attraction#, x-overshoot, y-overshoot	SVR	Distributions of temporal features (e.g., velocity) as well as spatial characteristics (e.g., direction change) are indicative of users' anxious state.	
Irritated, annoyed, reflective, neutral [103]	50	Mouse and touchscreen: (1) M13, M14 (sum, max, min, means, standard deviation, variance); (2) K18	SVM	Accuracy: 83.2%.	

DT = Decision Trees, SVM = Support Vector Machine, KNN = K-Nearest Neighbors, RF=Random Forest, CNN= Convolutional Neural Network, RNN= Recurrent Neural Network, SVC = Support Vector Classifier, ANN=Artificial Neural Network, MTL-NN = Multi-task Learning based Neural Network, SVR=Support Vector Regression.

from KMT dynamics can make the ITS systems lower-cost and easier-to-use.

C. TEAMWORK

As team works have become a popular work model in modern society, it is important to study the impact of emotions on teamwork. Researchers increasingly realized that emotions are not only intrinsic to human experience but also inherent to the situation where people interact with others or environment [2], [7], [9], [12]. Thus, studying emotions in teamwork settings has been a potential area [8]. Cambria [106] considered that recognizing emotions can improve communications among human workers which is important to define products and services. Specifically, the researchers [12], [13] focused on design teams and attempted to study the relationship between interaction dynamics and performance of engineering design teams. As a result, they proved that emotions are critical for both successful romantic relationships and high-performance teamwork.

D. USER AUTHENTICATION

In fact, keystroke dynamics was studied, widely used, and rapidly developed as a nature choice and complementary method in user authentication over the last 50 years [48],

[107]–[109]. This can be very practical because it has been proved that each user has a unique style when they type through keyboard [109]. More specifically, the time interval of successive keystroke, keystroke duration, key pressure or some other keystroke factors can be used to build up a unique signature for a user [108]. The study [38] proposed a biologic verification method based on pressure sensor equipped keyboards. Pressure sequences and other traditional keystroke features were implemented respectively to reinforce the usual password authentication. By conducting two experiments, the research [110] found that both keystroke and mouse dynamic features can measure some traits of personality. In order to continuously authenticate users after login authorization, the study [48] proposed a user-adaptive feature extraction method to collect keystroke dynamics from free typing text.

Furthermore, as the increasing spread of touchscreens on mobile phones, the study [40] focused on using touchscreen data for this purpose. It performed a large-scale study which showed that a specific user can be identified in 5 users with a relatively high precision. From the result, the conclusion has been made that touchscreen input patterns are distinguishing for each person. Also based on mobile touchscreen, the study [39] identified users almost immediately based on the way they performed while maintained users' convenience.

TABLE 8. Applications of emotion recognition based on kmt dynamics.

Application areas	Number of articles
Intelligent game controlling	3
Intelligent Tutoring System (ITS)	3
Intelligent systems	2
Human-Computer Interaction (HCI).	2
Healthcare scenarios	3
Authentication	2
Computer-mediated communication (CMC)	1
Perceptual computing	1
The Microsoft Word based scenario	1
Artistic activities	1

In this section, the Table 8 and the detailed analysis answered the RQs 5 and 6.

V. GAPS AND FUTURE WORK

Through this systematic analysis of the related work in emotion recognition based on KMT dynamics, some research gaps are identified. In this section, we discuss the following key research gaps and potential future work.

A. BUILDING A COMMON DATASET

Many researchers have investigated emotion recognition from KMT dynamics, and some of them reached a high accuracy, but it is meaningless to compare each of them since there is no common dataset as benchmark for comparison [49], [93]. This need was identified in 2009 [93] but still not be met yet [49]. The datasets used in literatures are usually built by individual research groups and the features contained in different datasets are either diverse or organized in different formations [104], making them difficult for using as benchmark. In addition, emotional states that have been recognized are also very diverse. For example, some researchers cared about happiness and sadness [76] while others about arousal and valence [78]. Thus, it is necessary to build a common dataset in the future for the purpose of comparing different algorithms. As for what features should be contained in this dataset, further work is needed to identify which features play a more significant role in KMT dynamics-based emotion recognition.

B. CONSIDERING USER FEATURES

User features such as age, gender, culture context, computer proficiency, some habits like left-handedness or even the body postures may also have impact on typing behavior [6], [78]. Ignoring these features may result in a wrong classification. However, most existing studies did not consider user features when building their own dataset. Besides, other factors come from the devices (e.g. the keyboard layout, the distance between keyboard and hands) are also important factors that may change participants' typing habits. Thus, using these elements as a supplement of KMT features may improve the recognition accuracy. In future work, more studies are needed to explore how to incorporate users features into the emotion recognition.

C. ASSESSING RECOGNITION PERFORMANCE USING BIOLOGICAL SIGNALS

Almost all the above research relies on self-reporting for the assessment of emotion recognition. In these studies, participants were requested to record their emotional states in the form of self-defined questionnaire or Self-Assessment-Manikin (SAM) questionnaire [92]. This method forces participants to break their typing behavior, which defeats the purpose of using keystroke dynamics as a non-intrusive method to recognize emotions. On the other hand, it is hard to record participants' every emotional state accurately, especially when the limited options of questionnaire do not contain their subtle emotions, or they are even not sure how to describe their emotions [6]. The biological signals are much more objective and precise in terms of reflecting people's emotional states [6], [111], [112]. For example, for those who do not like to show their emotions, recognizing their emotions through appearance or behavior habits would be deceptive. On the contrary, biological features like EEG signals could still reflect real emotion changes. Therefore, in the future, using biological signals as ground truth in parallel to the KMT dynamics should be a potential method that improves training data quality and help researchers evaluate their algorithm performance more objectively and accurately.

D. IMPROVING PRIVACY AND SAFETY THROUGH FEDERATED LEARNING

Another challenge is how to improve the safety of the data and protect the privacy of users. At present, almost all research collected rich data including personal information, and then sent those data to a server or cloud storage devices. This transmitting increases the risk of privacy information loss or theft. Thus, how to ensure the security of these private data is a primary problem that each researcher should consider. However, it is inadequately addressed in most past research. As the development of hardware performance, the computers, mobile phones and wearable devices have growing computational power, so it is increasingly attractive to store and compute emotion related data locally [113]. Federated learning is a machine learning setting which aims to train a model in the case that training data are stored in different remote clients. Under the federated learning concept, the update of the model is realized by aggregating all the updates of each device [114]. In future work, applying federated learning technology into affective computing will be an innovative method for privacy protection.

E. RECOGNIZING EMOTIONS THROUGH DIGITAL WRITING

It is a trend that touchscreen is used in smart devices with finger-writing capability. Thus, writing will no longer be restricted to paper and computer keyboard. On the other hand, the keyboard layout is designed for inputting English letters, for people whose native language is not English, they may find it is not easy to use, instead, writing on a

touchscreen may be much easier than typing on a normal keyboard. Therefore, in the future, writing on a touchscreen may give users similar experience as writing in paper with a pen or brush [115], [116]. For example, combining the traditional calligraphy with modern digital art provides a new way to create, present and preserve calligraphy works. It is also a hot topic to recognize emotions from handwriting on touchscreen [81], [115]. For example, the research [117] proposed an interactive system that can reflect calligraphers' emotions at the same time when they are writing. Up to now, few research combines KMT dynamics with calligraphy for emotion recognition. However, there is a real situation that some people are good at writing in their native language, but not good at inputting them through keyboard typing (for example, writing Chinese characters by hands rather than typing letters through keyboard). In this case, changing typing input model into writing model can be a convenient solution for most smart phone users. Thus, learning users' emotions when they are writing in a touchscreen through stylus or finger rather than typing by a touchscreen keyboard is a potential research direction.

VI. CONCLUSION

In this paper, we have systematically reviewed emotion recognition based on the KMT dynamics along the dimensions of emotion elicitation, features, recognition methods, and applications. Based on these dimensions, we have answered six research questions which most researchers want to ask, identified some key related application fields, current research gaps, and future research directions.

REFERENCES

- [1] K. R. Scherer, "What are emotions? And how can they be measured?" *Social Sci. Inf.*, vol. 44, no. 4, pp. 695–729, 2005.
- [2] V. Watzek, V. Anselmann, and R. H. Mulder, "Team learning and emotions during teamwork: A qualitative study," *Res. Papers Educ.*, vol. 34, no. 6, pp. 769–789, Nov. 2019.
- [3] P.-M. Lee, W.-H. Tsui, and T.-C. Hsiao, "The influence of emotion on keyboard typing: An experimental study using visual stimuli," *Biomed. Eng. OnLine*, vol. 13, no. 1, pp. 1–12, Jun. 2014.
- [4] A. Pentel, "Emotions and user interactions with keyboard and mouse," in *Proc. 8th Int. Conf. Inf., Intell., Syst. Appl. (IISA)*, Aug. 2017, pp. 420–425.
- [5] Q. Yi, S. Xiong, B. Wang, and S. Yi, "Identification of trusted interactive behavior based on mouse behavior considering web user's emotions," *Int. J. Ind. Ergonom.*, vol. 76, Mar. 2020, Art. no. 102903.
- [6] P. Freihaut and A. S. Göritz, "Using the computer mouse for stress measurement—An empirical investigation and critical review," *Int. J. Hum.-Comput. Stud.*, vol. 145, Jan. 2021, Art. no. 102520.
- [7] J. Luca and P. Tarricone, "Does emotional intelligence affect successful teamwork?" in *Proc. 18th Annu. Conf. Australas. Soc. Comput. Learn. Tertiary Educ.*, Melbourne, VIC, Australia, 2001.
- [8] S. G. Barsade and D. E. Gibson, "Why does affect matter in organizations?" *Acad. Manage. Perspect.*, vol. 21, no. 1, pp. 36–59, Feb. 2007.
- [9] H. A. Elfenbein, J. T. Polzer, and N. Ambady, "Team emotion recognition accuracy and team performance," in *Functionality, Intentionality and Morality (Research on Emotion in Organizations)*, vol. 3. Emerald Group Publishing, 2007, ch. 4, pp. 87–119.
- [10] E. Brynjolfsson, J. J. Horton, A. Ozimek, D. Rock, G. Sharma, and H.-Y. TuYe, "COVID-19 and remote work: An early look at US data," *Nat. Bur. Econ. Res., NBER Work. Paper* 27344, 2020.
- [11] C. A. Collazos, H. Fardoun, D. Alsekait, C. S. Pereira, and F. Moreira, "Designing online platforms supporting emotions and awareness," *Electronics*, vol. 10, no. 3, p. 251, Jan. 2021.
- [12] M. F. Jung and L. J. Leifer, "A method to study affective dynamics and performance in engineering design teams," in *Proc. DS 18th Int. Conf. Eng. Design (ICED), Impacting Soc. Through Eng. Design*, vol. 7. Human Behaviour in Design, Lyngby/Copenhagen, Denmark, 2011, pp. 1–10.
- [13] M. Jung, J. Chong, and L. Leifer, "Group hedonic balance and pair programming performance: Affective interaction dynamics as indicators of performance," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, 2012, pp. 829–838.
- [14] R. A. Calvo and S. D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications," *IEEE Trans. Affective Comput.*, vol. 1, no. 1, pp. 18–37, Jan. 2010.
- [15] S. Mukhopadhyay and S. Sharma, "Real time facial expression and emotion recognition using Eigen faces, LBPH and Fisher algorithms," in *Proc. 10th Int. Conf. Cloud Comput., Data Sci. Eng. LBPH Fisher Algorithms Confluence*, Jan. 2020, pp. 212–220.
- [16] S. Jaiswal, S. Virmani, V. Sethi, K. De, and P. P. Roy, "An intelligent recommendation system using gaze and emotion detection," *Multimedia Tools Appl.*, vol. 78, no. 11, pp. 14231–14250, Jun. 2019.
- [17] A. Masson, G. Cazenave, J. Trombini, and M. Batt, "The current challenges of automatic recognition of facial expressions: A systematic review," *AI Commun.*, vol. 33, nos. 3–6, pp. 113–138, 2020.
- [18] X. Fu, C. Xue, X. Li, Y. Zhang, and T. Cai, "A review of body gesture based affective computing computer graphics," *J. Comput.-Aided Des.*, vol. 32, no. 7, pp. 1052–1061, 2020.
- [19] F. Noroozi, C. A. Corneau, D. Kaminska, T. Sapinski, S. Escalera, and G. Anbarjafari, "Survey on emotional body gesture recognition," *IEEE Trans. Affect. Comput.*, vol. 12, no. 2, pp. 505–523, Apr. 2021.
- [20] V. Narayanan, B. M. Manoghar, V. S. Dorbala, D. Manocha, and A. Bera, "ProxEmo: Gait-based emotion learning and multi-view proxemic fusion for socially-aware robot navigation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2020, pp. 8200–8207.
- [21] S. Halovic, C. Kroos, and C. Stevens, "Adaptation aftereffects influence the perception of specific emotions from walking gait," *Acta Psycholog.*, vol. 204, Mar. 2020, Art. no. 103026.
- [22] J. Z. Lim, J. Mountstephens, and J. Teo, "Emotion recognition using eye-tracking: Taxonomy, review and current challenges," *Sensors*, vol. 20, no. 8, p. 2384, Apr. 2020.
- [23] C. Hirt, M. Eckard, and A. Kunz, "Stress generation and non-intrusive measurement in virtual environments using eye tracking," *J. Ambient Intell. Humanized Comput.*, vol. 11, no. 12, pp. 5977–5989, Dec. 2020.
- [24] J. Marín-Morales, J. L. Higuera-Trujillo, A. Greco, J. Guixeres, C. Llinàres, E. P. Scilingo, M. Alcañiz, and G. Valenza, "Affective computing in virtual reality: Emotion recognition from brain and heartbeat dynamics using wearable sensors," *Sci. Rep.*, vol. 8, no. 1, pp. 1–15, Dec. 2018.
- [25] S. Saganowski, A. Dutkowiak, A. Dziadek, M. Dziezyc, J. Komoszynska, W. Michalska, A. Polak, M. Ujma, and P. Kazienko, "Emotion recognition using wearables: A systematic literature review—work-in-progress," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PerCom Workshops)*, Mar. 2020, pp. 1–6.
- [26] H. Lu, M. Wang, and A. K. Sangaiyah, "Human emotion recognition using an EEG cloud computing platform," *Mobile Netw. Appl.*, vol. 25, no. 3, pp. 1023–1032, Jun. 2020.
- [27] M. Egger, M. Ley, and S. Hanke, "Emotion recognition from physiological signal analysis: A review," *Electron. Notes Theor. Comput. Sci.*, vol. 343, pp. 35–55, May 2019.
- [28] A. R. Aslam, T. Iqbal, M. Aftab, W. Saadeh, and M. A. B. Altaf, "A10.13u/classification 2-channel deep neural network-based SoC for emotion detection of autistic children," in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, Mar. 2020, pp. 1–4.
- [29] Y. Yang, Q. Wu, M. Qiu, Y. Wang, and X. Chen, "Emotion recognition from multi-channel EEG through parallel convolutional recurrent neural network," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2018, pp. 793–799.
- [30] M. D. Basar, A. D. Duru, and A. Akan, "Emotional state detection based on common spatial patterns of EEG," *Signal, Image Video Process.*, vol. 14, no. 3, pp. 473–481, Apr. 2020.
- [31] A. R. Aslam and M. A. B. Altaf, "An on-chip processor for chronic neurological disorders assistance using negative affectivity classification," *IEEE Trans. Biomed. Circuits Syst.*, vol. 14, no. 4, pp. 838–851, Aug. 2020.
- [32] D. M. Schuller and B. W. Schuller, "A review on five recent and near-future developments in computational processing of emotion in the human voice," *Emotion Rev.*, vol. 13, no. 1, pp. 44–50, Jan. 2021.

- [33] R. A. Khalil, E. Jones, M. I. Babar, T. Jan, M. H. Zafar, and T. Alhussain, "Speech emotion recognition using deep learning techniques: A review," *IEEE Access*, vol. 7, pp. 117327–117345, 2019.
- [34] M. Swain, A. Routry, and P. Kabisatpathy, "Databases, features and classifiers for speech emotion recognition: A review," *Int. J. Speech Technol.*, vol. 21, no. 1, pp. 93–120, 2018.
- [35] A. Seyedtabari, N. Tabari, and W. Zadrozny, "Emotion detection in text: A review," 2018, *arXiv:1806.00674*.
- [36] K. Sailunaz, M. Dhalwal, J. Rokne, and R. Alhajj, "Emotion detection from text and speech: A survey," *Social Netw. Anal. Mining*, vol. 8, no. 1, pp. 1–26, Dec. 2018.
- [37] C. Epp, M. Lippold, and R. L. Mandryk, "Identifying emotional states using keystroke dynamics," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, May 2011, pp. 715–724.
- [38] H. Lv and W.-Y. Wang, "Biologic verification based on pressure sensor keyboards and classifier fusion techniques," *IEEE Trans. Consum. Electron.*, vol. 52, no. 3, pp. 1057–1063, Aug. 2006.
- [39] A. De Luca, A. Hang, F. Brudy, C. Lindner, and H. Hussmann, "Touch me once and I know it's you!: Implicit authentication based on touch screen patterns," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, May 2012, pp. 987–996.
- [40] S. M. Kolly, R. Wattenhofer, and S. Welten, "A personal touch: Recognizing users based on touch screen behavior," in *Proc. 3rd Int. Workshop Sens. Appl. Mobile Phones (PhoneSense)*, 2012, pp. 1–5.
- [41] H. Huang, B. Cao, P. S. Yu, C.-D. Wang, and A. D. Leow, "DpMood: Exploiting local and periodic typing dynamics for personalized mood prediction," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2018, pp. 157–166.
- [42] R. Wampfler, S. Klingler, B. Solenthaler, V. R. Schinazi, and M. Gross, "Affective state prediction based on semi-supervised learning from smartphone touch data," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, Apr. 2020, pp. 1–13.
- [43] A. Kolakowska, "A review of emotion recognition methods based on keystroke dynamics and mouse movements," in *Proc. 6th Int. Conf. Hum. Syst. Interact. (HSI)*, Jun. 2013, pp. 548–555.
- [44] K. Bakhtiyari, M. Taghavi, and H. Husain, "Hybrid affective computing—Keyboard, mouse and touch screen: From review to experiment," *Neural Comput. Appl.*, vol. 26, no. 6, pp. 1277–1296, Aug. 2015.
- [45] R. Shikder, S. Rahaman, F. Afroze, and A. B. M. A. Al Islam, "Keystroke/mouse usage based emotion detection and user identification," in *Proc. Int. Conf. Netw., Syst. Secur. (NSysS)*, Jan. 2017, pp. 96–104.
- [46] Y. Zhong and Y. Deng, "A survey on keystroke dynamics biometrics: Approaches, advances, and evaluations," in *Recent Advances in User Authentication Using Keystroke Dynamics Biometrics*. Science Gate Publishing, 2015, ch. 1, pp. 1–22.
- [47] S. Zhang and P. Hui, "A survey on mobile affective computing," 2014, *arXiv:1410.1648*.
- [48] K. Shekhawat and D. P. Bhatt, "Recent advances and applications of keystroke dynamics," in *Proc. Int. Conf. Comput. Intell. Knowl. Economy (ICCIKE)*, Dec. 2019, pp. 681–684.
- [49] A. Maalej and I. Kallel, "Does keystroke dynamics tell us about emotions? A systematic literature review and dataset construction," in *Proc. 16th Int. Conf. Intell. Environ. (IE)*, Jul. 2020, pp. 60–67.
- [50] M. Exposito, J. Hernandez, and R. W. Picard, "Affective keys: Towards unobtrusive stress sensing of smartphone users," in *Proc. 20th Int. Conf. Hum.-Comput. Interact. With Mobile Devices Services Adjunct*, Sep. 2018, pp. 139–145.
- [51] A. F. M. N. H. Nahin, J. M. Alam, H. Mahmud, and K. Hasan, "Identifying emotion by keystroke dynamics and text pattern analysis," *Behaviour Inf. Technol.*, vol. 33, no. 9, pp. 987–996, Sep. 2014.
- [52] P. Zimmermann, S. Guttormsen, B. Danuser, and P. Gomez, "Affective computing—A rationale for measuring mood with mouse and keyboard," *Int. J. Occupational Saf. Ergonom.*, vol. 9, no. 4, pp. 539–551, 2003.
- [53] S. Hadjimichaelou, D. Iakovakis, V. Charisis, S. B. Dias, J. A. Diniz, J. Mercier, and L. J. Hadjileontiadis, "On capturing older adults' smartphone keyboard interaction as a means for behavioral change under emotional stimuli within i-PROGNOSIS framework," in *Proc. Int. Conf. Universal Access Hum.-Comput. Interact.*, Jul. 2017, pp. 346–356.
- [54] S. Salmeron-Majadas, R. S. Baker, O. C. Santos, and J. G. Boticario, "A machine learning approach to leverage individual keyboard and mouse interaction behavior from multiple users in real-world learning scenarios," *IEEE Access*, vol. 6, pp. 39154–39179, 2018.
- [55] J. Hernandez, P. Paredes, A. Roseway, and M. Czerwinski, "Under pressure: Sensing stress of computer users," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, Apr. 2014, pp. 51–60.
- [56] K. Hasegawa, H. Miyamoto, Y. Ashida, Y. Ban, R. Fukui, M. Inazawa, and S. Warisawa, "Estimation of affective state based on keystroke and typing vibration during computer-mediated communication," in *Proc. 14th Int. Joint Conf. Biomed. Eng. Syst. Technol.*, 2021, pp. 235–242.
- [57] M. Z. Soroush, K. Maghooli, S. K. Setarehdan, and A. M. Nasrabadi, "A review on EEG signals based emotion recognition," *Int. Clin. Neurosci. J.*, vol. 4, no. 4, pp. 118–129, Oct. 2017.
- [58] P.-M. Lee, W.-H. Tsui, and T.-C. Hsiao, "The influence of emotion on keyboard typing: An experimental study using auditory stimuli," *PLoS ONE*, vol. 10, no. 6, Jun. 2015, Art. no. e0129056.
- [59] K. Bakhtiyari, M. Taghavi, M. Taghavi, and J. Bentahar, "Ambiance signal processing: A study on collaborative affective computing," in *Proc. 5th Int. Conf. Web Res. (ICWR)*, Apr. 2019, pp. 35–40.
- [60] A. K. Km, B. Kiran, B. Shreyas, and S. J. Victor, "A multimodal approach to detect user's emotion," *Proc. Comput. Sci.*, vol. 70, pp. 296–303, Jan. 2015.
- [61] M. M. Bradley and P. J. Lang, "International affective picture system (IAPS)," in *Encyclopedia of Personality and Individual Differences*, V. Zeigler-Hill and T. Shackelford, Eds. London, U.K.: Springer, 2017, pp. 1–4.
- [62] E. S. Dan-Glauser and K. R. Scherer, "The Geneva affective picture database (GAPED): A new 730-picture database focusing on valence and normative significance," *Behav. Res. Methods*, vol. 43, no. 2, pp. 468–477, Jun. 2011.
- [63] L. Carretié, M. Tapia, S. López-Martín, and J. Albert, "EmoMadrid: An emotional pictures database for affect research," *Motivat. Emotion*, vol. 43, no. 6, pp. 929–939, Dec. 2019.
- [64] M. Wessa, P. Kanske, P. Neumeister, K. Bode, J. Heissler, and S. Schönfelder, "EmoPics: Subjective and psychophysiological evaluations of new imaging material for clinical-bio-psychological research," *J. Clin. Psychol. Psychotherapy, Suppl.*, vol. 1, pp. 11–77, Jan. 2010.
- [65] J. van der Schalk, S. T. Hawk, A. H. Fischer, and B. J. Doosje, "Moving faces, looking places: Validation of the Amsterdam dynamic facial expression set (ADFES)," *Emotion*, vol. 11, no. 4, pp. 907–920, 2011.
- [66] W. L. Yang, K. Makita, T. Nakao, N. Kanayama, M. G. Machizawa, T. Sasaoka, A. Sugata, R. Kobayashi, R. Hiramoto, S. Yamawaki, M. Iwanaga, and M. Miyatani, "Affective auditory stimulus database: An expanded version of the international affective digitized sounds (IADS-E)," *Behav. Res. Methods*, vol. 50, no. 4, pp. 1415–1429, 2018.
- [67] G. McKeown, M. Valstar, R. Cowie, M. Pantic, and M. Schröder, "The SEMAINE database: Annotated multimodal records of emotionally colored conversations between a person and a limited agent," *IEEE Trans. Affective Comput.*, vol. 3, no. 1, pp. 5–17, Jan./Mar. 2012.
- [68] O. Martin, I. Kotsia, B. Macq, and I. Pitas, "The eINTERFACE'05 audio-visual emotion database," in *Proc. 22nd Int. Conf. Data Eng. Workshops (ICDEW)*, 2006, p. 8.
- [69] J. Kossaifi, R. Walecki, Y. Panagakis, J. Shen, M. Schmitt, F. Ringeval, J. Han, V. Pandit, A. Toisoul, B. Schuller, K. Star, E. Hajiyev, and M. Pantic, "SEWA DB: A rich database for audio-visual emotion and sentiment research in the wild," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 3, pp. 1022–1040, Mar. 2021.
- [70] S. Carvalho, J. Leite, S. Galdo-Álvarez, and Ó. F. Gonçalves, "The emotional movie database (EMDB): A self-report and psychophysiological study," *Appl. Psychophysiol. Biofeedback*, vol. 37, no. 4, pp. 279–294, 2012.
- [71] M. Wöllmer, F. Weninger, T. Knaup, B. Schuller, C. Sun, K. Sagae, and L. P. Morency, "YouTube movie reviews: In, cross, and open-domain sentiment analysis in an audiovisual context," *IEEE Intell. Syst.*, vol. 28, no. 3, pp. 46–53, May/Jun. 2013.
- [72] I. Sneddon, M. McRorie, G. McKeown, and J. Hanratty, "The Belfast induced natural emotion database," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 32–41, Jan./Mar. 2012.
- [73] V. Pérez-Rosas, R. Mihalcea, and L.-P. Morency, "Utterance-level multimodal sentiment analysis," in *Proc. 51st Annu. Meeting Assoc. Comput. Linguistics*, 2013, pp. 973–982.
- [74] C. Busso, M. Bulut, C.-C. Lee, A. Kazemzadeh, E. Mower, S. Kim, J. N. Chang, S. Lee, and S. S. Narayanan, "IEMOCAP: Interactive emotional dyadic motion capture database," *Lang. Resour. Eval.*, vol. 42, no. 4, pp. 335–359, 2008.

- [75] J. Li, G. Ngai, H. V. Leong, and S. C. Chan, "Multimodal human attention detection for reading from facial expression, eye gaze, and mouse dynamics," *ACM SIGAPP Appl. Comput. Rev.*, vol. 16, no. 3, pp. 37–49, Sep. 2016.
- [76] H. R. Lv, Z. L. Lin, W. J. Yin, and J. Dong, "Emotion recognition based on pressure sensor keyboards," in *Proc. IEEE Int. Conf. Multimedia Expo.*, Jun. 2008, pp. 1089–1092.
- [77] Y. Gao, N. Bianchi-Berthouze, and H. Meng, "What does touch tell us about emotions in touchscreen-based gameplay?" *ACM Trans. Comput.-Hum. Interact.*, vol. 19, no. 4, pp. 1–30, Dec. 2012.
- [78] M. Trojahn, F. Arndt, M. Weinmann, and F. Ortmeier, "Emotion recognition through keystroke dynamics on touchscreen keyboards," in *Proc. 15th Int. Conf. Enterprise Inf. Syst.*, 2013, pp. 31–37.
- [79] Y. M. Lim, A. Ayesh, and M. Stacey, "Exploring direct learning instruction and external stimuli effects on learner's states and mouse/keystroke behaviours," in *Proc. 4th Int. Conf. User Sci. Eng. (i-USER)*, Aug. 2016, pp. 161–166.
- [80] S. Salmeron-Majadas, O. C. Santos, and J. G. Boticario, "An evaluation of mouse and keyboard interaction indicators towards non-intrusive and low cost affective modeling in an educational context," *Proc. Comput. Sci.*, vol. 35, pp. 691–700, Jan. 2014.
- [81] J. Han, G. Chernyshov, D. Zheng, P. Gao, T. Narumi, K. Wolf, and K. Kunze, "Sentiment pen: Recognizing emotional context based on handwriting features," in *Proc. 10th Augmented Hum. Int. Conf.*, Mar. 2019, pp. 1–8.
- [82] M. Fairhurst, C. Li, and M. Erbilek, "Exploiting biometric measurements for prediction of emotional state: A preliminary study for healthcare applications using keystroke analysis," in *Proc. IEEE Workshop Biometric Meas. Syst. Secur. Med. Appl. (BIOMS)*, Oct. 2014, pp. 74–79.
- [83] A. Kolakowska, "Recognizing emotions on the basis of keystroke dynamics," in *Proc. 8th Int. Conf. Hum. Syst. Interact. (HSI)*, Jun. 2015, pp. 291–297.
- [84] T. Yamauchi and K. Xiao, "Reading emotion from mouse cursor motions: Affective computing approach," *Cognit. Sci.*, vol. 42, no. 3, pp. 771–819, Apr. 2018.
- [85] T. Yamauchi, "Mouse trajectories and state anxiety: Feature selection with random forest," in *Proc. Humaine Assoc. Conf. Affect. Comput. Intell. Interact.*, Sep. 2013, pp. 399–404.
- [86] M. J. C. Samonte, L. A. Vea, R. M. C. San Jose, A. J. V. Lagoy, C. E. P. Manlapid, and A. M. P. Perez, "An affect detector model for gamers on a role-playing game through mouse movements," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2018, pp. 489–494.
- [87] P. Lali, M. Naghizadeh, H. Nasrollahi, H. Moradi, and M. S. Mirian, "Your mouse can tell about your emotions," in *Proc. 4th Int. Conf. Comput. Knowl. Eng. (ICCKE)*, Oct. 2014, pp. 47–51.
- [88] S. Ghosh, S. Goenka, N. Ganguly, B. Mitra, and P. De, "Representation learning for emotion recognition from smartphone keyboard interactions," in *Proc. 8th Int. Conf. Affect. Comput. Intell. Interact. (ACII)*, Sep. 2019, pp. 704–710.
- [89] H. Ceker and S. Upadhyaya, "Transfer learning in long-text keystroke dynamics," in *Proc. IEEE Int. Conf. Identity, Secur. Behav. Anal. (ISBA)*, Feb. 2017, pp. 1–6.
- [90] K. Bakhtiyari, M. Taghavi, and H. Husain, "Implementation of emotional-aware computer systems using typical input devices," in *Proc. 6th Asian Conf. Intell. Inf. Database Syst. (ACIIDS)*, 2014, pp. 364–374.
- [91] A. Lunski and R. K. Moore, "Affective computing and collaborative networks: Towards emotion-aware interaction," in *Proc. Work. Conf. Virtual Enterprises*, 2008, pp. 315–322.
- [92] M. M. Bradley and P. J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential," *J. Behav. Therapy Experim. Psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.
- [93] L. M. Vizer, L. Zhou, and A. Sears, "Automated stress detection using keystroke and linguistic features: An exploratory study," *Int. J. Hum.-Comput. Stud.*, vol. 67, no. 10, pp. 870–886, Oct. 2009.
- [94] M. G. Valdez, J. J. Merelo, A. H. Aguilera, and A. M. Soto, "Mining of keystroke and mouse dynamics to increase the engagement of students with programming assignments," in *Proc. Int. Joint Conf. Comput. Intell.*, Nov. 2019, pp. 41–61.
- [95] M. F. Jung, "Engineering team performance and emotion: Affective interaction dynamics as indicators of design team performance," Ph.D. dissertation, Dept. Mech. Eng., Stanford University, Stanford, CA, USA, 2011.
- [96] H. Lee, Y. Sang Choi, S. Lee, and I. P. Park, "Towards unobtrusive emotion recognition for affective social communication," in *Proc. IEEE Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2012, pp. 260–264.
- [97] W.-H. Tsui, P. Lee, and T.-C. Hsiao, "The effect of emotion on keystroke: An experimental study using facial feedback hypothesis," in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2013, pp. 2870–2873.
- [98] S. D. W. Gunawardhane, P. M. De Silva, D. S. B. Kulathunga, and S. M. K. D. Arunatileka, "Non invasive human stress detection using key stroke dynamics and pattern variations," in *Proc. Int. Conf. Adv. ICT Emerg. Regions (ICTer)*, Dec. 2013, pp. 240–247.
- [99] S. Ghosh, K. Hiware, N. Ganguly, B. Mitra, and P. De, "Emotion detection from touch interactions during text entry on smartphones," *Int. J. Hum.-Comput. Stud.*, vol. 130, pp. 47–57, Oct. 2019.
- [100] I. Zualkernan, F. Aloul, S. Shapsough, A. Hesham, and Y. El-Khorzaty, "Emotion recognition using mobile phones," *Comput. Electr. Eng.*, vol. 60, pp. 1–13, May 2017.
- [101] S. Ghosh, N. Ganguly, B. Mitra, and P. De, "Evaluating effectiveness of smartphone typing as an indicator of user emotion," in *Proc. 7th Int. Conf. Affect. Comput. Intell. Interact. (ACII)*, Oct. 2017, pp. 146–151.
- [102] A. Pentel, "Employing think-aloud protocol to connect user emotions and mouse movements," in *Proc. 6th Int. Conf. Inf., Intell., Syst. Appl. (IISA)*, Jul. 2015, pp. 1–5.
- [103] B. Schuller, G. Rigoll, and M. Lang, "Emotion recognition in the manual interaction with graphical user interfaces," in *Proc. IEEE Int. Conf. Multimedia Expo. (ICME)*, Jun. 2004, pp. 1215–1218.
- [104] L. Giancardo, A. Sánchez-Ferro, T. Arroyo-Gallego, I. Butterworth, C. S. Mendoza, P. Montero, M. Matarazzo, J. A. Obeso, M. L. Gray, and R. S. J. Estépar, "Computer keyboard interaction as an indicator of early Parkinson's disease," *Sci. Rep.*, vol. 6, no. 1, pp. 1–10, Oct. 2016.
- [105] D. Putwain, P. Sander, and D. Larkin, "Academic self-efficacy in study-related skills and behaviours: Relations with learning-related emotions and academic success," *Brit. J. Educ. Psychol.*, vol. 83, no. 4, pp. 633–650, Dec. 2013.
- [106] E. Cambria, "Affective computing and sentiment analysis," *IEEE Intell. Syst.*, vol. 31, no. 2, pp. 102–107, Mar./Apr. 2016.
- [107] W. G. De Ru and J. H. P. Eloff, "Enhanced password authentication through fuzzy logic," *IEEE Expert*, vol. 12, no. 6, pp. 38–45, Nov. 1997.
- [108] F. Monroe and A. D. Rubin, "Keystroke dynamics as a biometric for authentication," *Future Gener. Comput. Syst.*, vol. 16, no. 4, pp. 351–359, Feb. 2000.
- [109] R. Joyce and G. Gupta, "Identity authentication based on keystroke latencies," *Commun. ACM*, vol. 33, no. 2, pp. 168–176, Feb. 1990.
- [110] I. A. Khan, W.-P. Brinkman, N. Fine, and R. M. Hierons, "Measuring personality from keyboard and mouse use," in *Proc. 15th Eur. Conf. Cognit. Ergonom. Ergonom. Cool Interact. (ECCE)*, Jan. 2008, pp. 1–8.
- [111] W.-C. Li, C.-J. Yang, and W.-C. Fang, "A real-time emotion recognition system based on an AI system-on-chip design," in *Proc. Int. SoC Design Conf. (ISOCC)*, Oct. 2020, pp. 29–30.
- [112] W.-C. Fang, K.-Y. Wang, N. Fahier, Y.-L. Ho, and Y.-D. Huang, "Development and validation of an EEG-based real-time emotion recognition system using edge AI computing platform with convolutional neural network System-on-Chip design," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 9, no. 4, pp. 645–657, Dec. 2019.
- [113] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," *IEEE Signal Process. Mag.*, vol. 37, no. 3, pp. 50–60, May 2020.
- [114] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," 2016, *arXiv:1610.05492*.
- [115] V. Zagorskis, A. Kapenieks, and A. Gorbunovs, "Emotions identification utilizing periodic handwriting on mobile surfaces," *Periodicals Eng. Natural Sci.*, vol. 7, no. 1, pp. 228–237, 2019.
- [116] S. V. Kedar, D. S. Bormane, A. Dhadwal, S. Alone, and R. Agarwal, "Automatic emotion recognition through handwriting analysis: A review," in *Proc. Int. Conf. Comput. Commun. Control Autom.*, Feb. 2015, pp. 811–816.
- [117] C.-M. Wang and Y.-C. Chen, "Design of an interactive mind calligraphy system by affective computing and visualization techniques for real-time reflections of the Writer's emotions," *Sensors*, vol. 20, no. 20, p. 5741, Oct. 2020.



LIYING YANG received the B.Eng. degree in computer science and technology from Hebei Normal University, China, in 2016, and the M.Sc. degree in software engineering from Northwest A&F University, China, in 2019. She is currently pursuing the Ph.D. degree in industrial design with the Northumbria School of Design. Her research interests include emotion recognition and regulation in design teamwork based on keystroke, mouse, and touchscreen dynamics.



SHENG-FENG QIN (Member, IEEE) joined the Northumbria School of Design, in 2014.

Prior to this appointment, he worked at the Department of Design, Brunel University, as a Lecturer, from September 2001 to August 2007, and a Senior Lecturer in product/industrial design, from September 2007 to December 2013. He was a Lecturer/Associate Professor at East China Jiaotong University, China, from May 1988 to August 1996. He was an Academic Visitor with the School

of Mechanical and Manufacturing Engineering, University of Birmingham, from September 1996 to September 1997. He was a Research Assistant with the School of Product Design, Cardiff Metropolitan University, from September 1998 to February 1999. He was a Postdoctoral Research Fellow at the University of Loughborough, from March 2000 to August 2001. He has published more than 200 papers in journals and conferences. His research interests include digital design and manufacturing, smart product and service design, and design for emotion.

Prof. Qin was a recipient of the 2019 Newton Prize based on his collaborative research work with Prof. Cuixia Ma at the Institute of Software, Chinese Academy of Sciences on “Transforming service design and big data technologies into sustainable urbanization.”

• • •