Inferring Emotions from Touching Patterns

Mojgan Hashemian
INESC-ID & Instituto Superior Tecnico
Universidade de Lisboa
Lisbon, Portugal
mojgan.hashemian@tecnico.ulisboa.pt

Rui Prada
INESC-ID & Instituto Superior Tecnico
Universidade de Lisboa
Lisbon, Portugal
rui.prada@tecnico.ulisboa.pt

Pedro A. Santos
INESC-ID & Instituto Superior Tecnico
Universidade de Lisboa
Lisbon, Portugal
pedro.santos@tecnico.ulisboa.pt

João Dias
INESC-ID & Instituto Superior Tecnico
Universidade de Lisboa
Lisbon, Portugal
joao.dias@gaips.inesc-id.pt

Samuel Mascarenhas
INESC-ID & Instituto Superior Tecnico
Universidade de Lisboa
Lisbon, Portugal
samuel.mascarenhas@gaips.inesc-id.pt

***Abstract*—In this paper, we propose a feature-based model to recognize emotions via touching patterns of individuals playing a game on a typical tablet. In this work, novel features, such as *Angular Velocity/Acceleration, Angle, Curl, Area and number of strokes within a time window*, are introduced and the gold-standard of the data is determined automatically via subjects' facial expressions. The results show that the approach is promising and the model is able to recognize all the six basic emotions, with a performance of 71.92%±0.51. In addition, the recognition of valence and arousal reaches correlation coefficients equal to 0.76 and 0.78 respectively.**

***Index Terms*—Touch, Stroke, Emotion, Arousal, Valence, Automatic Recognition**

I. Introduction

There is a growing body of literature that recognizes the importance of automatic emotion recognition [1]. With this aim, researchers have investigated a variety of approaches using different information modalities from different sources. For instance, from facial expressions (for an early review see [2]), via speech [3], or physiological cues [4] (e.g. from heart rate variability [5], skin conductance [6], EEG [7]), or using vision-based to infer emotions from gestures and posture (for a survey see [8]). Finally, the advent of wearable technology made emotion recognition easier, for instance by analyzing bio-physiological signals [9], or gestures [10].

Another source of information which gives clues about our emotions, as a subset of our bodily movements, is the way we interact with touch-screens [11]. In comparison with previous methods, this channel is cheap (no need for peripheral sensors), and is totally non-intrusive (comparing with physiological-based approaches [12]). Additionally, this channel is robust and is present in any tablet [11]. Moreover, it requires a lower computation cost (comparing to facial expression recognition). It is more applicable than a speech-based approach, since there is no need to force the users to talk. Also, comparing with vision-based approaches, users' privacy remains intact.

These advantages motivated several researchers to investigate the relationship between emotion and touch. Though to date, the research on affective touch has tended to mostly focus on haptic devices [13], few studies have investigated recognizing emotional states using strokes on touchscreens. Yet, these few attempts are still in the early stage and further studies are required to improve their effectiveness. This motivated us to further investigate emotion recognition using tactile features.

One challenge is providing the ground truth or the gold standard of data. In so doing, the most straightforward approach is using a questionnaire. However, by definition, emotion is volatile [14], hence memorizing it may not be precise, especially during long periods of time. Also, the subjects may label the whole duration with a single label (the ongoing or the recently perceived emotion). However, emotions may change in a course of few seconds, hence they might not be fixed over the whole duration. Hence, using a single questionnaire at the end of an interaction is not sufficient to capture emotional variances for the whole period. A solution might be frequently applying questionnaires; but, this would be highly intrusive. On the other hand, emotions are highly subjective [15], hence annotations might not be reliable in general [16], neither self-annotation, as human emotion is a cognitive concept and is hard to recognize the exact emotion and its intensity [14].

In this work, we apply a recent automatic method of emotion recognition, using Facial Expressions, to infer the ground truth, to overcome the difficulties of using questionnaires. It is showed that facial expression achieves high accuracy in real time emotion recognition [17]. For instance, FaceReader is capable of recognition with 88% accuracy, while the human emotions recognition rate for the same datasets was 85%. Hence, FaceReader is as good as human coders at recognizing emotions [18]. We believe that using automatic labeling of emotions resolves the implication of using self-reports. First, it does not interrupt the task and secondly, it works in a real-time manner, hence emotions will not be forgotten, and more interestingly, every single instant could be labeled. Although putting a camera to record facial expressions might be intrusive and violates the privacy, the intrusion level is lower than frequent questionnaires. One possible drawback could be not

978-1-7281-3888-6/19/$31.00 ©2019 IEEE

achieving 100% accuracy, but we argue that this is true for any other technique, even the self-reports. Also, Facial Expression Recognition tools have been validated using either human coders or self reports, hence these tools may suffer the same source of uncertainty. In the worst case, the performance of our tool is bounded by the performance of the automated emotion recognition tool for the ground truth.

Additionally, here we focus on "Basic emotions" [19]. Basic emotions are universal and this fact gives us the advantage to use the same tool on different ethnicity. One may argue that not all the basic emotions, such as disgust, might be observable during gameplay. Also, emotion is related to the context, hence using a limited set of emotions may lead to a higher classification rate within that specific context. However, here we focus mostly on the universality of the work by considering basic emotions. Although this selection may lead to potential noise of unseen data in modeling, we compensate this by providing a more reliable source of ground truth in form of a tested automatic recognition system.

II. Related Work

As with any multidisciplinary field, this work can be placed in different categories. Here, we only focus on recent literature which attempts to recognize emotions using tactile behavior. For instance, an early work [11], aimed to recognize four affective states: 'irritation, annoyance, reflectiveness and neutral' from 825 data-points. A feature vector composed of 15 selected features among 220, gained 70.12% performance discriminating the four emotions (the highest accuracy attained 92.7 for irritation, and the least achieved 43.3 for recognition of reflectiveness). The features carry information about Cartesian coordinates, and z as the applied pressure. In a similar vein, another study [17] with 50 users from different cultures, proposed a model that recognizes four classes of emotion (neutral, fright, sadness, nervousness). The average performance gained 76.0% (the min classification rate gained 55.3 regarding nervousness, and the maximum reached 90.0 for fright). Comparing to the previous study, this model seems more reliable having been trained and tested on a larger population.

In the work closest to our study [13], 15 people participated in an experiment and played 20 levels of a modified version of Fruit Ninja game on an iPhone. To gather the ground truth, the participants were asked to fill in a questionnaire and then relabel their emotions by watching the video recorded their faces during the gameplay. The self-report contained four affective states: delighted, excited, annoyed and frustrated. The responses were mapped to valence-arousal emotional space [20], and the corresponding values for arousal and valence were extracted. A feature-set containing 16 features, including the number of strokes in each session, {average, median, min and max} of {length, speed, directionality index, pressure} was extracted. The best trained model, gained an accuracy of 77.0% for recognition of the four classes. The maximum performance in arousal detection reached up to 89.7%, and for valence detection the best rate was 86.0%.

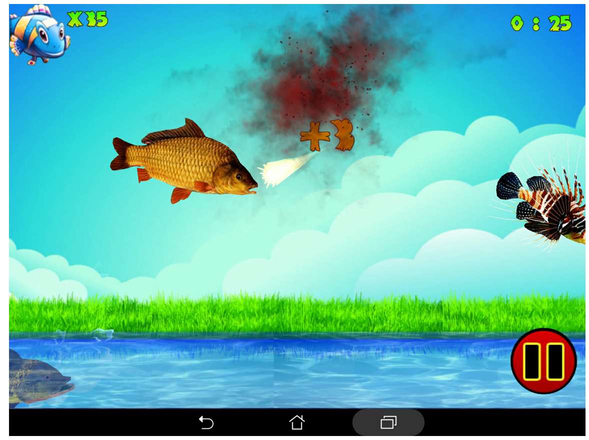

Fig. 1. Ninja Fish Game. In this picture a normal fish scores +3.

Other similar works have investigated the link between tactile behavior and other human factors such as effort [12] and more recently stress [21]. Although these studies do not directly relate to emotion recognition, the introduced features (startTime, stopTime, x, y points, acceleration, timeE-lapsed, distance, velocity, and relative changes in pressure) could be a reference for tactile-based approaches.

All the mentioned models used questionnaires to provide the ground truth; however, as discussed earlier, using a questionnaire is not a reliable way to measure emotions. Further, the aforementioned models could still be improved using more sophisticated features carrying more information about the dynamics of the performed stroke. Nevertheless, all together, these studies indicate that touching patterns consist of promising information for emotion recognition.

III. Data Collection

We sent an email to a random selection of the students including 320 people, stating that we will run a video-taped experiment for less than 10 minutes of playing a specific game on a tablet. No more information was added and we did not inform the subjects about the study goal. 25 people (16 male and 9 female) participated in the test voluntarily after filling the consent form. Participants' ages ranged between 19 and 45 (25.3 ± 6.1) years old. We ran the experiment during 4 days in an isolated room to prevent any distraction. We had to remove 5 cases from the dataset due to different sensor failures.

Preliminary results of our pilot studies with "Fruit Ninja" revealed that lots of people are excited during playing. Hence, we decided to ask the participants to play a game they never played and that entice a broad spectrum of emotions during gameplay. To choose the game, we looked at stroke-based games on GooglePlay, ranked by few people. We selected a game called "Ninja Fish" which takes at least 80 seconds (Figure 1). However, due to different events in the game, the duration might be different. We asked the subjects to play one level of the game, without telling them the game's rules to provide the opportunity to make them feel negative emotions such as boredom, anger, etc. There are three types of fishes in this game: slicing *Friendly* fishes has negative points, *Normal*

TABLE I
DISTRIBUTION OF STROKES REGARDING THE SIX BASIC EMOTIONS.

Emotion	Happy	Sad	Angry	Surprised	Scared	Disgusted
Count	654	369	126	33	20	155

fishes with normal score, and *Special* fishes which carry bonus points or freeze the time, giving the subject more time to catch more fishes. We expect to gather the basic emotions during the gameplay. For example, the first time a subject slices a friendly fish, s/he might get surprised or even angry by receiving unexpected negative points. Moreover, we may face a bored user who fails to remember the categories. Also, the more negative points might lead the player to lower arousal, or to experience more negative emotions, such as sadness. During the game, the user needs to remember the type of fishes and once s/he touches a friendly fish by mistake s/he might get disappointed, or sad, or even angry. Although one might argue that these factors are influenced by personality and game habits, but still the possibility to get negative emotions is high. Also, cutting a fish by fingers might be disgusting to some sensitive users, especially the ones with high level of imagination. The first explosion of fishes might arise scared emotion. Finally, normal fishes may make the player to be happy, also might lead to experiencing higher arousal.

The participants were asked to play on an 8-inch tablet. We fixed the tablet on the table to prevent potential movement noise. All the touch events were recorded using the ADB (Android Debugging Bridge). The subjects' interaction were videotaped by three cameras: one located on top of the table to record game events, one small LifeCam located in front of the user to record the subjects' facial expressions. Furthermore, another camera was placed next to the tablet to record the finger touching the screen for further analysis if necessary. To retrieve the gold-standard, we use FaceReader 6.0 which reports the intensity of emotions (happy, sad, angry, surprised, scared, disgusted) from 0 to 1, and extracts levels of valence (between -1 and 1) and arousal (between 0 to 1).

Each subject, performed different numbers of strokes on the tablet. A total number of 1357 strokes were recorded. Table I lists the distribution of the collected data in each emotion category. During each single stroke, the participant may show any of the 6 basic emotions in his/her face. To label each stroke, we use the most dominant emotion during the corresponding stroke interval. To be more specific, if the participant showed all the six emotions, the winner emotion is the one which had the highest intensity on the average. Similarly, we use the average of valence and arousal values recorded during that stroke as the ground truth for the corresponding models.

IV. METHOD

As discussed earlier, we argue that individuals experiencing different emotional states, show different patterns while working with a touchscreen. The justification behind this idea is that our motor behavior is highly influenced by the ongoing affective state [22]. Hence, the way we perform strokes carries information about our affective states. What follows is a description of our methodology to emotion recognition from touching patterns, more specifically dynamics of the performed strokes. The discriminative features we propose, which highlight the way the finger of the subject has moved on the screen, are categorized in 14 different categories as listed in Table II. In this table, X corresponds to the X coordination on the screen, Y corresponds to the Y coordination on the screen, P the pressure recorded, and T timestamp of the preceding recorded values. Note that the recorded pressure is the measurement reported by ADB which does not reflect the actual applied force, but rather an estimation based on the capacitance values obtained from the touchscreen sensor. To represent this continuous record, we define each stroke as follows: $(\{x_1, y_1, p_1, t_1\}, \{x_2, x_2, p_2, t_2\}, \ldots, \{x_i, y_i, p_i, t_i\}, \ldots, \{x_n, y_n, p_n, t_n\})$.

The major part of the selected features are in common with the recent literature. Other categories, i.e. Angular Velocity, Angular Acceleration, Angle, Curl, Area and number of strokes within a time window, are novel features introduced here. These new features are chosen to support the analysis of the dynamics of the touch movement. For instance, the 4th category, Positional Change (PC), examines the explicit changes in the position of the finger on the screen from three different metrics: 1. length, or Pythagorean Proposition, refers to the length of the traversed path considering backward moves (see Figure 2 (a)). 2. spanX/Y, refers to the maximum spanned path in direction of X/Y (See Figure 2(b)). 3. distance, considers only starting and the ending point (See Figure 2 (a)). To better examine such dynamics, we also considered the first and the second derivation of PCs in 5th and 6th category.

Other feature dealing with dynamics of movement are Velocity and Acceleration. To measure velocity, we divide the total displacement by duration of the stroke. Further, since the movement could be performed in non-straight line, we inspect the stroke by measuring its angular velocity and curl. Moreover, we calculate the first derivation of the curl to have an idea of the angular movement. To better examine the effectiveness of interval features (location and pressure), we consider descriptive statistics of each category (mean, minimum, maximum) as a single feature. Also, to better examine the number of performed strokes, we consider different sizes for a time window, i.e. 1, 3, 5 and 10 seconds long. These sizes were specified by multiple trial and error attempts. Regarding directional features, such as Positional Changes, we consider these numbers in the direction of the Cartesian axis (x and y), as well as the overall value. The amount of pressure that is applied can be influenced by affect; therefore, we expect to observe different level of pressure in higher intensities of anger comparing to less level of anger or even other emotions.

Additionally, we consider features that are contextual in terms of the actual gameplay. For instance, as the game reaches the end, the participant might be more aroused. Also, in this game, when a fish is getting out of the screen the chance to catching it decreases; hence the corner coordinates play a more tricky roles in games. The final feature set in Table II,

TABLE II
INITIAL FEATURE VECTOR CATEGORIZED IN 14 MAIN CATEGORIES.

Category	#	Abbreviation	Description
Time	1	Duration	Duration of the stroke
	2	Dist2prev	Time elapsed since the previous stroke
	3	timeElappsed	Time elapsed from the start of the interaction
Location	4-8	xMin,xMax,xMean,xMedian,xSTD	Min,Max,Mean,Median, StandardDeviation recorded X
	9-13	yMin,yMax,yMean,yMedian,ySTD	Min,Max,Mean,Median, StandardDeviation recorded Y
Pressure	14-18	pMin,pMax,pMean,pMedian,pSTD	Min,Max,Mean,Median, StandardDeviation recorded Pressure
Positional Changes (PC)	19	lengthT	sum of Pythagorean Propositions
	20	spanX	total span in direction of X
	21	spanY	total span in direction of Y
	22	distanceX	difference of X position from start to the end
	23	distanceY	difference of Y position from start to the end
	24	displacement	total displacement
PC first derivation	25	1stDVPCx	first derivation of X changes
	26	1stDVPCy	first derivation of Y changes
	27	1stDVPP	first derivation of the diameter of the Pythagorean Propositions
PC 2nd derivation	28	2ndDVPCx	second derivation of X changes
	29	2ndDVPCy	second derivation of Y changes
	30	2ndDVPP	second derivation of the diameter of the Pythagorean Propositions
Velocity	31	velocity	average speed
Acceleration	32	acceleration	average acceleration
Angular velocity	33	wj_Mean	Mean Angular velocity
	34	wj_Min	Minimum Angular velocity
	35	wj_Max	Maximum Angular velocity
Curl	36	curlX	Curve in direction of X
	37	curlY	Curve in direction of Y
First derivation of Curl	38	1stDVcurlx	average Angular Acceleration X (based on curve)
	39	1stDVcurly	average Angular Acceleration Y (based on curve)
Angle	40	angleM	the angle of the middle point of the stroke
	41	angle1	the angle between mid point and starting point
	42	angle2	the angle between mid point and the ending point
# of strokes	43-46	numOfStrkW1,3,5,10	# strokes in 1,3,5,10s time window
Area	47	Area	Area

would be large with a size of 47. So, to prevent overfitting and enhance the classification performance, we use a Feature Selection algorithm, to select highly correlated features with the class label, and remove redundant and irrelevant features. This step reduces the dimensionality of the data and allows learning algorithms to operate faster and more effectively [23].

V. RESULTS

To analyze the data, we used Weka [24] data mining tool to select the features, train and test the model (version 3.8 of Weka, contains 57 different classifiers by default and we have tried all of them). We used 10-fold cross-validation technique to evaluate the trained models. To select the most relevant features, in case of basic emotions which constitute a discrete set, we used InformationGain algorithm (Table III). We select features ranked higher than 0. Regarding Valence and Arousal, which constitute a continuous spectrum, we used the ReleifAttributeEval algorithm to select the features. This algorithm calculated a weight between -1 and $+1$ for each feature and higher positive weights indicating more predictive attributes. Hence, we applied this algorithm several times till the attempt which the ranks are higher than 0.

Among all trained models, we selected the ones gaining the highest classification accuracy together with minimum variance and Root Mean Square Error (RMSE). In case of emotions, a meta-classifier, AttributeSelectedClassifier with a J48 base-classifier, performs better than other Weka's classifiers. Specifically, on average the model reaches 71.92 ± 0.51, and 0.27 RMSE. The average Confusion Matrix is listed in Table V, the best classification rate corresponds to happiness (82.42%) and the worst is scared (40%). In case of Arousal, RandomForrest algorithm reaches a Correlation Coefficient of 0.78 ± 0.002 with 0.1 RMSE. In case of Valence, the model gained 0.76 ± 0.003 Correlation Coefficient on the average (RMSE 0.2).

To evaluate the model, we trained a model using Leave-One-

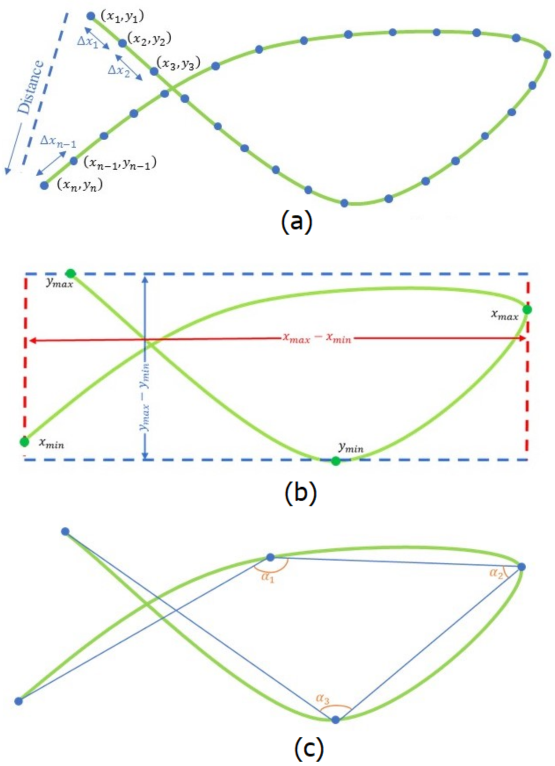

Fig. 2. (a) To measure the length of the traversed path, we calculate the sum of the Euclidean distance of each consecutive point. While ‘displacement’ (the dashed line) considers only the starting and ending point of each stroke. (b) To have an idea of the spanned area we considered the positional differences of x,y recorded values. (c) 3 points were selected to measure the angle between them in order to examine the level of rotation performed by the subject.

Subject-Out (LOSO) Cross-Validation and gained $52.16\% \pm 12.37$. Also, to evaluate the effectiveness of the novel features, we trained other models using only conventional features, i.e. {mean, min, max, median} of {number of strokes, stroke length, stroke speed, Directness Index, Pressure}. In case of emotions (Table VI), the model reaches $46.62\% \pm 0.72$ (RMSE 0.31). Considering Arousal levels, a model trained using conventional features gained 0.58 ± 0.006 (RMSE 0.1). Turning to Valence, the model reaches on the average for 10 runs, a Correlation Coefficient of 0.48 ± 0.007 with 0.2 RMSE. Finally, to reduce the negative effect of imbalanced data, we trained another model ignoring the sparse classes, i.e. Surprised and Scared. This model gains $72.77\% \pm 0.78$ (RMSE 0.33).

VI. Discussion

Our results revealed that the models perform well in general, but when viewed by class, in some classes the performance is less satisfactory than expected. One source of uncertainty, could be the imbalanced data we gathered. Although we used a meta-classifier, which is less sensitive to imbalanced data [25], however, collecting more data especially for the unseen classes enhances the performance.

TABLE III
Features Selection (InformationGain) output. The features which ranked higher than zero were selected.

Rank	Feature Abbreviation	Rank	Feature Abbreviation
0.5196	curly	0.0271	2ndDVPCy
0.5179	curlx	0.0267	distanceY
0.2607	numOfStrkW10	0.0264	angleM
0.1731	numOfStrkW5	0.0258	spanY
0.1419	numOfStrkW3	0.0249	Duration
0.1213	timeElappsed	0.0247	ySTD
0.1115	pMean	0.0247	mSTD
0.1083	pMax	0.0216	1stDVcurly
0.0882	numOfStrkW1	0.0216	1stDVPCy
0.0746	1stDVPP	0.0204	acceleration
0.0746	2ndDVPP	0.0178	yMax
0.0746	velocity	0.0177	angle1
0.0703	distanceX	0	wj_Max
0.0697	length	0	wj_Min
0.0667	1stDVcurlx	0	pMedian
0.0651	Area	0	xMin
0.0623	xSTD	0	angle2
0.0617	spanX	0	pMin
0.0595	2ndDVPCx	0	yMin
0.0542	1stDVPCx	0	yMean
0.0452	displacement	0	xMedian
0.0383	dist2prev	0	xMean
0.0383	xMax	0	yMedian
0.0289	wj_Mean		

TABLE IV
Features Selection (ReliefAttributeEval) output. For each model, the features which ranked higher than zero were selected. Features ranked less than 0.001 are not listed.

Arousal		Valence	
Rank	Feature	Rank	Feature
0,024962	timeElappsed	0,02224	wj$_{Mean}$
0,016085	curlx	0,01715	timeElappsed
0,016080	curly	0,00938	numOfStrkW10
0,007266	pMean	0,0065	wj$_{Max}$
0,006762	pMax	0,00498	acceleration
0,006028	Dist2prev	0,00466	numOfStrkW5
0,005236	numOfStrkW10	0,0043	curlx
0,004755	distanceY	0,00429	curly
0,004392	distanceX	0,00401	angle1
0,003909	xSTD	0,0035	numOfStrkW3
0,002673	Duration	0,0024	xMax
0,002620	2ndDVPCy	0,0022	pMax
0,002580	yMedian	0,00204	pMean
0,002580	pMedian	0,00169	numOfStrkW1
0,002384	2ndDVPCx		
0,002003	numOfStrkW5		

Comparing previously proposed methods with this study is tricky, because not only the data used is different, but also the training techniques, experiment setups, emotional

TABLE V
EMOTION RECOGNITION CLASSIFICATION CONFUSION MATRIX

	Classified as					
	Happy	Sad	Angry	Surprised	Scared	Disgusted
Happy	**539**	78	14	7	3	13
Sad	103	**239**	16	0	0	11
Angry	41	24	**52**	1	2	6
Surprise	12	2	1	**17**	1	0
Scared	6	1	5	0	**8**	0
Disgusted	20	10	5	0	0	**120**

TABLE VI
SEMI-BALANCED CLASSIFIER CONFUSION MATRIX. THE SPARSE CLASSES HAVE BEEN REMOVED TO REACH A MORE BALANCED DATASET.

Classified as ->	Happy	Sad	Angry	Disgusted
	530	87	22	15
	79	**261**	15	14
	43	27	**50**	6
	23	11	3	**118**

labels, etc. are different. Nevertheless, as seen in the result section, the emotion recognition model trained using solely the conventional features introduced in other literature, reached an accuracy of 46.62 ± 0.72 (RMSE 0.31), while using the novel features suggested here the accuracy reaches 71.92 ± 0.51 (RMSE 0.27). In case of modeling Arousal, a model trained with conventional features reaches an average Correlation Coefficient of 0.59 ± 0.006 (RMSE 0.1), while involving the novel features the Correlation coefficient is increased. Regarding valence, the model trained using conventional features reached an average Correlation Coeficient of 0.4846 ± 0.007 (RMSE 0.19), while using the novel features enhances the accuracy up to 0.7629 ± 0.003. These numbers endorse the effectiveness of the novel features introduced, i.e. timing features, acceleration of the stroke, curl, angular velocity and acceleration, angle, area, and time window counting strokes.

To check the impact of using contextual features that grasp some aspects of concrete gameplay, we trained another model, without the contextual features (i.e. elapsed time and position in x,y). In case of Emotional labels, a model trained using context-free features reached an accuracy of $66.23\% \pm 0.93$ (RMSE 0.27). Regarding Arousal levels, such model reaches an average Correlation Coefficient of 0.7411 ± 0.0039 (RMSE 0.1) and in case of Valence, the model gained 0.6278 ± 0.005 (RMSE 0.2). The accuracy is a bit lower, which supports the effectiveness of contextual features. Nevertheless, the model without those features, hence more generic, is still promising in recognition of Emotion, Arousal and Valence.

Also, the models perform better than baselines. Specifically, zeroR classifier, which simply predicts the majority without any predictability power, reaches lower rate. Considering Emotions, zeroR classifiers gains 48.19% as the baseline. Regarding Arousal levels, ZeroR classifier reaches the Correlation Coefficient of -0.0912 (RMSE 0.2). In case of Valence, ZeroR gains -0.0757 (RMSE 0.3). These results show that the models proposed in this study, work way to far more promising than random models. Table VII summarizes these comparisons. Yet, a more precise comparison could be done by applying our model on the same datasets used in other studies.

TABLE VII
EXAMINING THE EFFECTIVENESS OF THE FEATURES.

Model	Novel Features	Context-Free	Conv. Features	ZeroR
Emotions	71.92%±0.51	66.23%±0.93	46.62%±0.07	48.19%
Arousal	0.7796±0.002	0.7411±0.004	0.5894±0.006	-0.0912
Valence	0.7629±0.002	0.6278±0.005	0.4846±0.006	-0.0757

It is important to bear in mind the possible bias of imbalanced data in the reported results. As shown in the results section, removing the sparse classes enhanced the accuracy of the model. More specifically, the model could not recognize the members of these two classes, due to fewer learned instances. On the other hand, in case of Arousal and Valence, the Correlation Coefficient decreased a little bit. These decrease in the accuracy might be due to overfitting of the models toward the more populated classes.

VII. CONCLUSION AND FUTURE DIRECTIONS

In this paper, we proposed a model to recognize users' emotional state while playing a game on a typical tablet. In final words, we conclude that using a cheap sensor available in all tablets with low level of computation cost, the proposed model provides the opportunity to detect emotions, arousal and valence. This approach has the capability of being implemented in real-time (context-free apps) and semi-real-time (context-sensitive) applications. While our findings offer many interesting insights into emotion recognition using tactile features, our study has limitation that points to follow up studies. The first and foremost step could be increasing the game contextual features, using video-taped events happened during gameplay, as a source of triggering emotions. Further, applying other methods of feature selection/reduction, such as the Principal Component Analysis or Correlation Based techniques, may improve the performance. Also, we plan to add more novel features, such as, kinetic energy of strokes, angular kinetic energy, and the level of expertise that might be another interesting features to consider. Another step worth to follow is expanding the model to detect and analyze strokes received from different fingers on a multi-touch screen. Also using devices having inbuilt sensors capable of measuring real pressure (the one used in [21]) would provide more information about strokes. The study could be repeated using other implicit measurements of emotions (ex. Self Assessment Manikin) that allow for linguistic and cultural issues or biases related to the introspective verbalisation. Also, combing different methodologies (e.g., physiological indices, self-report, implicit measures, facial expressions) may increase the reliability of emotions measure. Note that the approach might not be applicable which works merely with other sensors such as gyroscope.

Acknowledgment

This work was funded by the EC H2020 project RAGE (Realising an Applied Gaming Eco-System); http://www.rageproject.eu/; Grant agreement No 644187, the AMIGOS project (PTDC/EEISII/7174/2014), and was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2019.

References

[1] H. Gunes, *Automatic, dimensional and continuous emotion recognition*. IGI Global, 2010, ch. 8.

[2] B. Fasel and J. Luettin, "Automatic facial expression analysis: a survey," *Pattern recognition*, vol. 36, no. 1, pp. 259–275, 2003.

[3] S. G. Koolagudi and K. S. Rao, "Emotion recognition from speech: a review," *International journal of speech technology*, vol. 15, no. 2, pp. 99–117, 2012.

[4] A. Haag, S. Goronzy, P. Schaich, and J. Williams, "Emotion recognition using bio-sensors: First steps towards an automatic system," in *Tutorial and research workshop on affective dialogue systems*. Springer, 2004, pp. 36–48.

[5] J. Anttonen and V. Surakka, "Emotions and heart rate while sitting on a chair," in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2005, pp. 491–499.

[6] A. Nakasone, H. Prendinger, and M. Ishizuka, "Emotion recognition from electromyography and skin conductance," in *Proc. of the 5th International Workshop on Biosignal Interpretation*. Citeseer, 2005, pp. 219–222.

[7] S. Lokannavar, P. Lahane, A. Gangurde, and P. Chidre, "Emotion recognition using eeg signals," *Emotion*, vol. 4, no. 5, 2015.

[8] A. Kleinsmith and N. Bianchi-Berthouze, "Affective body expression perception and recognition: A survey," *IEEE Transactions on Affective Computing*, vol. 4, no. 1, pp. 15–33, 2013.

[9] Y. Fu, H. V. Leong, G. Ngai, M. X. Huang, and S. C. Chan, "Physiological mouse: toward an emotion-aware mouse," *Universal Access in the Information Society*, vol. 16, no. 2, pp. 365–379, 2017.

[10] N. Carbonaro, A. Greco, G. Anania, G. Dalle Mura, R. Tognetti, P. Scilingo, and D. De Rossi, "Unobtrusive physiological and gesture wearable acquisition system: a preliminary study on behavioral and emotional correlations," 2012.

[11] B. Schuller, G. Rigoll, and M. Lang, "Emotion recognition in the manual interaction with graphical user interfaces," in *Multimedia and Expo, 2004. ICME'04. 2004 IEEE International Conference on*, vol. 2. IEEE, 2004, pp. 1215–1218.

[12] D. H. Shanabrook, I. Arroyo, and B. P. Woolf, "Using touch as a predictor of effort: what the ipad can tell us about user affective state," in *International Conference on User Modeling, Adaptation, and Personalization*. Springer, 2012, pp. 322–327.

[13] Y. Gao, N. Bianchi-Berthouze, and H. Meng, "What does touch tell us about emotions in touchscreen-based gameplay?" *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 19, no. 4, p. 31, 2012.

[14] P. E. Ekman and R. J. Davidson, *The nature of emotion: Fundamental questions.* Oxford University Press, 1994.

[15] L. F. Barrett, B. Mesquita, K. N. Ochsner, and J. J. Gross, "The experience of emotion," *Annu. Rev. Psychol.*, vol. 58, pp. 373–403, 2007.

[16] I. B. Mauss and M. D. Robinson, "Measures of emotion: A review," *Cognition and emotion*, vol. 23, no. 2, pp. 209–237, 2009.

[17] K. Bakhtiyari, M. Taghavi, and H. Husain, "Implementation of emotional-aware computer systems using typical input devices," in *Asian Conference on Intelligent Information and Database Systems*. Springer, 2014, pp. 364–374.

[18] P. Lewinski, T. M. den Uyl, and C. Butler, "Automated facial coding: Validation of basic emotions and facs aus in facereader." *Journal of Neuroscience, Psychology, and Economics*, vol. 7, no. 4, p. 227, 2014.

[19] P. Ekman, "An argument for basic emotions," *Cognition & emotion*, vol. 6, no. 3-4, pp. 169–200, 1992.

[20] J. Posner, J. A. Russell, and B. S. Peterson, "The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology," *Development and psychopathology*, vol. 17, no. 03, pp. 715–734, 2005.

[21] M. Exposito, R. W. Picard, and J. Hernandez, "Affective keys: towards unobtrusive stress sensing of smartphone users," in *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. ACM, 2018, pp. 139–145.

[22] P. Zimmermann, S. Guttormsen, B. Danuser, and P. Gomez, "Affective computinga rationale for measuring mood with mouse and keyboard," *International journal of occupational safety and ergonomics*, vol. 9, no. 4, pp. 539–551, 2003.

[23] M. A. Hall, "Correlation-based feature selection of discrete and numeric class machine learning," 2000.

[24] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: an update," *ACM SIGKDD explorations newsletter*, vol. 11, no. 1, pp. 10–18, 2009.

[25] W.-J. Lin and J. J. Chen, "Class-imbalanced classifiers for high-dimensional data," *Briefings in bioinformatics*, p. bbs006, 2012.