



**IJIRCCCE**

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

**Volume 12, Issue 8, August 2024**

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

**Impact Factor: 8.625**



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com



# Facial Expression Recognition with Two-Branch Disentangled Generative Adversarial Network

Selvi P, M. Naveena

Assistant Professor, Department of CSE, Jayam College of Engineering and Technology, Dharmapuri,  
Tamil Nadu, India

PG Student, Department of Master of Computer Application, Jayam College of Engineering and Technology,  
Dharmapuri, Tamil Nadu, India

**ABSTRACT:** Facial Expression Recognition (FER) is a challenging task in computer vision as features extracted from expressional images are usually entangled with other facial attributes, e.g., poses or appearance variations, which are adverse to FER. To achieve a better FER performance, we propose a model named Two-branch Disentangled Generative Adversarial Network (TDGAN) for discriminative expression representation learning. Different from previous methods, TDGAN learns to disentangle expressional information from other unrelated facial attributes. To this end, we build the framework with two independent branches, which are specific for facial and expressional information processing respectively. Correspondingly, two discriminators are introduced to conduct identity and expression classification. By adversarial learning, TDGAN is able to transfer an expression to a given face. It simultaneously learns a discriminative representation that is disentangled from other facial attributes for each expression image, which is more effective for FER task. In addition, a self-supervised mechanism is proposed to improve representation learning, which enhances the power of disentangling. Quantitative and qualitative results in both in-the-lab and in-the-wild datasets demonstrate that TDGAN is competitive to the state-of-the-art methods.

## I. INTRODUCTION

The notion of smart cities is founded on developing advanced, automated and connected intelligent industrial applications that conveniently create a sustainable and livable city for its dwellers. Even though paramount importance is directed towards smart applications in a smart city environment, limited attention has been paid to identify the very component which keeps the city alive, it's citizens. In the advancement of intelligent industrialization towards what is being called the fourth industrial revolution, it is important to utilize Artificial Intelligence (AI) to understand citizens' emotions and perceptions towards industrial applications. Even in cognitive automation, it is mentioned that the further optimization and development of many smart applications rely on the subjective wishes of the user indicating the importance of recognizing the citizens' perspective. Although attempts have been made to measure citizens' emotional reactions via physical sensors the use of sensors poses privacy and management challenges, when it concerns continuous monitoring. In addition, such physical sensors are not able to sense the opinions and emotional reactions of citizens that continuously influence smart city initiatives. Nevertheless, citizens express their opinions, observations, and perceptions daily via social media conversations. This paves a pathway and an opportunity to create smart observation systems which enable to feel the 'emotional pulse' of the city via emotions expressed collectively by citizens in publicly available data sources such as social media conversations.

### 1.1. AI-BASED EMOTION OBSERVATION FRAMEWORK

The advent of social media and the massive amounts of data generated via these platforms create a large pool of data that encompasses emotions, feelings, and thoughts of citizens. With the exponential growth in social media, the wealth of information generated on social media can be viewed as a 'big data' source capturing the citizens' voice. Studies show that social media channels, with their large repositories of user-generated data, provide an opportunity to gain insights into citizens' well-being, defined as the 'emotion pulse' of a nation. In this study, we relate the concept of emotional pulse to the smart city context in order to understand the perceptions of the citizens towards the smart city environment and applications. As all smart applications are built around the day-to-day lives of citizens and with the



## International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

purpose of serving the citizens better, it is important to understand and ‘sense’ their perception in order to design future innovations and policies. Despite this importance, utilizing citizen opinion and their emotional reactions related to smart city applications is currently limited to mere sentiment and emotion extraction and does not deeply explore the emergence and shift of emotional reactions using social media data.

Given the fact that free-flowing social media content has not been deeply explored in industrial informatics research with the objective of detecting the emotional pulse of citizens in a smart city context, we propose the first practical use of social media to capture the pulse of the city by developing an AI-based emotion observation framework. The proposed framework is able to monitor the emotional pulse of citizens using social media conversations, thus derive the emergence and shift of emotions as well as examine the negativity of public opinion in a smart city context.

### 1.2. COGNITIVE INTERNET OF THINGS

In the recent era, artificial intelligence (AI) is being used to support numerous solutions for human beings, such as healthcare, autonomous transportation, and so on. Cognitive computing is represented as a next-generation application AI-based solutions which provide human-machine interaction with personalized interactions and services that imitate human behavior. On the other hand, a large volume of data is generated from smart city applications such as healthcare, smart transportation, retail industry, and freighting. There is always a concern on how to efficiently manage the large volume of generated data. Recently many existing researches discussed the analysis of the large quantity of data using cognitive computing; however, these researches are failed to handle the certain problems, namely scalability, and flexibility of data gathered in a smart city environment. Data captured from millions of sensors can be cross implemented across various cognitive computing applications to ensure real-time responses. In this paper, we study the cognitive internet of things (CIoT) and propose a CIoT-based smart city network (CIoT-Net) architecture which describes how data gathered from smart city applications can be analyzed using cognitive computing and handle the scalability and flexibility problems. We discuss various technologies such as AI and big data analysis to implement the proposed architecture. Finally, we describe the possible research challenges and opportunities while implementing the proposed architecture.

### 1.3. OPTIMIZATION ALGORITHM

- Optimization is the problem of finding a set of inputs to an objective function that results in a maximum or minimum function evaluation.
- It is the challenging problem that underlies many machine learning algorithms, from fitting logistic regression models to training artificial neural networks.
- There are perhaps hundreds of popular optimization algorithms, and perhaps tens of algorithms to choose from in popular scientific code libraries. This can make it challenging to know which algorithms to consider for a given optimization problem.
- In this tutorial, you will discover a guided tour of different optimization algorithms.

After completing this tutorial, you will know:

- Optimization algorithms may be grouped into those that use derivatives and those that do not.
- Classical algorithms use the first and sometimes second derivative of the objective function.
- Direct search and stochastic algorithms are designed for objective functions where function derivatives are unavailable.
- Optimization refers to a procedure for finding the input parameters or arguments to a function that result in the minimum or maximum output of the function.
- The most common type of optimization problems encountered in machine learning are continuous function optimization, where the input arguments to the function are real-valued numeric values, e.g. floating point values. The output from the function is also a real-valued evaluation of the input values.



## International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

### II. LITERATURE SURVEY

**Title-** Conceptualizing Smart City with Dimensions of Technology, People, and Institutions

**Author-** Artemis Psaltoglou

**Year-**2018

**Description** -This conceptual paper discusses how we can consider a particular city as a smart one, drawing on recent practices to make cities smart. A set of the common multidimensional components underlying the smart city concept and the core factors for a successful smart city initiative is identified by exploring current working definitions of smart city and a diversity of various conceptual relatives similar to smart city. The paper offers strategic principles aligning to the three main dimensions (technology, people, and institutions) of smart city:

**Title-** Conceptualizing Smart City with Dimensions of Technology, People, and Institutions

**Author-** Taewoo Nam & Theresa A. Pardo

**Year-**2011

**Description** – This conceptual paper discusses how we can consider a particular city as a smart one, drawing on recent practices to make cities smart. A set of the common multidimensional components underlying the smart city concept and the core factors for a successful smart city initiative is identified by exploring current working definitions of smart city and a diversity of various conceptual relatives similar to smart city. The paper offers strategic principles aligning to the three main dimensions (technology, people, and institutions) of smart city: integration of infrastructures and technology-mediated services, social learning for strengthening human infrastructure, and governance for institutional improvement and citizen engagement.

**Title-** Cognitive Automation—Survey of Novel Artificial General Intelligence Methods for the Automation of Human Technical Environments

**Author-** Dietmar Bruckner, HeimoZeilinger, Dietmar Dietrich,

**Year-**2011

**Description** – Automation, the utilization of control and information technologies for reducing the need for human intervention in the production process is about to meet Cognition-the science concerned with human thinking-and related sciences. More and more processes require analysis and insights that allow controlling them beyond the mere execution of rules and beyond prefitted controllers in order to automatically keep them within the desired conditions. Automatic and flexible decision making based on challenging conditions such as increasing amounts of information, lacking prior knowledge of data, incomplete, missing or contradicting data, becomes the key challenges for future automation technologies.

**Title-** Emotional artificial intelligence in children’s toys and devices: Ethics, governance and practical remedies

**Author-** [Lupton and Williamson](#),

**Year-**2017

**Description** – This article examines the social acceptability and governance of emotional artificial intelligence (emotional AI) in children’s toys and other child-oriented devices. To explore this, it conducts interviews with stakeholders with a professional interest in emotional AI, toys, children and policy to consider implications of the usage of emotional AI in children’s toys and services. It also conducts a demographically representative UK national survey to ascertain parental perspectives on networked toys that utilise data about emotions. The article highlights disquiet about the evolution of generational unfairness, that encompasses injustices regarding the datafication of childhood, manipulation, parental vulnerability, synthetic personalities, child and parental media literacy, and need for improved governance. It concludes with practical recommendations for regulators and the toy industry.

**Title** -Emotions of COVID-19: Content Analysis of Self-Reported Information Using Artificial Intelligence

**Author-** [AchiniAdikari](#), BSc, [RashmikaNawaratne](#), BSc, PhD,#1 [Daswin De Silva](#), BSc, PhD, [SajaniRanasinghe](#), BSc, [OshadiAlahakoon](#), BSc, PhD, and [DammindaAlahakoon](#), BSc, PhD

**Year-**2021

**Description-** The COVID-19 pandemic has disrupted human societies around the world. This public health emergency was followed by a significant loss of human life; the ensuing social restrictions led to loss of employment, lack of



## International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

interactions, and burgeoning psychological distress. As physical distancing regulations were introduced to manage outbreaks, individuals, groups, and communities engaged extensively on social media to express their thoughts and emotions. This internet-mediated communication of self-reported information encapsulates the emotional health and mental well-being of all individuals impacted by the pandemic.

### III. EXISTING SYSTEM

Facial Expression Recognition (FER) is one of the most popular research topics in the field of computer vision, which aims to recognize six basic facial expressions (i.e., angry, disgust, fear, happy, sad and surprise) from images or videos. Previous works on FER usually go by two steps: researchers first extract expressional features to represent the given image/video and then train/use a classifier to recognize different expressions based on the extracted features. Conventional methods utilize some hand-crafted features to represent an expressional images. In recent years, deep-learning algorithms are also introduced to solve the FER task. Since deep-learning models learn features automatically and always be trained in an end-to-end manner, the performance of them is usually better and therefore be regarded as a promising solution. However, FER is still far from being solved. On one hand, expressional actions usually occur in specific local areas (e.g., the neighbourhood of eyes or mouth), while features extracted from expression-unrelated regions can be redundant to FER and may deteriorate the overall performance.

#### DISADVANTAGES

- Despite its popularity, current softmax loss-based approach does not explicitly reward intra-class compactness and inter-class separation, and identity-related factors remain major obstacles for FER.
- Machine recognition usually is based on similarity metrics, but those metrics may be more sensitive to identity than expressions. To decouple these two types of similarity and exploit the appearance information, substantial efforts have been dedicated to extracting features by learning.
- Given that the expressions are formed by relaxing or contracting some facial muscles that result in temporally deformed facial features, identity-disentangled representations for FER normally separate a face with expression into a main component neutral face that encodes identity cues and an action component that encodes motion cues (such as movements of eye brows, cheeks, lips, eyelids and nose) which are related to the AUs and FACS.
- FER is certainly not unique among computer vision applications that have to cope with nuisance factors causing variability in the data. Deep metric learning approaches have been shown to be successful for person and vehicle identification tasks, which also exhibit large intra-class variations. The initial work in this domain involves training a Siamese network.

#### PROPOSED SYSTEM

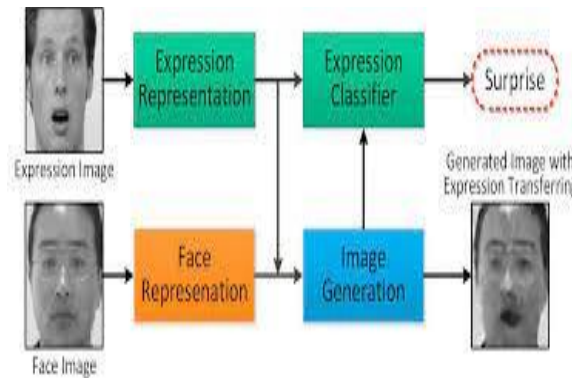
- We propose a model named TDGAN that is able to learn disentangled and discriminative expressional representation for FER task. By adopting adversarial learning, TDGAN is encouraged to disentangle expressional information from other redundant facial attributes or variations, which results in a better performance on FER task.
- The proposed TDGAN is able to conduct expression transferring. Since representations of the input image pair can be effectively fused, TDGAN is able to synthesize an image with expected expression by modifying some specific regions of the input face image. TDGAN can be trained in some small expression datasets.
- By jointly training with an auxiliary face dataset, TDGAN is encouraged to learn the data distribution of both expression and face datasets. This makes it possible to work well even though there are only limited training samples of labelled expression images.
- TDGAN outperforms many existing FER methods on both in-the-lab and in-the-wild datasets. Visualization and recognition results demonstrate the effectiveness of our model



## International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

### BLOCK DIAGRAM



### IV. MODULES DESCRIPTION

#### Classification of emotions

It should be evident that waiting for a consensus on how experts classify emotions should be quite unrealistic currently because the study of emotions is still incomplete. However, there is some acceptance of classifying emotions according to the emotion function, i.e., depending on the goal that each emotion has, as an adaptive, a motivational, an informative or a social emotion.

#### Smart city applications and citizens

The design and the conceptualization of a smart city necessitate several factors including technology, policies, economy, governance and people communities. As the current approaches in smart cities strive to create an allconnected environment it is imperative to understand the smart initiatives from citizens' perspective which would be helpful for policymakers and industrial leaders when developing future strategies. Although citizen engagement is known as a key component in most definitions of smart cities, limited research has been carried out to investigate the practicality and applicability of understanding public participation

#### Artificial Emotional Intelligence

This article examines Artificial Emotion Intelligence (AEI) and its application in social robots. It argues that AEI and social robotics intensify practices of data-capture and algorithmic governance by extending the spatial reach of digital surveillance deeper into intimate spaces and individual psyches, with the goal of manipulating human emotional and behavioural responses. The analysis demonstrates the need to more thoroughly engage the multiplicity of theoretical and applied approaches to building artificial intelligence, to question assumptions as to the kinds of intelligence being created, and to consider how a diversity of AI systems infiltrate and reshape the spaces of everyday life.

#### Convolutional Neural Network (CNN) Layers

While recurrent neural networks extract orderly, sequential information, convolution neural networks are able to capture indistinct combinations of features. In the classification model, a 1D-CNN layer was added next to take the output of each separate GRU layer. The hyperparameters of the CNN layer were as follows; kernel size =3, filters =64 and the ReLU was used as the activation function. The outputs of the CNN layers were then directed through separate average pooling and max-pooling layers and the outputs of the pooling layers were concatenated. After concatenating the outcomes of the pooling layers, the dense layer was used to produce 6 outputs for each label.

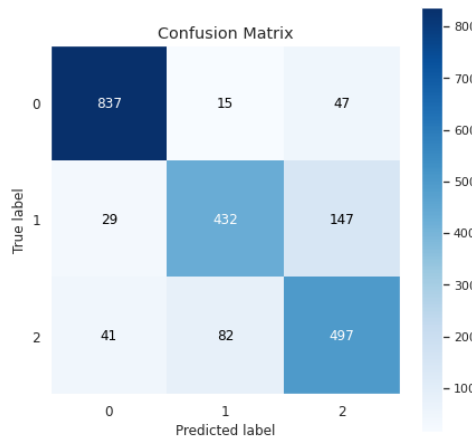
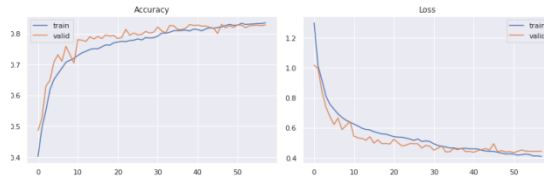
### V. RESULT AND DISCUSSION

Facial Expression Recognition (FER) in the wild is extremely challenging due to occlusions, variant head poses, face deformation and motion blur under unconstrained conditions.



## International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Although substantial progresses have been made in automatic FER in the past few decades, previous studies were mainly designed for lab-controlled FER. Real-world occlusions, variant head poses and other issues definitely increase the difficulty of FER on account of these information-deficient regions and complex backgrounds. Different from previous pure CNNs based methods, we argue that it is feasible and practical to translate facial images into sequences of visual words and perform expression recognition from a global perspective. Therefore, we propose the Visual Transformers with Feature Fusion (VTFF) to tackle FER in the wild by two main steps. First, we propose the attentional selective fusion (ASF) for leveraging two kinds of feature maps generated by two-branch CNNs. The ASF captures discriminative information by fusing multiple features with the global-local attention. The fused feature maps are then flattened and projected into sequences of visual words. Second, inspired by the success of Transformers in natural language processing, we propose to model relationships between these visual words with the global self-attention. The proposed method is evaluated on three public in-the-wild facial expression datasets (RAF-DB, FERPlus and AffectNet). Under the same settings, extensive experiments demonstrate that our method shows superior performance over other methods, setting new state of the art on RAF-DB with 88.14%, FERPlus with 88.81% and AffectNet with 61.85%. The cross-dataset evaluation on CK+ shows the promising generalization capability of the proposed method



### VI. CONCLUSION

In this paper we have proposed a unified expression transfer, editing and recognition architecture, TER-GAN, which has two objectives: 1). to extract efficient and disentangled expression and identity features from input images, and 2). to employ the extracted expression and identity representations for realistic looking expression synthesis that preserves the identity information of the target (given) image. This goal is achieved by explicitly encoding the expression information from a source image and extracting identity information from a target image by using two different



## International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

dedicated encoders, and these two feature vectors are then combined to generate an expression image by employing the decoder part of TER-GAN. In order to further improve the expression and identity feature extraction process, we have introduced novel expression and identity consistency losses. Experimental results show that the proposed method can be used for efficient facial expression transfer and facial expression editing, and the disentangled feature representation can be used for facial expression recognition.

### FUTURE WORK

This article examines the social acceptability and governance of emotional artificial intelligence (emotional AI) in children's toys and other child-oriented devices. To explore this, it conducts interviews with stakeholders with a professional interest in emotional AI, toys, children and policy to consider implications of the usage of emotional AI in children's toys and services. It also conducts a demographically representative UK national survey to ascertain parental perspectives on networked toys that utilise data about emotions. The article highlights disquiet about the evolution of generational unfairness, that encompasses injustices regarding the datafication of childhood, manipulation, parental vulnerability, synthetic personalities, child and parental media literacy, and need for improved governance. It concludes with practical recommendations for regulators and the toy industry.

### REFERENCES

- [1] T. Nam and T. A. Pardo, "Conceptualizing Smart City with Dimensions of Technology, People, and Institutions," in Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times, New York, NY, USA, 2011, pp. 282–291.
- [2] A. Vanolo, "Is there anybody out there? The place and role of citizens in tomorrow's smart cities," *Futures*, vol. 82, pp. 26–36, Sep. 2016.
- [3] D. Bruckner, H. Zeilinger, and D. Dietrich, "Cognitive Automation—Survey of Novel Artificial General Intelligence Methods for the Automation of Human Technical Environments," *IEEE Trans. Ind. Inform.*, vol. 8, no. 2, pp. 206–215, May 2012.
- [4] V. C. C. Roza and O. A. Postolache, "Citizen emotion analysis in Smart City," in 2016 7th International Conference on Information, Intelligence, Systems Applications (IISA), Jul. 2016, pp. 1–6.
- [5] M. Li, E. Ch'ng, A. Chong, and S. See, "The new eye of smart city: Novel citizen Sentiment Analysis in Twitter," in 2016 International Conference on Audio, Language and Image Processing (ICALIP), Jul. 2016, pp. 557–562.
- [6] D. Milne, C. Paris, H. Christensen, P. Batterham, and B. O'Dea, "We Feel: Taking the emotional pulse of the world," Aug. 2015, Accessed: Oct. 08, 2019. [Online]. Available: <https://publications.csiro.au/rpr/pub?pid=csiro:EP153791>.
- [7] E. Diaz-Aviles, C. Orellana-Rodriguez, and W. Nejdl, "Taking the Pulse of Political Emotions in Latin America Based on Social Web Streams," in 2012 Eighth Latin American Web Congress, Oct. 2012, pp. 40–47.
- [8] D. Doran, S. Gokhale, and A. Dagnino, "Human Sensing for Smart Cities," in Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, New York, NY, USA, 2013, pp. 1323–1330.
- [9] B. Granier and H. Kudo, "How are citizens involved in smart cities? Analysing citizen participation in Japanese 'Smart Communities'," *Inf. Polity*, vol. 21, no. 1, pp. 61–76, Jan. 2016.
- [10] "Twitter suspended 166,153 accounts for terrorism content in second...", Reuters, May 09, 2019.
- [11] B. Guthier, R. Alharthi, R. Abaalkhail, and A. El Saddik, "Detection and Visualization of Emotions in an Affect-Aware City," in Proceedings of the 1st International Workshop on Emerging Multimedia Applications and Services for Smart Cities, New York, NY, USA, 2014, pp. 23–28.
- [12] "(16) TweetAlert: Semantic analytics in social networks for citizen opinion mining in the city of the future | J. Villena-Román | Request PDF," ResearchGate. [https://www.researchgate.net/publication/290332744\\_TweetAlert\\_Semantic\\_analytics\\_in\\_social\\_networks\\_for\\_citizen\\_opinion\\_mining\\_in\\_the\\_city\\_of\\_the\\_future](https://www.researchgate.net/publication/290332744_TweetAlert_Semantic_analytics_in_social_networks_for_citizen_opinion_mining_in_the_city_of_the_future) (accessed Jun. 26, 2019).
- [13] P. Zeile, B. Resch, J.-P. Exner, and G. Sagl, "Urban Emotions: Benefits and Risks in Using Human Sensory Assessment for the Extraction of Contextual Emotion Information in Urban Planning," in Planning Support Systems





## International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

and Smart Cities, S. Geertman, Jr. Ferreira Joseph, R. Goodspeed, and J. Stillwell, Eds. Cham: Springer International Publishing, 2015, pp. 209–225.

[14] A. Cacho et al., “Social Smart Destination: A Platform to Analyze User Generated Content in Smart Tourism Destinations,” in *New Advances in Information Systems and Technologies*, 2016, pp. 817–826.

[15] M. Balduini et al., “A Case Study of Active, Continuous and Predictive Social Media Analytics for Smart City,” in *Proceedings of the Fifth International Conference on Semantics for Smarter Cities - Volume 1280*, Aachen, Germany, Germany, 2014, pp. 31–46, Accessed: Jun. 26, 2019. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2878779.2878784>.

[16] M. Behrens, N. Valkanova, A. F. gen. Schieck, and D. P. Brumby, “Smart Citizen Sentiment Dashboard: A Case Study Into Media Architectural Interfaces,” in *Proceedings of The International Symposium on Pervasive Displays*, New York, NY, USA, 2014, pp. 19:19–19:24.

[17] Y. Bae and H. Lee, “Sentiment analysis of twitter audiences: Measuring the positive or negative influence of popular twitterers,” *J. Am. Soc. Inf. Sci. Technol.*, vol. 63, no. 12, pp. 2521–2535, Dec. 2012.

[18] R. Plutchik, *The Emotions*. University Press of America, 1991.

[19] S. M. Mohammad and P. D. Turney, “Crowdsourcing a Word-Emotion Association Lexicon,” Aug. 2013, Accessed: Oct. 17, 2018. [Online]. Available: <https://arxiv.org/abs/1308.6297>.

[20] R. W. Picard, *Affective Computing*. Cambridge, MA, USA: MIT Press, 1997.



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  [ijircce@gmail.com](mailto:ijircce@gmail.com)



[www.ijircce.com](http://www.ijircce.com)

Scan to save the contact details