A Deep Learning Approach To Analyze The Sentiment Of Online Game Users

Agung Wahana
Department of Informatics
UIN Sunan Gunung Djati
Bandung, Indonesia
wahana.agung@uinsgd.ac.id

Wildan Budiawan Zulfikar
Department of Informatics
UIN Sunan Gunung Djati
Bandung, Indonesia
wildan.b@uinsgd.ac.id

Wildan Najah Wildiansyah
Department of Informatics
UIN Sunan Gunung Djati
Bandung, Indonesia
wildan.najah@gmail.com

Aldy Rialdy Atmadja
Department of Informatics
UIN Sunan Gunung Djati
Bandung, Indonesia
aldy@if.uinsgd.ac.id

Diena Rauda Ramdania
Department of Informatics
UIN Sunan Gunung Djati
Bandung, Indonesia
diena.rauda@uinsgd.ac.id

Beki Subaeki
Department of Information System
Universitas Sangga Buana
Bandung, Indonesia
beki.subaeki@usbypkp.ac.id

***Abstract*— The policies to be set by online game developers can depend on user sentiment. This study aims to obtain information about the sentiment of online game users as one of the bases for decision making. The Convolutional Neural Network (CNN) algorithm is a method of deep learning that is used to obtain classification results regarding game user sentiment. The methodology used is the Cross-Industry Standard Process for Data Mining. Then preprocessing and weighting was carried out using GloVe. The accuracy value of the CNN algorithm is 81%. And it shows 80.4% is positive and 19.6% is negative. The results of this study can be used to assist in decision-making, which is then seen from the opinions of game players.**

Keywords—deep learning, CNN, convolutional neural network, social media, dota

I. Introduction

Nowadays, microblogging is a communication tool that is very widely used by internet users. It has evolved into a source of various types of information. Apart from that, it is a social media service where people post messages about their opinion on various topics, discuss current issues, complain, and express positive or negative sentiments for the products or services they use in their daily life. Many companies also create microblog polls to get general sentiment for their products. Often times companies study user reactions and reply to users on microblogging. Twitter is one of the microblogging websites with more than 300 million active users and more than 300 million tweets per day [1]–[5].

Sentiment analysis is a field of study that analyzes opinions, sentiments, judgments, evaluations, attitudes, and emotions in text data using a person's text analysis techniques related to a particular topic, service, product, individual, organization, or activity. Sentiment analysis is carried out to identify a person's opinion or comment on a problem that has positive or negative sentiments and can be used as a reference in improving the quality of products, services, individuals, organizations, or certain activities [6]–[10].

Deep learning is a sub-field in machine learning research that deals with artificial intelligence algorithms that are inspired by brain structures and functions called neural networks [11], [12]. Deep Learning is about learning multiple levels of representation and abstraction which help to understand and recognize data such as images, sounds, and text [11]. In previous studies, the main algorithms used in text classification were Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) [13], [14]. CNN (Convolutional Neural Networks) is one of the deep learning models that was created for computer vision and produced impressive results in image recognition, and in recent years this CNN model has proven effective for NLP (natural language processing) and has had very good results [4], [15], [16].

Video games are an interactive entertainment that grows along with technological advances and the availability of electronic devices for users is very easy to access because they can be played on computers, laptops, tablet PCs, or smartphones. This shows that the video game industry is experiencing a very rapid increase, due to the presence of internet access that can connect all game players around the world to play together or can be called online multiplayer games, with various game categories [17]–[20]. Such as the action-shooting, strategy, fighting, adventure, MOBA, and puzzle categories that can be played via a computer, laptop, smartphone, Xbox, Nintendo, or Playstation. The advent of online multiplayer games has created further avenues for sentiment analysis research.

Currently, one of the most popular online games is Dota 2 (Defense of the Ancients). Dota 2 allows players to connect, interact, cooperate virtually and accommodate multiple players anywhere. This game consists of two teams of five fighters each (heroes) defending their team arena. Dota 2 is a very popular but very complex online game [21].

Dota 2 has millions of users around the world. It also provides active competition for professional players through various leagues and tournaments that are scheduled. The

978-1-6654-4402-6/21/$31.00 ©2021 IEEE

biggest competition organized by Valve offers a total prize of up to tens of millions of US dollars. Each of these tournaments is broadcast on the internet as well as on television networks and seen by millions of people. Online games like Dota 2 can easily reach teenagers, adults, men, and women who are interested in online gaming. Since its introduction on July 9, 2013, Dota 2 has been developed to meet the needs of players around the world[21]–[23].

The purpose of this study is to provide information about users' sentiments towards their products. This study implements the Convolutional Neural Networks algorithm to classify a tweet related to the Dota 2 online video game as a positive or negative sentiment class. In addition, it calculates the accuracy of the Convolutional Neural Networks algorithm for sentiment analysis on Twitter social media related to the Dota 2 online video game. The results of this study can be used by online video game developers, especially Valve, which is a company that develops Dota 2 online video games in determining directions, goals, policies, etc. based on user sentiment.

II. Methodology

This study use CRISP-DM as a main framework. The first stage in the CRISP-DM framework is to develop business understanding. At this stage, the focus is on defining the problem to be solved or the question to be answered. In analyzing the data understanding the business and its specification problems is the most important stage.

A. *Business Understanding*

Important research has been conducted over the decades on the negative effects of play, including addiction, depression, and aggression, and we certainly don't suggest that this should be ignored. However, to understand the impact of video games on the development of children and adolescents, a more balanced perspective is needed.

This game can strengthen various cognitive skills such as spatial navigation, reasoning, memory, and perception. This is especially true for shooter video games. A 2013 meta-analysis found that playing shooter video games increased a player's capacity to think about objects in the same three dimensions as academic courses designed to enhance the same skills. This has important implications for education and career development, as previous research has established the strength of spatial skills for achievement in science, technology, engineering, and mathematics.

Playing video games can also help children develop problem-solving skills, say the authors. The more adolescents reported playing strategic video games, the more they improved in problem-solving and school grades the following year. Children's creativity is also enhanced by playing anything. Simple games that are easily accessible and can be played quickly, such as "Angry Birds," can elevate a player's mood, promote relaxation and ward off anxiety, the study said. If playing video games only makes people happier, this appears to be a fundamental emotional benefit to consider. Many think that gamers are socially isolated. More than 70 percent of gamers play with a friend, and millions of people around the world participate in massive virtual worlds through video games such as "DoTA 2", "Rising Force Online", "Farmville", and "World of Warcraft". Multiplayer games become virtual social communities, where decisions must be made quickly about who to trust or reject and how to lead the group. People who play video games, even if they are violent, which encourages cooperation are more likely to help others when playing the game than those who play the same games competitively.

The use of data in the form of text comments or tweets from microblogging Twitter is often used to process the sentiments of each comment or tweet. The sentiment is an attitude, thought, or judgment driven by feelings. Sentiment analysis, which is also known as opinion mining, studies people's sentiment towards a particular entity. The internet is a resourceful place when it comes to sentimental information. From a user perspective, people can post their content via various social media, such as microblog forums, or online social media sites. Online video games have negative and positive impacts, therefore this sentiment analysis about online video games on Twitter social media is carried out to find out people's direct opinions about online video games.

B. *Data Understanding*

This study uses the Convolutional Neural Networks algorithm, so it requires a lot of data to train it to be able to analyze positive and negative sentiments. Therefore, an additional dataset is used to train this Convolutional Neural Networks model, the dataset used is data from Stanford that has been collected by Andrew Maas. This dataset has 25 000 data trains which have evenly labeled 12500 positive data and 12500 negative data. Then it will be tested with a dataset of Twitter tweets containing the keyword Dota using the Twitter API which has 1441 datasets. When using Twitter social media there are various symbols such as "#" (hashtag), "@" (at), links, and other symbols that must be removed for better use. Meanwhile, the preprocessing method in English uses the SpaCy library.

C. *Data Preparation*

At this stage each data must be labeled, for the sentiment analysis model each comment can be labeled "1" for positive and "0" for negative as in table 1.

TABLE I. Sample of Documents

No	Content	Label
1	Next level Morphling morph fail lol"	0
2	The ESL India Winter Season LAN is now over and we hope you had a great time!! Which game were you most excited to watch!?	1

Natural Language Processing (NLP) is a subfield of computer science. This area focuses on how to create computer programs to process and analyze large amounts of natural language data. It is very difficult to do because the process of learning and understanding language is very complex. Tokenization is the process of tokenizing or separating a series of words in a sentence or paragraph into a list of tokens or single word pieces that stand alone as in table 2.

TABLE II. SAMPLE OF TOKENIZING

No	Before Tokenizing	After Tokenizing
1	the dota2 squad is preparing for the upcoming LA qualifier with our new sports psychologist	['the', 'dota2', 'squad', 'preparing', 'for', 'the', 'upcoming', 'LA', 'qualifier', 'with', 'our', 'new', 'sports', 'psychologist']
2	The ESL India Winter Season LAN is now over and we hope you had a great time!! Which game were you most excitedto watch!?	['The', 'ESL', 'India', 'Winter', 'Season', 'LAN', 'is', 'now', 'over', 'and', 'we', 'hope', 'you', 'had', 'a', 'great', 'time', '!', '!', ' ', 'Which', 'game', 'were', 'you', 'most', 'excited', 'to', 'watch', '!', '?']

Case folding is the next step. This step converts all letters into lowercase letters. In this process, the characters "A" - "Z" contained in the data are converted into characters "a" - "z". After that is the cleaning data. In this section, cleaning of data collected from the Twitter API client is carried out to separate the words and those that are difficult for the model to analyze and process. This type of data is difficult to understand and process for sentiment analysis models. Therefore, data must be cleaned such as links, '@', '#', and others.

D. Modeling

To create a sentiment analysis model that can predict comments including positive or negative sentences, this study uses the Convolutional Neural Networks (CNN) algorithm. Convolutional Neural Networks (CNN) in recent years have shown breakthrough results in Natural Language Processing (NLP), in particular, sentence classification, which is to classify phrases with about 20 to 50 tokens into predetermined categories.

When applied to text instead of images, it has a 1-dimensional array that represents the text. In natural language processing, most jobs with deep learning methods have involved learning vector word representations through neural language models and work composition during the study of word vectors for classification. Word vectors, where words are projected from a sparse 1-of-V encoding (where V is the size of the vocabulary) onto low-dimensional vector space via hidden layers, are feature extractors that encode the semantic features of words in their dimensions. In such a dense representation, semantically close words are also close in euclidean or cosine distances in a low-dimensional vector space.

In this study, CNN was trained simply with one layer of word vector convolution obtained from the unsupervised neural language model. CNN architecture is changed to convolutional operation and 1D unification. One of the most common tasks of NLP in which CNN is used is sentence classification, that is, classifying sentences into a predefined set of categories taking into account n-grams, i.e. words or word sequences, or also characters or word sequences. character.

1. 1-D Convolutions

It is given a sequence of words $w1: n = w1, \dots, wn$ where each is associated with an implanted dimensional vector d. The 1D-width convolution k is the result of moving a sliding window of size k over a sentence and applying the same convolution or kernel filter to each window in sequence, that is, the point-product between the embedding set of vectors in the given window and the weight vector u, which then often followed by a non-linear activation function g. Considering the word window $wi, \dots, wi + k$ the i-th window composite vector is (1):

$$x_i = [w_i, w_{i+1}, \dots, w_{i+k}] \in R^{k \times d} \quad (1)$$

$$r_i = g(x_i \cdot u) \in R \quad (2)$$

In practice, it is common to apply more filters, $u1, \dots, ul$, which can then be represented as a vector multiplied by a matrix U and by adding the bias term b (3) by (4).

$$r_i = g(x_i \cdot U + b) \quad (3)$$

$$r \in R^l,\ x_i \in R^{k \times d}, U \in R^{k \cdot d \times l} \text{ dan } b \in R^l \quad (4)$$

An example of sentence convolution in combined-vector notation is shown in Figure 1.

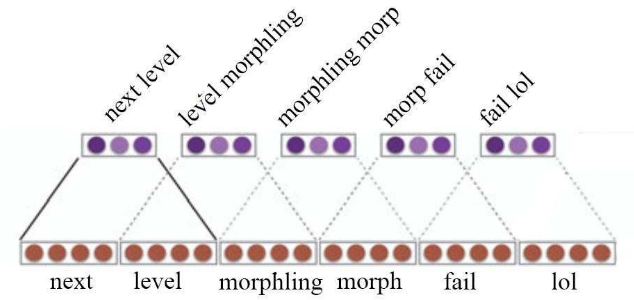

Fig. 1. Example of Convolution of sentences in vector-combined notation

2. Channel

The Convolutional Neural Network assumes that it is processing black and white images, and therefore we have one matrix that represents the intensity of the scale. CNN applies several paradigm channels to image processing as well as to text processing. For example, for a certain phrase or text window, one channel could be a sequence of words, another channel a corresponding POS tag sequence, and the third a word form.

Applying convolution to words will result in m vector w, applying it over PoS tag will result also vector m, and the same for shape, again vector m. These three different channels can then be combined either by addition (5).

$$p_i = words_{1:m} + pos_{1:m} + shapes_{1:m} \quad (5)$$

3. Pooling

The union operation is used to join the vectors generated from different convolution windows into a single l-dimensional vector. This is done again by taking the observed max or average value in the vector resulting from the convolution. Ideally, this vector will capture the most relevant features of the document.

4. Fully Connected

The two processes previously described namely: convolution and pooling, can be thought of as a feature extractor, then we pass this feature, usually as a vector reconstructed from a single line, further down the network, for example, a multi-layer perceptron to be created. trained for classification. Model architecture with two channels for an example sentence describes in fig 2.

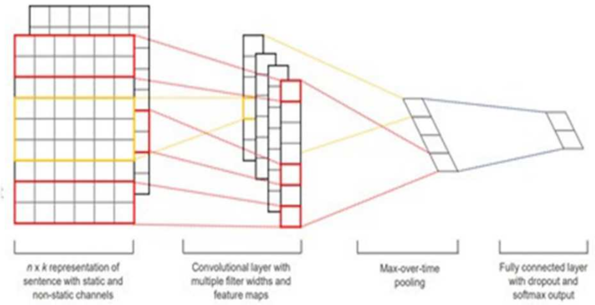

Fig. 2. Model architecture with two channels for an example sentence

III. Result and Discussion

Tests conducted using 25,000 training data with 12,500 positive data and 12,500 negative data. Then tested using 1441 documents. This test is carried out 5 epochs because with a large dataset this model has 2,620,801 parameters that can be trained so that it takes a long time and this model has 2 iterations for each epoch. A large number of datasets and parameters trained in 5 epochs is sufficient to maximize the results of this model.

TABLE III. Testing Result

Epoch	**Accuracy**
1	77.43%
2	84.66%
3	85.68%
4	87.18%
5	86.78%

The test carried out is a test that uses test data from Twitter which contains the keyword Dota. The results of the tested data are shown in fig 4.

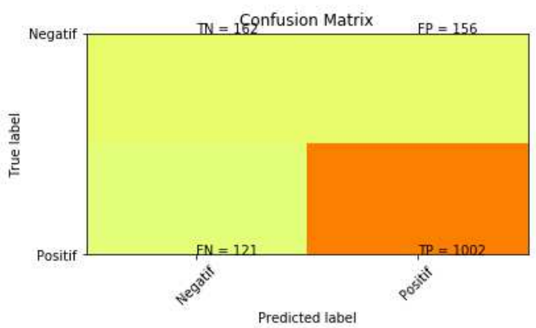

Fig. 3. Confusion Matrix and Accuracy output

Based on fig 4, the precision for positive and negative classes can be calculated as in (6) and (7) below:

$$Precission_{(positive)} = \frac{1{,}002}{1{,}002 + 156} = 0.87 \quad (6)$$

$$Precission_{(negative)} = \frac{162}{162 + 121} = 0.57 \quad (7)$$

Recall or sensitivity is the proportion of true positive cases predicted to be positive correctly. The results of recall calculations can be seen in (8) and (9) below:

$$Recall_{(positive)} = \frac{1{,}002}{1{,}002 + 121} = 0.89 \quad (8)$$

$$Recall_{(negative)} = \frac{162}{162 + 156} = 0.51 \quad (9)$$

The next calculation is the calculation of F1-Score or F-Measure, which is one of the classification evaluation calculations that combines recall and precision. The F1-Score value can be found in (10) and (11).

$$f1 - score_{(positive)} = 2\,x\,\frac{(0.87\,x\,0.89)}{0.87 + 0.89} = 0.88 \quad (10)$$

$$f1 - score_{(negative)} = 2\,x\,\frac{(0.57\,x\,0.51)}{0.57 + 0.51} = 0.54 \quad (11)$$

Testing of the first Convolutional Neural Networks resulted in a precision (precision) of positive 0.87, and negative 0.57. So that the average value of the precision is avg = (0.87+ 0.57) / 2 = 0.72. In addition, the average recall was 0.70 and the average f1-score was 0.71. And support from positive as many as 1,123 documents and negative as many as 318 documents. So that the total documents classified in the first experiment were 1,441 documents.

The last one is calculating the accuracy of the sentiment classifying application. Accuracy can be found by dividing the number of correct classifications with all documents that are classified, as in (12).

$$Accuration = \frac{1{,}002 + 162}{1{,}002 + 121 + 162 + 156}\,x\,100\% = 81\% \quad (12)$$

Based on the calculation, an accuracy value of 81% is obtained from the test data from Twitter containing the keyword Dota. After labeling the data generated from Twitter data related to the online video game Dota 2, taking into account positive and negative sentiments. The factors that affect positive precision are how many are truly positive and false positive, while negative is how many are truly negative and false negative. The factors that influence positive recall are how many are truly positive and false negative, while negative recall is how many are truly negative and false positive. And the factors that affect accuracy involve all true negative, true positive, false negative, and false positive. From the labeling results, it is found that positive sentiment has more numbers than positive sentiment.

The prediction of positive sentiment is more accurate than negative sentiment. Predictions of positive sentiment are more accurate because there are far more positive datasets that were tested than negative.

IV. Conclusion

Based on the test results on Twitter data that has the keyword Dota, it is 81%. With 81% accuracy, this program can be used to analyze Dota Online Video Game sentiments. From the results of this study, the sentiment analysis program using the Convolutional Neural Networks method got 80.4% positive and 19.6% negative results so it can be concluded that Dota 2 Online Video Game has positive sentiments from Twitter users.

Further works, you should check repeatedly manually on the dataset that will be carried out by training and testing so that the processed data is completely clean from noise. It is also recommended to use a dataset for training and do more testing so that the accuracy results obtained are more accurate. The dataset sourced from game reviews is worth considering for more detailed results and information.

In terms of weighting, it is recommended to use features other than GloVe in the weighting and embedding of words, such as Wordnet, FastText, and Word2Vec. Technically it is recommended to use a cloud server to speed up the deep learning process, such as Amazon Web Service AMI for Deep Learning, Google Cloud Platform Virtual Machine for Deep Learning, etc.

REFERENCES

[1] Z. Doshi, S. Nadkarni, K. Ajmera, and N. Shah, "TweerAnalyzer: Twitter Trend Detection and Visualization," in *2017 International Conference on Computing, Communication, Control and Automation, ICCUBEA 2017*, 2018.

[2] S. A. Phand and J. A. Phand, "Twitter sentiment classification using stanford NLP," in *Proceedings - 1st International Conference on Intelligent Systems and Information Management, ICISIM 2017*, 2017, vol. 2017-January, pp. 1–5.

[3] J. Song, S. Lee, and J. Kim, "Inference Attack on Browsing History of Twitter Users Using Public Click Analytics and Twitter Metadata," *IEEE Trans. Dependable Secur. Comput.*, vol. 13, no. 3, pp. 340–364, May 2016.

[4] S. Kamiş and D. Goularas, "Evaluation of Deep Learning Techniques in Sentiment Analysis from Twitter Data," in *Proceedings - 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications, Deep-ML 2019*, 2019, pp. 12–17.

[5] W. B. Zulfikar, M. Irfan, C. N. Alam, and M. Indra, "The comparation of text mining with Naive Bayes classifier, nearest neighbor, and decision tree to detect Indonesian swear words on Twitter," in *2017 5th International Conference on Cyber and IT Service Management (CITSM)*, 2017, pp. 1–5.

[6] P. V. Shah and P. Swaminarayan, "Sentiment Analysis—An Evaluation of the Sentiment of the People: A Survey," in *Lecture Notes on Data Engineering and Communications Technologies*, vol. 52, Springer Science and Business Media Deutschland GmbH, 2021, pp. 53–61.

[7] N. K. Singh, D. S. Tomar, and A. K. Sangaiah, "Sentiment analysis: a review and comparative analysis over social media," *J. Ambient Intell. Humaniz. Comput.*, vol. 11, no. 1, pp. 97–117, Jan. 2020.

[8] M. Biba and M. Mane, "Sentiment analysis through machine learning: An experimental evaluation for Aslbanian," *Adv. Intell. Syst. Comput.*, vol. 235, pp. 195–203, 2014.

[9] J. Serrano-Guerrero, J. A. Olivas, F. P. Romero, and E. Herrera-Viedma, "Sentiment analysis: A review and comparative analysis of web services," *Inf. Sci. (Ny).*, vol. 311, pp. 18–38, Aug. 2015.

[10] M. Irfan, Jumadi, W. B. Zulfikar, and Erik, "Implementation of Fuzzy C-Means algorithm and TF-IDF on English journal summary," in *2017 Second International Conference on Informatics and Computing (ICIC)*, 2017, pp. 1–5.

[11] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, "Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach," in *Procedia Computer Science*, 2018, vol. 132, pp. 679–688.

[12] D. Roy, P. Panda, and K. Roy, "Tree-CNN: A hierarchical Deep Convolutional Neural Network for incremental learning," *Neural Networks*, vol. 121, pp. 148–160, Jan. 2020.

[13] S. Agustin, H. Tjandrasa, and R. V. H. Ginardi, "Deep Learning-based Method for Multi-Class Classification of Oil Palm Planted Area on Plant Ages Using Ikonos Panchromatic Imagery," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 10, no. 6, p. 2200, Dec. 2020.

[14] N. S. Sani, M. A. Rahman, A. A. Bakar, S. Sahran, and H. M. Sarim, "Machine learning approach for Bottom 40 Percent Households (B40) poverty classification," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 8, no. 4–2, pp. 1698–1705, 2018.

[15] X. Bai, "Text classification based on LSTM and attention," in *2018 13th International Conference on Digital Information Management, ICDIM 2018*, 2018, pp. 29–32.

[16] Y. Heryadi and H. L. H. S. Warnars, "Learning temporal representation of transaction amount for fraudulent transaction recognition using CNN, Stacked LSTM, and CNN-LSTM," in *2017 IEEE International Conference on Cybernetics and Computational Intelligence, CyberneticsCOM 2017 - Proceedings*, 2018, vol. 2017-November, pp. 84–89.

[17] M. Groen, N. Kiel, A. Tillmann, and A. Weßel, "Ethical Questions in Digital Game Cultures—A Multi-Perspective View," 2020, pp. 1–10.

[18] A. Matallaoui, N. Hanner, and R. Zarnekow, "Introduction to Gamification: Foundation and Underlying Theories," 2017, pp. 3–18.

[19] M. Irfan, Jumadi, W. B. Zulfikar, and Erik, "Implementation of Fuzzy C-Means algorithm and TF-IDF on English journal summary," in *2017 Second International Conference on Informatics and Computing (ICIC)*, 2017, pp. 1–5.

[20] W. B. Zulfikar, D. R. Ramdania, A. Afif, and W. Uriawan, "A prototype of serious game for lecturing simulation," in *Journal of Physics: Conference Series*, 2019, vol. 1402, no. 6, p. 066103.

[21] B. Meneses-Claudio and A. Roman-Gonzalez, "Study of the Brain Waves for the differentiation of Gamers category between a Newbie and a Hardcore in the game Dota 1," in *Congreso Argentino de Ciencias de la Informatica y Desarrollos de Investigacion, CACIDI 2018*, 2018.

[22] A. Katona *et al.*, "Time to die: Death prediction in dota 2 using deep learning," in *IEEE Conference on Computatonal Intelligence and Games, CIG*, 2019, vol. 2019-August.

[23] A. D. Wijaya, F. Baskara, M. R. Kusuma, R. S. Bernhard, and R. Senjaya, "Ease of use analysis of in-game interface design in DotA 2 for beginners using SAW method," in *Proceedings of 2015 International Conference on Information and Communication Technology and Systems, ICTS 2015*, 2016, pp. 123–126.