International Journal of Electrical and Computer Engineering (IJECE)
Vol. 14, No. 1, February 2024, pp. 698~710
ISSN: 2088-8708, DOI: 10.11591/ijece.v14i1.pp698-710

Emoji's sentiment score estimation using convolutional neural network with multi-scale emoji images

Theerawee Kulkongkoon, Nagul Cooharojananone, Rajalida Lipikorn
Machine Intelligence and Multimedia Information Technology Laboratory, Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, Bangkok, Thailand

Article Info

Article history:

Received May 27, 2023
Revised Aug 29, 2023
Accepted Sep 6, 2023

Keywords:

Convolutional neural network
Emoji classification
Majority voting with probability
Multi-scale images
Sentiment analysis

ABSTRACT

Emojis are any small images, symbols, or icons that are used in social media. Several well-known emojis have been ranked and sentiment scores have been assigned to them. These ranked emojis can be used for sentiment analysis; however, many new released emojis have not been ranked and have no sentiment score yet. This paper proposes a new method to estimate the sentiment score of any unranked emotion emoji from its image by classifying it into the class of the most similar ranked emoji and then estimating the sentiment score using the score of the most similar emoji. The accuracy of sentiment score estimation is improved by using multi-scale images. The ranked emoji image data set consisted of 613 classes with 161 emoji images from three different platforms in each class. The images were cropped to produce multi-scale images. The classification and estimation were performed by using convolutional neural network (CNN) with multi-scale emoji images and the proposed voting algorithm called the majority voting with probability (MVP). The proposed method was evaluated on two datasets: ranked emoji images and unranked emoji images. The accuracies of sentiment score estimation for the ranked and unranked emoji test images are 98% and 51%, respectively.

Corresponding Author:

Rajalida Lipikorn
Machine Intelligence and Multimedia Information Technology Laboratory, Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University
Bangkok, Thailand
Email: rajalida.l@chula.ac.th

1. INTRODUCTION

Recently, social media can be considered as part of our everyday life that allows people to express their feelings with freedom of choices, such as opinion, emotion, experience sharing, and so on. The number of social media users is growing significantly since many users use social media as a new communication channel to exchange information with each other. Their posts on social media can be either private or public depending on their preference in many different formats, such as text, images, voice, and video. The contents in social media can be either user-generate-content (UGC) or multi-format content, such as text with images, text with emojis, images with emojis, and text with images and emojis. Thus, there must be lots of information within one content that is quite difficult to analyze the feeling of a person who posted the content using traditional methods. Many researchers are interested in analyzing the sentiment from the contents and classifying whether the sentiment is positive, neutral, or negative. Emojis are graphical symbols that have been widely used to represent emotion in social media besides texts and images. Most of the existing emojis have been ranked and sentiment scores have been assigned to them; however, many new emojis have been released and they have never been ranked yet. Several existing methods rank an emoji based on the entire content or the name of an emoji, thus if the content or the name does not contain any word that represents emotion then sentiment of that emoji will be classified as neutral which might not always be true. It should be better if a new emoji can be classified as one of the ranked emojis and use its sentiment score to estimate the sentiment score of a new emoji. This sentiment score can then be used for sentiment analysis.

Research on sentiment analysis has been focused on developing new methods for analyzing sentiment from emojis. Several methods that enhanced sentiment analysis using emojis have been proposed. Alzubaidi *et al.* [1] proposed sentiment analysis on social media posts using machine learning algorithms with emojis to classify emotions into neutral and negative categories. Zhao *et al.* [2] developed a machine learning-based approach to perform sentiment analysis from emojis for different social platforms. Novak *et al.* [3] investigated the sentiment of emojis, analyzed their use in different contexts and cultures. Zhang *et al.* [4] used deep learning for sentiment analysis on large-scale data sets that achieved state-of-the-art results. Eisner *et al.* [5] proposed Emoji2vec, a method for learning emoji representations from their description, which could potentially improve sentiment analysis of new emojis. Additionally, Sun *et al.* [6] developed the sentiment analysis method that incorporated both emojis and text information for more accurate results. Barbieri *et al.* [7] also introduced a multi-modal sentiment analysis method that combined visual and textual features of emojis to improve the performance of sentiment analysis. Their research offered insights into the use of emojis for sentiment analysis and provided potential solutions for analyzing sentiment in multi-format content. A novel approach called Chinese emoji-embedding long short-term memory (CEmo-LSTM) model based on bi-directional LSTM (Bi-LSTM), that integrated emojis into sentiment analysis algorithms, achieved high accuracy in analyzing online Chinese texts, particularly during the corona virus disease 2019 (COVID-19) pandemic. By combining emojis and emoji usage with the sentiment analysis model, the model can effectively handle emotion mining tasks, providing a promising way to enhance sentiment analysis in micro-texts and online conversations [8]. Chauhan *et al.* [9] presented a deep learning-based framework for detecting sarcasm in text, which incorporated emojis as an important feature. The multitask framework treated sarcasm as the primary task while considering emotion and sentiment analysis as auxiliary tasks. The SEEmoji MUStARD dataset was used for evaluation, and the proposed framework achieved state-of-the-art performance, outperforming existing methods in terms of accuracy and precision in sarcasm detection. Lou *et al.* [10] introduced the EA-Bi-LSTM model, which incorporated emojis into sentiment analysis for Chinese microblog posts. The model employed an attention mechanism to weigh words based on associated emojis that can effectively capture the influence of emojis on sentiment polarity. The model's effectiveness was compared to the baseline models and the results highlighted potential applications in enhancing sentiment analysis accuracy.

Besides, the new methods for evaluating emoji sentiment lexica and emoji resources have been presented. Milagros *et al.* [11] presented a gold standard dataset of manually annotated English Twitter data that can be used for evaluating emoji sentiment lexica generated from various online resources. The proposed unsupervised methodology involved correlation analysis between sentiment scores from different resources and the gold standard dataset. The results provide insights into the quality of emoji sentiment lexica and the potential for improving emoji-based sentiment analysis. Eunhee *et al.* [12] explored the impact of emojis on user engagement in brand-related user-generated content on Instagram by analyzing a dataset of 1,000 user generated content (UGC) posts. Many factors like the presence, number, and type of emojis used, as well as the valence of accompanying text were examined and it was found that emojis have a positive influence on consumer engagement, with emotional emojis being particularly effective. However, an excessive use of emojis may not lead to additional benefits and could overwhelm the audience.

Other methods for understanding emojis in communication and language processing have been proposed as well. Allan and Budd [13] investigated how alexithymia affects verbal and non-verbal expressivity in text messaging by focusing on the use of words and emojis. Their research suggested that emojis can compensate for verbal deficits in individuals with alexithymia, potentially helping them overcome communication challenges in computer-mediated settings. Scheffler *et al.* [14] examined the processing of emojis in sentence comprehension and their ability to replace words without affecting comprehension. Participants' reading times were measured during self-paced reading experiments, and their research demonstrated that emojis can be easily integrated into sentence interpretation, though certain types of emojis may exhibit visual ambiguity. Pfeifer *et al.* [15] investigated the influence of facial emojis on the emotional significance of text messages and subsequent text processing. They concluded that facial emojis can significantly impact perceived sender emotion and facilitate downstream text processing. However, the subjective nature of facial emojis may lead to varied interpretations of emotional meaning.

There are also other research topics on emojis, such as emoji prediction and sense disambiguation. Shardlow *et al.* [16] developed a semantic network for emojis to examine their multiple meanings or polysemous nature by analyzing a corpus of tweets and using a deep learning-based disambiguation algorithm. The proposed network identified separable information among different classes of emoji senses, paving the way for improving emoji prediction and disambiguation in natural language processing systems.

Lee *et al.* [17] introduced a deep neural network model called the MultiEmo that combined emoji prediction and emotion detection tasks. The model was evaluated using two datasets: the Twitter dataset for emoji prediction and the GoEmotion dataset for emotion detection. MultiEmo achieved superior performance compared to existing models, demonstrating its effectiveness in predicting emoji and emotion labels. The evaluation results provide insights into the relationship between emojis and emotions, offering interpretability for the prediction results and shedding light on the decision-making process of the model. Kaye *et al.* [18] investigated the impact of emojis on word processing and recall. Through self-paced reading experiments, the experiments examined whether emojis can replace words without affecting sentence comprehension and whether they facilitate the retrieval of complete lexical entries. The findings indicate that emojis can be integrated into sentence interpretation, but their use may exhibit visual ambiguity. Additionally, familiarity with emojis may influence processing efficiency.

Recently, the sentiment of an emoji can be analyzed from the name of an emoji or the content of the post; however, this sentiment analysis might not be appropriate if the name of an emoji does not contain any emotion or if the content contains a variety of emotions. An alternative method proposed in this research is to estimate the sentiment score of an emoji from its image. The main ideas are to classify a new emoji as belonging to the class of the most similar ranked emoji image and estimate the sentiment score from the classification results. Several methods have been proposed for image classification and convolutional neural network (CNN) is among the most widely used neural networks for image classification due to its ability to automatically extract features from images. The performances of CNNs were evaluated in various image classifications on the Modified National Institute of Standards and Technology database (MNIST) and the Canadian Institute for Advanced Research (CIFAR) CIFAR-10 datasets [19]–[21]. Additionally, the comparison between different CNN models, including AlexNets, GoogLeNet, and ResNet50, was discussed with testing results on ImageNet, CIFAR-10, and CIFAR-100 datasets [22], [23] The use of CNNs for medical image classifications, such as diabetic retinopathy analysis and wound image classification have been explored in several literatures [24]–[26]. Moreover, data augmentation techniques that were used to increase and balance the size of a dataset including transfer learning, web data augmentation, and different image transformation methods were also proposed [27], [28].

Burnik and Kneževic [29] utilized a CNN to classify emoticons with high accuracy despite a small dataset. Elastic transformations were employed for artificial dataset expansion, and background noise was added to enhance dataset variance and accuracy. They also suggested that additional improvements in CNN models could be achieved by collecting more data and exploring the use of generative adversarial networks for generating emoticons. Akter *et al.* [30] presented a system for detecting and categorizing hand-drawn emojis into eight distinct classes using a CNN model. A local dataset of 4,000 images was generated and 97% accuracy can be achieved. Their research provided valuable insights into image recognition and classification for hand-drawn emojis and had potential applications in social media platforms.

Noord *et al.* [31] proposed a method for enhancing image classification performance by utilizing both scale-variant and scale-invariant features with multi-scale images. Their approach was evaluated on a dataset of digital photographic reproductions of print artworks. The multi-scale convolutional neural network outperformed the single-scale CNN in terms of classification accuracy. They also suggested to incorporate a scale-invariant and scale-variant representation in CNNs to improve image recognition performance. Liu *et al.* [32] have found that deep CNNs applied to small datasets like CIFAR-10 could lead to overfitting. They proposed the visual geometry group (VGG) VGG-16 network which is the modification of VGG that could prevent overfitting and achieved an error rate of 8.45%. They also suggested that deep convolutional neural networks (D-CNNs) can be used for small datasets with appropriate modifications. He *et al.* [33] examined the training procedure refinements for image classification models using an ablation study to evaluate the impact of these refinements and concluded that combining them could improve the performance of various CNN models. Blot *et al.* [34] proposed a modification to the standard convolutional block of CNN to transfer more information layer by layer while maintaining invariance which could be achieved by using a MaxMin strategy with modified activation function before pooling to exploit positive and negative high scores in convolution maps. The modified CNN outperformed the standard CNN on two classic datasets which are MNIST and the CIFAR-10.

According to the existing literatures, sentiment score of an emoji has never been estimated from an emoji image. This paper thus proposes a new method for sentiment score estimation of an unranked emoji from multi-scale images using the convolutional neural network together with the proposed majority voting with probability (MVP) algorithm. An unranked emoji is classified by CNN in different scales and the sentiment score obtained from the classification of each scale is then used to estimate the sentiment score of an unranked emoji by the proposed MVP algorithm. The remaining of this paper is organized as follows. Section 2 presents an overview of the proposed method. Results and discussions are presented in Section 3. Section 4 draws conclusions and presents future work.

2. METHOD

Sentiment score of an unranked emoji is commonly estimated by classifying it as one of the most similar emojis from its name or Unicodes and use the sentiment score of a ranked emoji but this can cause the classification to be imprecise because some emoji names or Unicodes do not express any emotion, and they sometimes use different names even though they express the same emotion. This paper proposes a new method that can overcome these problems by using CNN and MVP algorithm to classify and estimate the sentiment score of an unranked emoji. In order to classify emojis from their images, emoji images were extracted from two sources [3], [35] before passing them through the proposed method. Figure 1 shows the proposed method that is divided into three steps: data collection, pre-processing, emoji image classification, and sentiment score estimation.

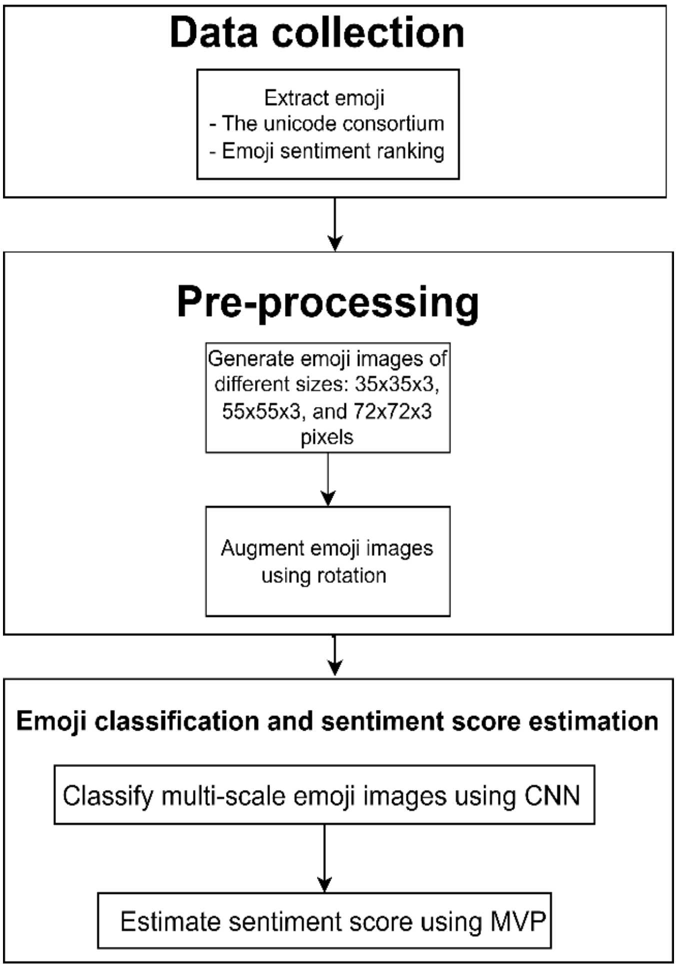
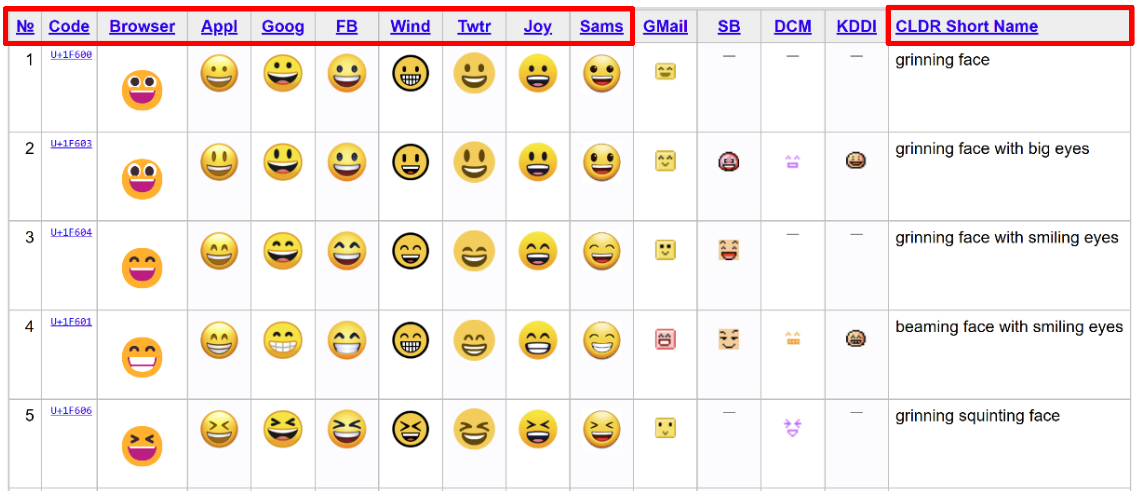


Figure 1. Flowchart of the proposed emoji classification and sentiment score estimation method

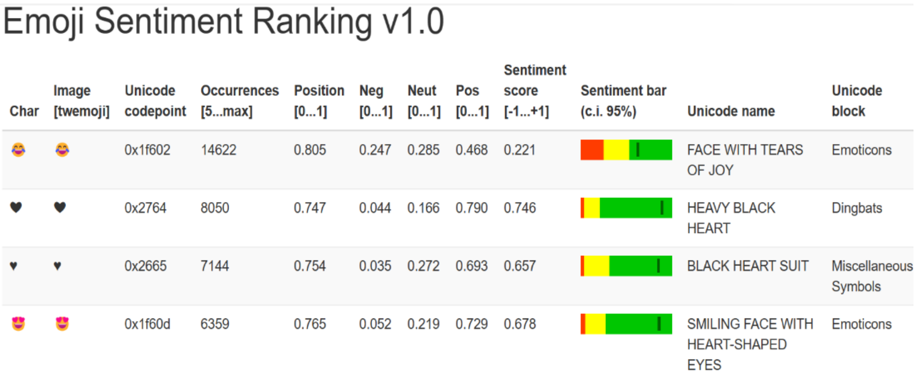
2.1. Data collection

Data collection was performed by extracting data from two different sources: the Unicode consortium website [35] and the emoji sentiment ranking [3]. Emoji images were extracted from the Unicode consortium website using Python and the Beautiful soup library [35] and were stored in .csv format. The web-derived emoji images are version 13.1 which contains approximately 1,816 classes of emojis, and each class contains seven emoji images from seven platforms which are Windows, Facebook, Apple, Twitter, Google, JoyPixels, and Samsung. For example, ,  are emojis from the same class but different platforms, the first emoji image is from iOS and the second one is from Windows platforms. Sample emoji images from different platforms are shown in Figure 2. These emoji images were used for emoji classification. On the other hand, sentiment scores of the ranked emojis were obtained from the emoji sentiment ranking [3] which contains the sentiment scores and polarities of 752 emojis as shown in Figure 3.



№	Code	Browser	Appl	Goog	FB	Wind	Twtr	Joy	Sams	GMail	SB	DCM	KDDI	CLDR Short Name
1	U+1F600										—	—	—	grinning face
2	U+1F603													grinning face with big eyes
3	U+1F604											—	—	grinning face with smiling eyes
4	U+1F601													beaming face with smiling eyes
5	U+1F606										—		—	grinning squinting face

Figure 2. The content of the Unicode consortium [35]



Emoji Sentiment Ranking v1.0

Char	Image [twemoji]	Unicode codepoint	Occurrences [5...max]	Position [0...1]	Neg [0...1]	Neut [0...1]	Pos [0...1]	Sentiment score [-1...+1]	Sentiment bar (c.i. 95%)	Unicode name	Unicode block
		0x1f602	14622	0.805	0.247	0.285	0.468	0.221		FACE WITH TEARS OF JOY	Emoticons
		0x2764	8050	0.747	0.044	0.166	0.790	0.746		HEAVY BLACK HEART	Dingbats
		0x2665	7144	0.754	0.035	0.272	0.693	0.657		BLACK HEART SUIT	Miscellaneous Symbols
		0x1f60d	6359	0.765	0.052	0.219	0.729	0.678		SMILING FACE WITH HEART-SHAPED EYES	Emoticons

Figure 3. The content of emoji sentiment ranking [3]

2.2. Pre-processing

Emoji data used for evaluation consist of 1,816 classes of emoji images obtained from seven platforms that are the most widely used. Figure 2 shows sample color emoji images in *.png* format with the size of 72x72x3 pixels. However, some emojis do not express any emotion and they have not been included in the emoji sentiment ranking [3], thus only emoji images that are in the sentiment ranking were extracted from the Unicode consortium website by using Unicode to match them with those in the sentiment ranking. As a result, there are 613 classes of emoji images that could be matched with emojis in the sentiment ranking. The original emoji images of these 613 classes were cropped to two-fourths and three-fourths of the original size, from the center of an image to generate three multi-scale image datasets of size 35x35x3 pixels, 55x55x3 pixels, and 72x72x3 pixels as shown in Figures 4(a), 4(b), and 4(c), respectively. Emoji images of each class were then augmented to generate training data by rotating each image with the increment of 15 degrees as shown in Figures 5(a), 5(b), and 5(c), respectively. As a result, each class contains 161 emoji images.

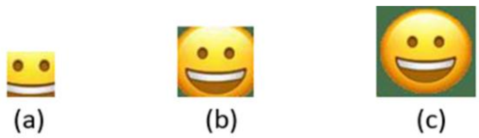


Figure 4. Samples of emoji images in three different sizes: (a) 35x35x3 pixels, (b) 55x55x3 pixels, and (c) 72x72x3 pixels

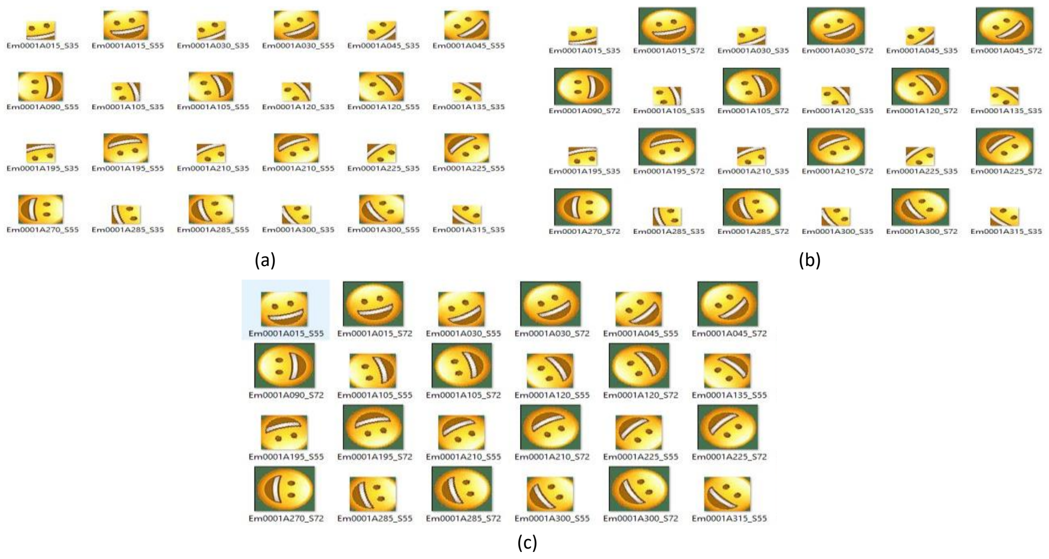


Figure 5. Samples of augmented emoji images: (a) for 35x35x3 and 55x55x3 pixels, (b) for 35x35x3 and 72x72x3 pixels, and (c) for 55x55x3 and 72x72x3 pixels

2.3. Emoji classification and sentiment score estimation

Emoji classification and sentiment score estimation was performed in two steps: emoji image classification and sentiment score estimation. In the first step, an unranked emoji was classified using the convolutional neural network with multi-scale images that could effectively classified emojis based on their visual features. In the second step, sentiment score estimation was performed using the proposed MVP algorithm that yielded better estimation especially when an emoji was classified as belonging to multiple classes.

2.3.1. Emoji classification using convolutional neural network

The main advantages of CNN are that the architecture of CNN can be modified to suit any particular problems [36]–[41] and its ability to automatically extract features from images. In this paper, CNN was used to classify unranked emojis as one of the most similar ranked emojis from emoji images instead of their names or Unicodes. The architecture of CNN used for emoji classification consisted of one input layer, three convolutional layers with max pooling operations, followed by two fully connected layers and one output layer. The first convolutional layer and the second convolutional layer consisted of 256 filters of size 7x7 pixels, the last convolutional layer consisted of 128 filters of size 3x3 pixels, while all max pooling operations used a 2x2 window and stride 2. For fully connected layers, there were two dense layers with 4,096 nodes and 613 nodes, respectively.

The CNN was trained using the batch size of 16, 70 epochs with early stopping to prevent overfitting, categorical cross entropy as loss function, Adam optimization algorithm as learning algorithm with the learning rate of 0.0001, and SoftMax as the activation function for the output layer as shown in Figure 6. The classification of each unranked emoji was repeated seven times with seven multi-scale datasets: three times for the data set of each individual scales, three times for three data sets obtained from the combinations of two scales, and one time for the data set obtained from the combination of three scales. The classification of each iteration returned the class of the most similar emoji with the probability of being in that class as the output which was used to estimate the sentiment score of an unranked emoji in the next step.

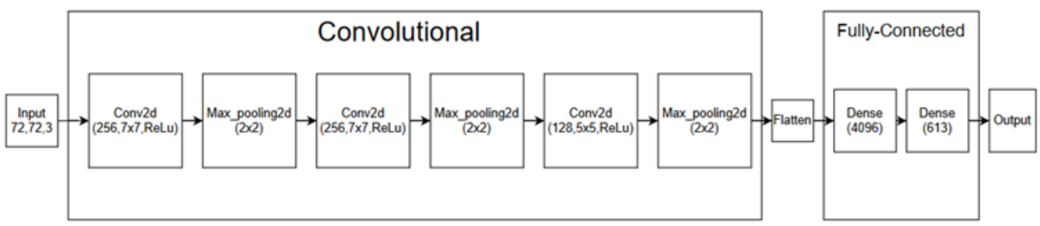


Figure 6. The CNN architecture for emoji classification

2.3.2. Sentiment score estimation using the majority voting with probability algorithm

After obtaining the classification results from seven multi-scale datasets, the sentiment scores of the classes of an unranked emoji in seven scales were used to improve the sentiment score estimation of an unranked emoji by the proposed MVP algorithm. The proposed MVP algorithm estimated the sentiment score of an unranked emoji in two steps. The first step estimated the sentiment score of an unranked emoji by using the sentiment score of a class that has majority vote among seven classification results. The majority vote means the sentiment score of the class that is returned as the output with the highest frequency is assigned as sentiment score of an unranked emoji. However, there might be cases when there was no majority, i.e., the frequencies of two or more classes were equal. This problem was solved in the second step by returning the sentiment score of the class with the highest probability of being in that class as the sentiment score of an unranked emoji. The proposed MVP algorithm is as shown in Algorithm 1:

Algorithm 1. The propose MVP algorithm

```
If FreqClass is maximum and unique
       return sentiment score of that class as sentiment score of an unranked emoji
else
       return sentiment score of a class with the highest probability of being in that
class as sentiment score of an unranked emoji
```

where *FreqClass* is frequency of the classification results

3. RESULTS AND DISCUSSION

The proposed method was evaluated on seven datasets of multi-scale emoji images with seven different training sets of multi-scale images. The first training set (S1) consisted of 613 classes of emoji images in their original size (72x72x3 pixels) with 127 images in each class for the total of 77,851 images. The second training set (S2) consisted of 613 classes of 55x55x3 emoji images with 127 images in each class for the total of 77,851 images. The third training set (S3) consisted of 613 classes of 35x35x3 emoji images with 127 images in each class for the total of 77,851 images. The fourth training set (S4) consisted of 613 classes of the combination of 72x72x3 emoji images and 55x55x3 emoji images with 254 images in each class for the total of 155,702 images. The fifth training set (S5) consisted of 613 classes of the combination of 72x72x3 emoji images and 35x35x3 emoji images with 254 images in each class for the total of 155,702 images. The sixth training set (S6) consisted of 613 classes of the combination of 55x55x3 emoji images and 35x35x3 emoji images with 254 images in each class for the total of 155,702 images and the last training set (S7) consisted of 613 classes of the combination of 72x72x3 emoji images with 55x55x3 emoji images and 35x35x3 emoji images with 381 images in each class for the total of 233,553 images.

The test images were divided into two sets: the first set consisted of the ranked emoji images and the second set consisted of the unranked emoji images. The first set was used to evaluate the performance of the proposed method which was then used to estimate the sentiment scores of unranked emojis in the second set. The first test set consisted of 613 emoji images that were randomly extracted from each class of the training set. The classifications of the test images in the first set were performed seven times using the same test images and the same CNN model with seven different training sets mentioned earlier. Thus, each iteration contained seven classification results with sentiment scores, and the final sentiment score was obtained by applying the proposed MVP algorithm to the sentiment scores obtained from seven classifications with multi-scale training sets.

Figures 7 to 10 show examples of possible classification results that were obtained where the first column shows the test emoji image, the second column shows the classification results, the third column shows the Unicode common locale data repository (CLDR) short name of the class, and the last column shows the multi-scale classification results. Figure 7 is an example of an emoji that was classified correctly by the proposed method on all seven datasets with 100% accuracy. Figure 8 shows an example of an emoji that was classified as belonging to three different classes. In this case, the proposed MVP method was applied to determine the result, i.e., an emoji was finally classified as belonging to the *face with tears of joy* class with the highest probability of being in this class with the majority vote from 5 out of 7 votes. Figure 9 shows another example of the test image that is similar to the training images from many different classes, but the proposed MVP method can correctly classify it as belonging to the *cat with wry smile* class from its highest probability. The sentiment score of the class was returned as the sentiment score of the test image.

However, there are some cases that the proposed method gave the wrong results as shown in Figure 10. In Figure 10, the test image was classified as belonging to the *angry face* when it actually belongs to the *pouting face*. The reason that the classification result was incorrect is because the dominant features of the test image, which are the eyes and the mouth, are very similar to those of the training image in the *angry face* class.

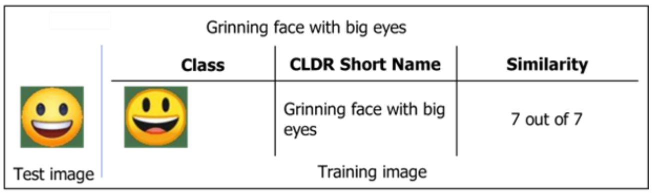


Figure 7. An example of an emoji image that was classified correctly from all scales

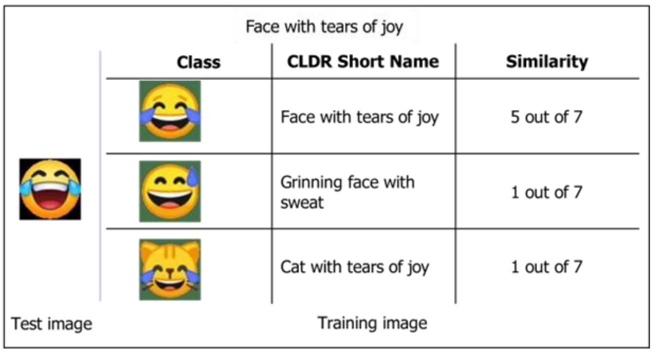


Figure 8. An example of an emoji image that was classified correctly from the majority vote 5 out of 7

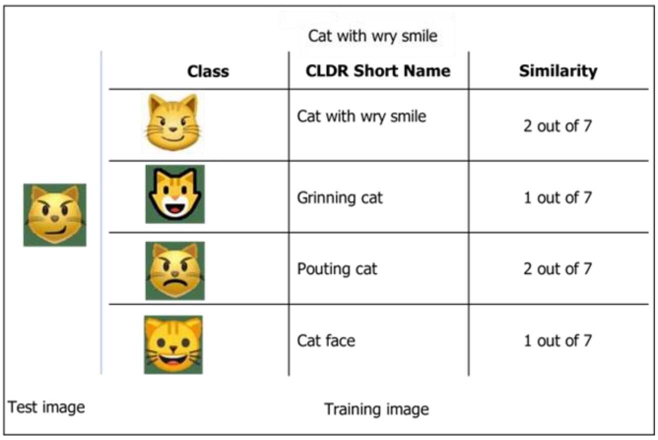


Figure 9. An example of an emoji image that has no majority vote

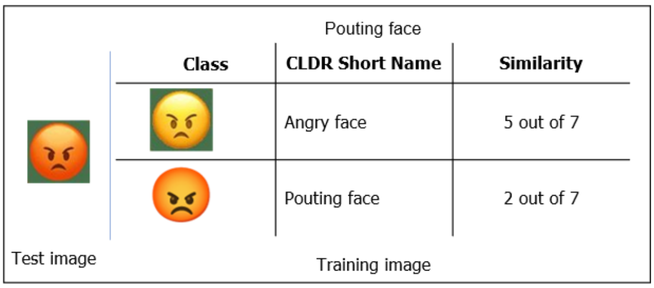


Figure 10. An example of an emoji image that has the majority vote in the wrong class

The classifications and estimations with seven multi-scale data sets were iterated nine times with different test sets. Table 1 shows the accuracies obtained from using each training set where the first seven columns show the accuracies from the classifications and estimations without using MVP algorithm and the last column shows the accuracies from applying MVP algorithm to classification results from seven multi-scale data sets. It can be noticed that the proposed MVP algorithm yields higher accuracies for all iterations and the overall accuracy can be increased to 98.40%.

Table 1. The classification and estimation accuracies from using the ranked emojis as test data

Type Round	S1	S2	S3	S4	S5	S6	S7	MVP
R1	93.15	96.25	96.41	95.6	94.78	97.23	95.43	**98.53**
R2	93.47	97.23	95.92	96.41	95.6	97.55	97.88	**98.53**
R3	93.8	95.11	96.9	95.92	94.29	95.92	95.92	**97.72**
R4	95.76	96.25	96.08	96.41	96.25	97.23	96.57	**98.53**
R5	94.13	96.08	95.6	96.25	94.94	96.57	96.08	**98.21**
R6	91.84	96.25	96.41	94.45	95.92	97.06	96.74	**98.53**
R7	92.5	95.11	96.74	93.8	94.94	96.9	97.55	**98.21**
R8	94.29	96.25	97.23	96.08	96.74	96.9	96.74	**98.53**
R9	93.15	97.88	98.69	96.74	96.25	97.72	97.55	**98.86**
Average	93.57	96.27	96.66	95.74	95.52	97.01	96.72	**98.40**

The second test set consisted of 42 unranked emoji images that were not in any class of the training sets. The images in the second test set were classified using the same CNN architecture as the first test set and the classification and estimation were performed in the exact same way as the first test set. However, it is not possible to measure the actual accuracy because there is no ground truth data for the second test set; therefore, the accuracy of the proposed method was estimated by visual assessment. Figure 11 shows an example of an emoji that was classified correctly by the classifications from seven multi-scale data sets and the sentiment score of the *face with tears of joy* class was assigned as the sentiment score of the test image. However, there are some cases when the classification results are not all agreed. Figure 12 shows an example of an emoji that was classified as belonging to three different classes, but the proposed MVP method classified it as belonging to the *neutral face* class with the highest probability and majority vote which seems reasonable. Figure 13 shows an example of an emoji that was classified as belonging to the *knocked-out face* class with the highest probability and the majority vote from 3 out of 7. This might be because both the test image and the *knocked-out face* training image have eyebrows and have eyes that are different from eyes of other training images. Figure 14 is another good example of how the proposed MVP algorithm can be used to classify and estimate the sentiment score of an unranked emoji when CNN classified it as belonging to totally different classes from seven multi-scale training data sets and the final result was obtained by using the probability because there was no majority vote.

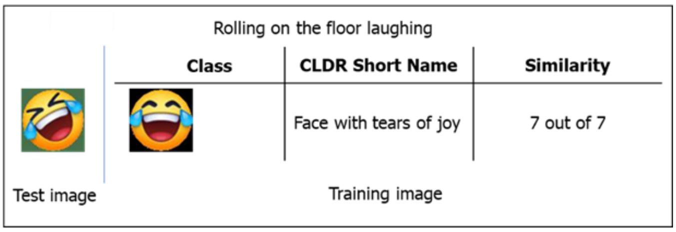


Figure 11. An example of an unranked emoji image that was classified into the same class for all scales

Table 2 shows the classification and estimation results that were evaluated by visual assessment. The overall accuracy for each training set was calculated from the average of the accuracies obtained from repeating the experiments nine times. The features that were used for visual assessment include physical appearance, emotion, and color. It can be noticed that the classification results from using the proposed MVP method are about the same for all experiments whereas the classification results from using CNN without the proposed MVP algorithm depend on the training data.

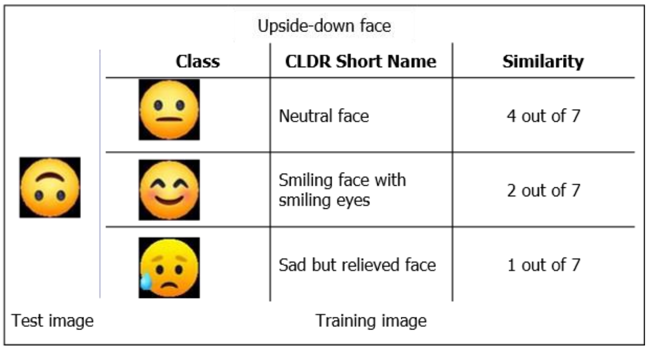


Figure 12. An example of an unranked emoji image that was classified by using majority vote and probability from the proposed MVP algorithm

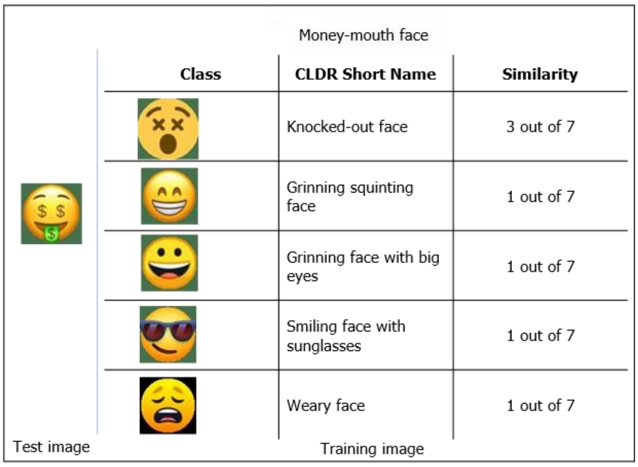


Figure 13. An example of an unranked emoji image that was classified by using majority vote and probability from the proposed MVP algorithm

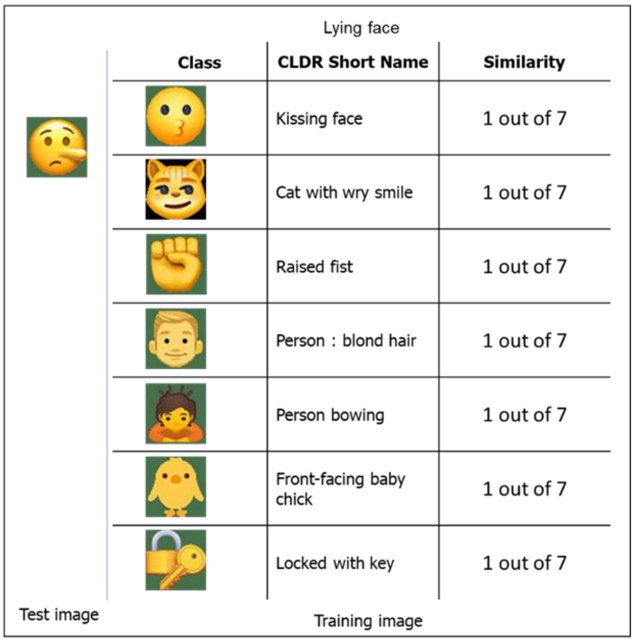


Figure 14. An example of an unranked emoji image that was classified as belonging to seven different classes

Table 2. The classification and estimation results from using unranked emojis as test data

Type Round	S1	S2	S3	S4	S5	S6	S7	MVP
R1	50	52.38	50	50	42.86	50	**57.14**	**57.14**
R2	50	42.86	42.86	54.76	57.14	**59.52**	50	45.24
R3	47.62	45.24	**57.14**	54.76	54.76	45.24	45.24	50
R4	57.14	52.38	54.76	54.76	54.76	**59.52**	54.76	57.14
R5	**57.14**	47.62	54.76	50	**57.14**	47.62	42.86	47.62
R6	52.38	**61.9**	47.62	**61.9**	54.76	59.52	57.14	47.62
R7	38.1	50	45.24	**52.38**	50	50	38.1	**52.38**
R8	57.14	**61.9**	57.14	52.38	**61.9**	45.24	52.38	57.14
R9	47.62	45.24	42.86	**50**	42.86	42.86	47.62	45.24
Average	50.79	51.06	50.26	**53.44**	52.91	51.06	49.47	51.06

4. CONCLUSION

Nowadays, social media has become an important communication channel and part of our daily lives where people use social media to express their feelings, opinions or share experiences. There are various forms of content that are posted on social media, such as text, images, video, and emoji. This paper proposes a new method that can estimate the sentiment score of any new emoji by classifying it from its multi-scale images using convolutional neural network and estimating the sentiment score using the proposed MVP algorithm. CNN is used to solve the classification problem when the name of an emoji does not indicate any emotion; whereas the MVP algorithm is used to improve the classification performance when the test emoji image is classified as belonging to more than one class. The performance of the proposed method was evaluated on two test sets. The first test set contained 613 ranked emojis that were extracted from the training set and the second test set contained 42 new or unranked emojis. For the first test set, the classification results of the proposed method yield higher accuracies for all seven multi-scale data sets and the overall accuracy increases to 98.40%. For the second test set, the classification results had to be evaluated visually and it can be noticed that the results are acceptable although there are some classification results that might be difficult to determine. For example, the classification result of the new emoji image whose features are so much different from the training images is not easy to determine whether the result is acceptable because it depends on individual judgement. Some emojis express ambiguous emotions which are very subjective. However, the overall accuracy for the second test set based on visual evaluation is approximately 51.06%.

In the future, it would be valuable to enhance classification accuracy by refining the proposed algorithm to better differentiate between visually similar emojis. Additionally, more efforts might be focused on proposing new features that can be used to classify new emoji images and a new method that can optimally estimate sentiment score. This could lead to a more robust emoji classification and sentiment score estimation.

REFERENCES

[1] M. Saleh Alzubaidi and F. Bourennani, "Sentiment analysis of tweets using emojis and texts," in *ACM International Conference Proceeding Series*, ACM, Mar. 2021, pp. 111–114, doi: 10.1145/3459955.3460608.

[2] P. Zhao, J. Jia, Y. An, J. Liang, L. Xie, and J. Luo, "Analyzing and predicting emoji usages in social media," in *The Web Conference 2018 - Companion of the World Wide Web Conference, WWW 2018*, {ACM} Press, 2018, pp. 327–334, doi: 10.1145/3184558.3186344.

[3] P. K. Novak, J. Smailović, B. Sluban, and I. Mozetič, "Sentiment of emojis," *PLoS ONE*, vol. 10, no. 12, Art. no. e0144296, Dec. 2015, doi: 10.1371/journal.pone.0144296.

[4] L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 4, Mar. 2018, doi: 10.1002/widm.1253.

[5] B. Eisner, T. Rocktäschel, I. Augenstein, M. Bošnjak, and S. Riedel, "emoji2vec: Learning emoji representations from their description," in *EMNLP 2016 - Conference on Empirical Methods in Natural Language Processing, Proceedings of the 4th International Workshop on Natural Language Processing for Social Media, SocialNLP 2016*, Association for Computational Linguistics, 2016, pp. 48–54, doi: 10.18653/v1/w16-6208.

[6] L. Sun *et al.*, "Multimodal emotion recognition and sentiment analysis via attention enhanced recurrent model," in *MuSe 2021 - Proceedings of the 2nd Multimodal Sentiment Analysis Challenge, co-located with ACM MM 2021*, ACM, Oct. 2021, pp. 15–20, doi: 10.1145/3475957.3484456.

[7] F. Barbieri, M. Ballesteros, F. Ronzano, and H. Saggion, "Multimodal emoji prediction," in *NAACL HLT 2018 - 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, Association for Computational Linguistics, 2018, pp. 679–686, doi: 10.18653/v1/n18-2107.

[8] C. Liu *et al.*, "Improving sentiment analysis accuracy with emoji embedding," *Journal of Safety Science and Resilience*, vol. 2, no. 4, pp. 246–252, Dec. 2021, doi: 10.1016/j.jnlssr.2021.10.003.

[9] D. S. Chauhan, G. V. Singh, A. Arora, A. Ekbal, and P. Bhattacharyya, "An emoji-aware multitask framework for multimodal sarcasm detection," *Knowledge-Based Systems*, vol. 257, Art. no. 109924, Dec. 2022, doi: 10.1016/j.knosys.2022.109924.

[10] L. O. U. Yinxia, Y. Zhang, L. I. Fei, T. Qian, and J. I. Donghong, "Emoji-based sentiment analysis using attention networks,"

ACM Transactions on Asian and Low-Resource Language Information Processing, vol. 19, no. 5, pp. 1–13, Jun. 2020, doi: 10.1145/3389035.
[11] M. Fernández-Gavilanes, E. Costa-Montenegro, S. García-Méndez, F. J. González-Castaño, and J. Juncal-Martínez, "Evaluation of online emoji description resources for sentiment analysis purposes," *Expert Systems with Applications*, vol. 184, Art. no. 115279, Dec. 2021, doi: 10.1016/j.eswa.2021.115279.
[12] E. Ko, D. Kim, and G. Kim, "Influence of emojis on user engagement in brand-related user generated content," *Computers in Human Behavior*, vol. 136, Art. no. 107387, Nov. 2022, doi: 10.1016/j.chb.2022.107387.
[13] H. Allan and M. J. Budd, "A case for emojis, more or less: An analysis of word and emoji expressivity in text messaging for high and low alexithymia levels," *Computers in Human Behavior*, vol. 147, Oct. 2023, doi: 10.1016/j.chb.2023.107845.
[14] T. Scheffler, L. Brandt, M. de la Fuente, and I. Nenchev, "The processing of emoji-word substitutions: A self-paced-reading study," *Computers in Human Behavior*, vol. 127, p. 107076, Feb. 2022, doi: 10.1016/j.chb.2021.107076.
[15] V. A. Pfeifer, E. L. Armstrong, and V. T. Lai, "Do all facial emojis communicate emotion? The impact of facial emojis on perceived sender emotion and text processing," *Computers in Human Behavior*, vol. 126, p. 107016, Jan. 2022, doi: 10.1016/j.chb.2021.107016.
[16] M. Shardlow, L. Gerber, and R. Nawaz, "One emoji, many meanings: A corpus for the prediction and disambiguation of emoji sense," *Expert Systems with Applications*, vol. 198, Art. no. 116862, Jul. 2022, doi: 10.1016/j.eswa.2022.116862.
[17] S. E. Lee, D. Jeong, and E. Park, "MultiEmo: Multi-task framework for emoji prediction," *Knowledge-Based Systems*, vol. 242, Art. no. 108437, Apr. 2022, doi: 10.1016/j.knosys.2022.108437.
[18] L. K. Kaye *et al.*, "Exploring the (lack of) facilitative effect of emoji for word processing," *Computers in Human Behavior*, vol. 139, Art. no. 107563, Feb. 2023, doi: 10.1016/j.chb.2022.107563.
[19] T. Guo, J. Dong, H. Li, and Y. Gao, "Simple convolutional neural network on image classification," in *2017 IEEE 2nd International Conference on Big Data Analysis, ICBDA 2017*, IEEE, Mar. 2017, pp. 721–724, doi: 10.1109/ICBDA.2017.8078730.
[20] R. Doon, T. Kumar Rawat, and S. Gautam, "Cifar-10 classification using deep convolutional neural network," in *1st International Conference on Data Science and Analytics, PuneCon 2018 - Proceedings*, IEEE, Nov. 2018, doi: 10.1109/PUNECON.2018.8745428.
[21] R. C. Calik and M. F. Demirci, "Cifar-10 image classification with convolutional neural networks for embedded systems," in *Proceedings of IEEE/ACS International Conference on Computer Systems and Applications, AICCSA*, IEEE, Oct. 2019, doi: 10.1109/AICCSA.2018.8612873.
[22] N. Sharma, V. Jain, and A. Mishra, "An analysis of convolutional neural networks for image classification," *Procedia Computer Science*, vol. 132, pp. 377–384, 2018, doi: 10.1016/j.procs.2018.05.198.
[23] F. Sultana, A. Sufian, and P. Dutta, "Advancements in image classification using convolutional neural network," in *Proceedings - 2018 4th IEEE International Conference on Research in Computational Intelligence and Communication Networks, ICRCICN 2018*, IEEE, Nov. 2018, pp. 122–129, doi: 10.1109/ICRCICN.2018.8718718.
[24] S. Wan, Y. Liang, and Y. Zhang, "Deep convolutional neural networks for diabetic retinopathy detection by image classification," *Computers and Electrical Engineering*, vol. 72, pp. 274–282, Nov. 2018, doi: 10.1016/j.compeleceng.2018.07.042.
[25] J. Zhang, Y. Xie, Q. Wu, and Y. Xia, "Medical image classification using synergic deep learning," *Medical Image Analysis*, vol. 54, pp. 10–19, May 2019, doi: 10.1016/j.media.2019.02.010.
[26] B. Rostami, D. M. Anisuzzaman, C. Wang, S. Gopalakrishnan, J. Niezgoda, and Z. Yu, "Multiclass wound image classification using an ensemble deep CNN-based classifier," *Computers in Biology and Medicine*, vol. 134, Art. no. 104536, Jul. 2021, doi: 10.1016/j.compbiomed.2021.104536.
[27] D. Han, Q. Liu, and W. Fan, "A new image classification method using CNN transfer learning and web data augmentation," *Expert Systems with Applications*, vol. 95, pp. 43–56, Apr. 2018, doi: 10.1016/j.eswa.2017.11.028.
[28] A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in *2018 International Interdisciplinary PhD Workshop, IIPhDW 2018*, IEEE, May 2018, pp. 117–122, doi: 10.1109/IIPHDW.2018.8388338.
[29] K. Burnik and D. Bjelobrk Kneževic, "Using convolutional neural networks for emoticon classification," in *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2019 - Proceedings*, IEEE, May 2019, pp. 1614–1618, doi: 10.23919/MIPRO.2019.8757189.
[30] M. Akter, M. S. Hossain, and K. Andersson, "Hand-drawn emoji recognition using convolutional neural network," in *Proceedings of 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering, WIECON-ECE 2020*, IEEE, Dec. 2020, pp. 147–152, doi: 10.1109/WIECON-ECE52138.2020.9397933.
[31] N. van Noord and E. Postma, "Learning scale-variant and scale-invariant features for deep image classification," *Pattern Recognition*, vol. 61, pp. 583–592, Jan. 2017, doi: 10.1016/j.patcog.2016.06.005.
[32] S. Liu and W. Deng, "Very deep convolutional neural network based image classification using small training sample size," in *Proceedings - 3rd IAPR Asian Conference on Pattern Recognition, ACPR 2015*, IEEE, Nov. 2016, pp. 730–734, doi: 10.1109/ACPR.2015.7486599.
[33] T. He, Z. Zhang, H. Zhang, Z. Zhang, J. Xie, and M. Li, "Bag of tricks for image classification with convolutional neural networks," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2019, pp. 558–567, doi: 10.1109/CVPR.2019.00065.
[34] M. Blot, M. Cord, and N. Thome, "Max-min convolutional neural networks for image classification," in *Proceedings - International Conference on Image Processing, ICIP*, IEEE, Sep. 2016, pp. 3678–3682, doi: 10.1109/ICIP.2016.7533046.
[35] Unicode, "Full emoji list, v13.1." Unicode, http://www.unicode.org/emoji/charts/full-emoji-list.html (accessed Mar. 03, 2021).
[36] P. Kim, *Convolutional neural network*, MATLAB dee. Springer, 2017.
[37] T. Wang, D. J. Wu, A. Coates, and A. Y. Ng, "End-to-end text recognition with convolutional neural networks," in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, 2012, pp. 3304–3308.
[38] Y.-L. Boureau, J. Ponce, and Y. Lecun, *A theoretical analysis of feature pooling in visual recognition*. 2010.
[39] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse Rectifier neural networks," in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, G. Gordon, D. Dunson, and M. Dudík, Eds., in Proceedings of Machine Learning Research, vol. 15. Fort Lauderdale, FL, USA: PMLR, 2011, pp. 315–323.
[40] A. Ghosh, A. Sufian, F. Sultana, A. Chakrabarti, and D. De, "Fundamental concepts of convolutional neural network," in *Intelligent Systems Reference Library*, Springer International Publishing, 2019, pp. 519–567, doi: 10.1007/978-3-030-32644-9_36.
[41] S. Sharma, S. Sharma, and A. Athaiya, "Activation functions in neural networks," *International Journal of Engineering Applied Sciences and Technology*, vol. 04, no. 12, pp. 310–316, May 2020, doi: 10.33564/IJEAST.2020.v04i12.054.

BIOGRAPHIES OF AUTHORS

Theerawee Kulkongkoon received the B.B.A. in business information technology and M.Sc. in information technology from Dhurakij Pundit University, in 2007 and 2011, respectively. He is currently pursuing the Doctoral's degree in computer science and information technology with Chulalongkorn University. His research interests include image processing, natural language processing, and data mining. He can be contacted at email: theerawee.k@student.chula.ac.th.

Nagul Cooharojananone is an associate professor at Chulalongkorn University, Thailand. He received his B.S. degree in computer science from Mahidol University. He received his M. Eng and Ph.D. in information and communication engineering from the University of Tokyo. His research interests include computer vision, multimedia technology, and user behavior analysis. He can be contacted at email: nagul.c@chula.ac.th.

Rajalida Lipikorn received the B.Sc. in applied mathematics and M.Sc. in computer science from California State University, Northridge, U.S.A. in 1987 and 1992, respectively. She received her PhD. in bio-applications and system engineering from Tokyo University of Agriculture and Technology, Tokyo, Japan in 2002. She has been an associate professor at Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University since 2015. Her research interests include data science, computer graphics, medical image processing, modeling, and big data analytics. She can be contacted at email: rajalida.l@chula.ac.th.