

Facial recognition using deep learning

Abdulrazak Yahya Saleh^{a,1,*}, Kirthanaa A/P Jiva Rattinami^{a,2}

^a Faculty of Cognitive Science and Human Development, Universiti Malaysia Sarawak, 94300, Sarawak, Malaysia

¹ ysahabdulrazak@unimas.my*

ABSTRACT

In this article, the researcher presented the results of recognition of four emotional states (happy, sad, angry, and disgust) based on facial expressions. A deep learning method with a Convolutional Neural Network algorithm for recognizing problems has been proven very effective way to overcome the recognition problem. A comparative study is carried out using MUAD3D dataset from Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak for evaluating accuracy performance of this dataset. More discussion is provided to prove the effectiveness of the Convolutional Neural Network in recognition problems.

Keywords:
Classification
Facial Recognition
Convolutional
Neural Network

I. Introduction

Facial expression is one of the non-verbal communication methods to understand the mood/mental state of a person [1][2]. Facial recognition technology work utilizes several measurements and technologies to scan faces, including thermal imaging, 3D face mapping, cataloguing unique features, analyzing geometric proportions of facial features, mapping distance between key facial features, and skin surface texture analysis [3]. The aim of this research, presented in the article, is to recognize four basic emotional states: happy, sad, angry, and disgust based on facial expressions. Convolutional Neural Network can improve the accuracy performance in related tasks and few recent works on facial expression recognition successfully utilize Convolutional Neural Network for feature extraction and inference [4][5][6][7][8].

II. Method

This section reviews the important foundation of developing Convolutional Neural Network and discusses how the algorithm has been utilized for enhancement. In the first part, an introduction of the Convolutional Neural Network has been presented. The second part focuses on the data set used in this research. Finally, the third part focuses on the proposed method.

A. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) one of the most popular techniques used in image recognition and computer vision systems today. The historical roots traced back to the 1980s, when Kunihiko Fukushima proposed a neural network architecture inspired by the feline visual processing system [9][10][11][12][13][14]. The primary aim of Convolutional is to extract features from image where preserves spatial relationship between pixels by learning image features using Small Square of input data [15][16][17]. Added on, Convolutional Neural Network composed of learnable weights and biases. Each neuron receives several inputs, takes a weighted sum over them, pass it through an activation function and responds with an output [18][19]. There are five structures of the Convolutional Neural Network which are convolutional layer used to detect features, non-linear layer, pooling or down-sampling layer decreases the number of weights and controls over-fitting, flattening layers function as set up data for classical neural network, and fully-connected layer is the standard neural network used for classifications [20][21].

Moreover, in the article indicted by [22] stated there are four fundamental functions of the Convolution Neural Network. First, input layers customarily hold the pixel values of the image. Then, Convolutional Neural Network decides the output of neurons, which are connected to local regions of the input through their weight and connected region. Withal verbally expressed that rectified linear unit (ReLU) apply activation function such as sigmoid to the output of the anterior layers. Next, pooling layer function as down-sample gives input dimensionally and reduces the number of



parameters within the activation. Determinately, ANN performs the same obligations and engenders class scores from the activation to be utilized as relegation. This is known as the fully-connected layers functions. The Convolutional Neural Network consists of four main steps; convolutional, subsampling, activation and fully connected. First step functions as a receiver of input signals which is called as convolutional filters. Convolutional is the process, where network tries to label the input signal by referring to what it has learned in the past. Convolutional also was known as translational invariant property.

The second step is subsampling. The input from convolutional layer has been smoother in order to reduce the sensitivity of noise and variations filters. Furthermore, third step known as activation layer where it controls the signal flow from one layer to another layer. The output layer at this step is strongly interconnected with past references which can activate more neurons. Besides, Rectified Linear Unit (ReLU) is known as one of the common functions which uses in training to increase the speed. Finally, in the last step, all the networks will be fully connected where neurons are connected into subsequent layers [23][24][25]. Figure 1 illustrates the Convolutional Neural Network when combined with numerous convolutional and pooling layers.

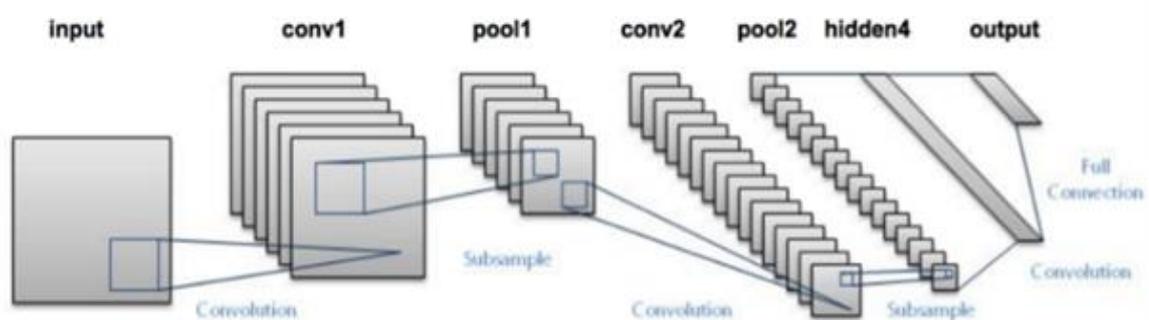


Fig. 1. Fig. 1. The layers are put together in the Convolutional Neural Network algorithm

B. Data Collection and Preparation

The dataset collected from the Faculty of Computer Science and Information Technology, UNIMAS have been divided into two files, test and train. Each image in the folder has the label part of the file name. The train folder contains 80 images of happy emotion and 80 images of sad emotion. The test folder consists of 20 images of happy emotion and 20 images of sad emotions. The same procedures were applied for the emotion of anger and disgust. This section demonstrates the different facial expressions of the participants on low, medium, and high intensities. The context below shows the figures which involves the different types of images under different intensities for different types of emotions, as well as the facial expressions of the participants from the 2 different ethnic groups; the students from the Chinese and Indian ethnic groups. As shown in Figures 2 to Figure 5.



Fig. 2. A: High, B: medium, and C: low intensities of happy emotion by a Chinese student



Fig. 3. A: High, B: medium, and C: low intensities of happy emotion by an Indian student

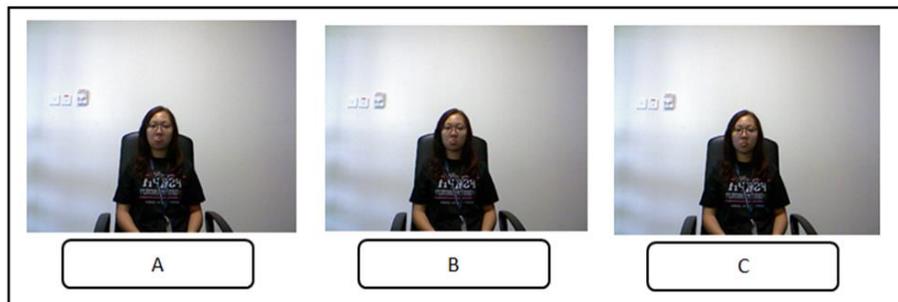


Fig. 4. A: High, B: medium, and C: low intensities of sad emotion by a Chinese student

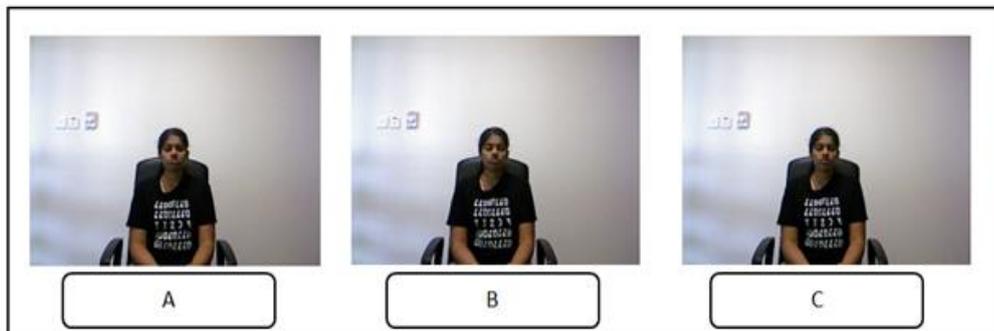


Fig. 5. A: High, B: medium, and C: low intensities of sad emotion by an Indian student

III. Results and Discussion

This section displays the results of the proposed method Convolutional Neural Network. The results of the Convolutional Neural Network are measured in terms of accuracy for the happy, sad, angry, and disgust Emotions. The experiments are run 12 epochs in the training and testing for all datasets with mix intensities of the four emotions. From observation, (Figure 6 to Figure 7) shows the plotted graph of training and validation. The result is the accuracy is higher than the loss. Although, the graph that plotted shows noise, however the accuracy has been achieved.

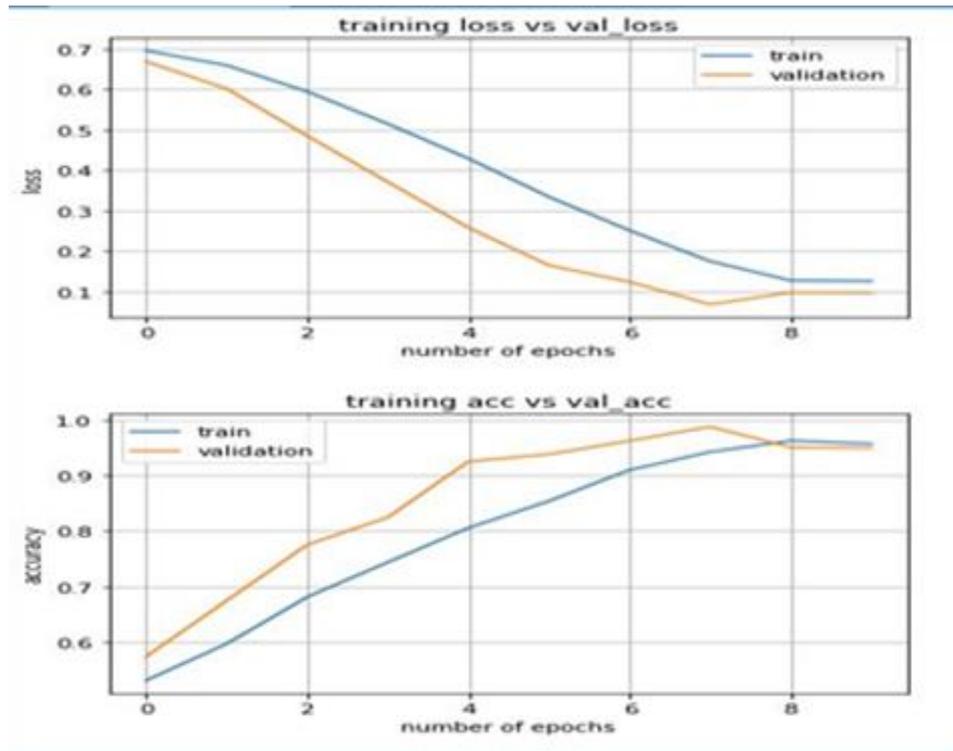


Fig. 6. The plotted graph of training loss, validation loss, training accuracy, and validation accuracy for mix intensities of angry and disgust emotions

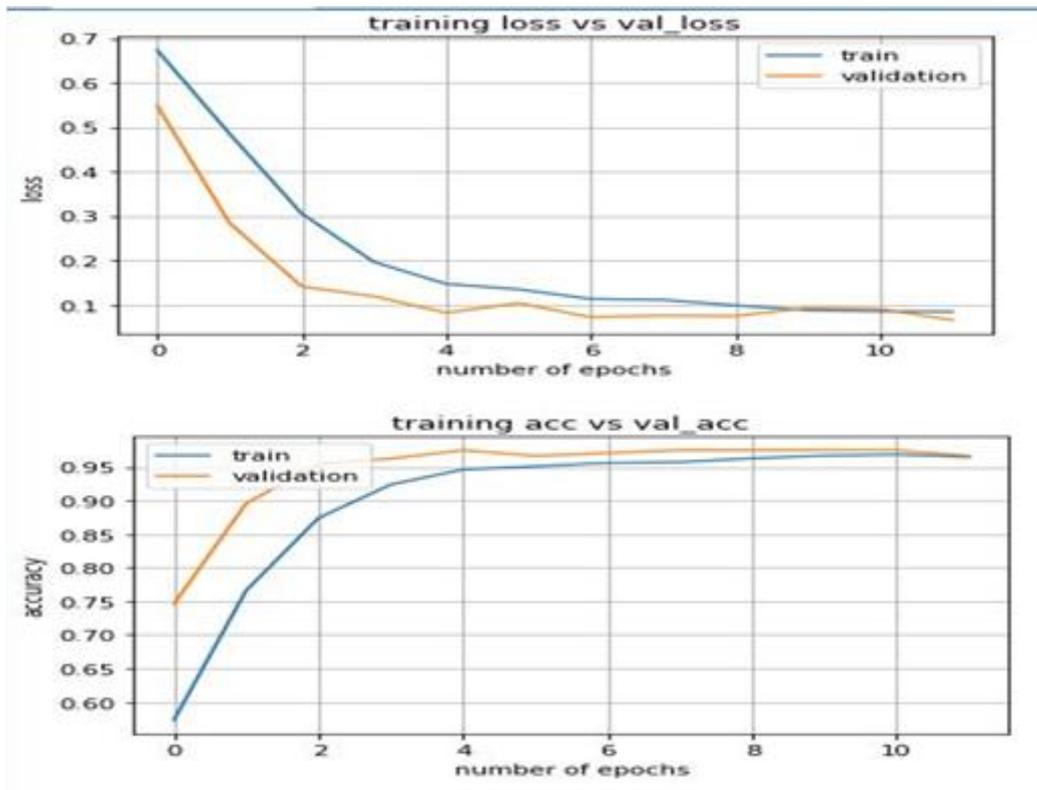


Fig. 7. The plotted graph of training loss, validation loss, training accuracy, and validation accuracy for the mix intensities of happy and sad emotions

The table 1 and 2 shows the results for happy or sad emotions and angry or disgust emotions. The accuracy for the mix intensities of happy or sad emotions is 0.9652 whereas the accuracy for the angry or disgust mix intensities is 0.9896. This indicates the disgust or angry emotions' classification is much better than the happy or sad emotions. Furthermore, training accuracy explained on how the model learns to map the input and output while validation accuracy generalizes the data of Convolutional Neural Networks [26][27][28].

Table 1. Summary of the results for Happy and Sad Emotions

Intensities	Low	Medium	High	Mix
Epoches	18	12	10	12
Time taken to complete (second)	56	60	52	183
Step taken to complete (meter per second)	700	750	656	762
Loss (end point)	0.0509	0.1801	0.0709	0.0841
Accuracy (end point)	0.9820	0.9180	0.9828	0.9652
Validation loss (end point)	0.0857	0.0987	0.1172	0.0655
Validation accuracy (end point)	0.09748	0.9380	0.9496	0.9667

Table 2. Summary of the results for Angry and Disgust Emotions

Intensities	Low	Medium	High	Mix
Epoches	8	10	12	12
Time taken to complete (second)	58	50	48	172
Step taken to complete (meter per second)	731	624	596	718
Loss (end point)	0.1945	0.1224	0.0430	0.0310
Accuracy (end point)	0.9351	0.9593	0.9867	0.9896
Validation loss (end point)	0.1874	0.0967	0.0323	0.0478
Validation accuracy (end point)	0.9501	0.9496	0.9874	0.9875

IV. Conclusion

In this paper, the Convolutional Neural Network was proposed to determine the facial expression recognition for enhancing the accuracy performance. This research has been contributed to predict the student's emotions validated by the accuracy performance. The future research can progress on this topic by using different types of deep learning methods to add more functions and predict the emotions. Moreover, the data must be more than 4000 images to train, validate, and test more accurately to get more accurate performance.

Acknowledgment

This research is supported and funded by RIMC of University Malaysia Sarawak (UNIMAS), under the special Grant Scheme (F04/SpGS/1547/2017).

References

- [1] J. Kumari, R. Rajesh, and K. M. Pooja, "Facial Expression Recognition: A Survey," *Procedia Comput. Sci.*, vol. 58, pp. 486–491, 2015 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1877050915021225>
- [2] I. M. Revina and W. R. S. Emmanuel, "A Survey on Human Face Expression Recognition Techniques," *J. King Saud Univ. - Comput. Inf. Sci.*, Sep. 2018 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1319157818303379>
- [3] F. Becerra-Riera, A. Morales-González, and H. Méndez-Vázquez, "Facial marks for improving face recognition," *Pattern Recognit. Lett.*, vol. 113, pp. 3–9, Oct. 2018 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0167865517301423>
- [4] W. Deng, Y. Fang, Z. Xu, and J. Hu, "Facial landmark localization by enhanced convolutional neural network," *Neurocomputing*, vol. 273, pp. 222–229, Jan. 2018 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0925231217313668>
- [5] Y. Liu, X. Yuan, X. Gong, Z. Xie, F. Fang, and Z. Luo, "Conditional convolution neural network

- enhanced random forest for facial expression recognition,” *Pattern Recognit.*, vol. 84, pp. 251–261, Dec. 2018 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0031320318302516>
- [6] A. T. Lopes, E. de Aguiar, A. F. De Souza, and T. Oliveira-Santos, “Facial expression recognition with Convolutional Neural Networks: Coping with few data and the training sample order,” *Pattern Recognit.*, vol. 61, pp. 610–628, Jan. 2017 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0031320316301753>
- [7] M. McCurrie, F. Beletti, L. Parzianello, A. Westendorp, S. Anthony, and W. J. Scheirer, “Convolutional Neural Networks for Subjective Face Attributes,” *Image Vis. Comput.*, vol. 78, pp. 14–25, Oct. 2018 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0262885618301069>
- [8] D. Sáez Trigueros, L. Meng, and M. Hartnett, “Enhancing convolutional neural networks for face recognition with occlusion maps and batch triplet loss,” *Image Vis. Comput.*, vol. 79, pp. 99–108, Nov. 2018 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0262885618301562>
- [9] S. Brahimi, N. Ben Aoun, and C. Ben Amar, “Boosted Convolutional Neural Network for object recognition at large scale,” *Neurocomputing*, vol. 330, pp. 337–354, Feb. 2019 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0925231218313596>
- [10] C. Du, S. Gao, Y. Liu, and B. Gao, “Multi-focus image fusion using deep support value convolutional neural network,” *Optik (Stuttg.)*, vol. 176, pp. 567–578, Jan. 2019 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0030402618313925>
- [11] Y. Fu and C. Aldrich, “Flotation froth image recognition with convolutional neural networks,” *Miner. Eng.*, vol. 132, pp. 183–190, Mar. 2019 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0892687518305533>
- [12] Y. Seo and K. Shin, “Hierarchical convolutional neural networks for fashion image classification,” *Expert Syst. Appl.*, vol. 116, pp. 328–339, Feb. 2019 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0957417418305992>
- [13] B. B. Traore, B. Kamsu-Foguem, and F. Tangara, “Deep convolution neural network for image recognition,” *Ecol. Inform.*, vol. 48, pp. 257–268, Nov. 2018 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1574954118302140>
- [14] H. Yang *et al.*, “Asymmetric 3D Convolutional Neural Networks for action recognition,” *Pattern Recognit.*, vol. 85, pp. 1–12, Jan. 2019 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0031320318302632>
- [15] P. V. Arun, I. Herrmann, K. M. Budhiraju, and A. Karnieli, “Convolutional network architectures for super-resolution/sub-pixel mapping of drone-derived images,” *Pattern Recognit.*, vol. 88, pp. 431–446, Apr. 2019 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0031320318304217>
- [16] H. Tang, B. Xiao, W. Li, and G. Wang, “Pixel convolutional neural network for multi-focus image fusion,” *Inf. Sci. (Ny.)*, vol. 433–434, pp. 125–141, Apr. 2018 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0020025517311647>
- [17] M. Yan, J. Guo, W. Tian, and Z. Yi, “Symmetric convolutional neural network for mandible segmentation,” *Knowledge-Based Syst.*, vol. 159, pp. 63–71, Nov. 2018 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0950705118302983>
- [18] H. Wu, J. Weng, X. Chen, and W. Lu, “Feedback weight convolutional neural network for gait recognition,” *J. Vis. Commun. Image Represent.*, vol. 55, pp. 424–432, Aug. 2018 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1047320318301445>
- [19] Q. Zhang, M. Zhang, T. Chen, Z. Sun, Y. Ma, and B. Yu, “Recent advances in convolutional neural network acceleration,” *Neurocomputing*, vol. 323, pp. 37–51, Jan. 2019 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0925231218311007>
- [20] A. J. O’Toole, C. D. Castillo, C. J. Parde, M. Q. Hill, and R. Chellappa, “Face Space Representations in Deep Convolutional Neural Networks,” *Trends Cogn. Sci.*, vol. 22, no. 9, pp. 794–809, Sep. 2018 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1364661318301463>
- [21] G. Yao, T. Lei, and J. Zhong, “A review of Convolutional-Neural-Network-based action recognition,” *Pattern Recognit. Lett.*, vol. 118, pp. 14–22, Feb. 2019 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0169254518301463>

- <https://linkinghub.elsevier.com/retrieve/pii/S0167865518302058>
- [22] K. O’Shea and R. Nash, “An Introduction to Convolutional Neural Networks,” pp. 1–11, 2015 [Online]. Available: <http://arxiv.org/abs/1511.08458>
- [23] J. Gu *et al.*, “Recent advances in convolutional neural networks,” *Pattern Recognit.*, vol. 77, pp. 354–377, May 2018 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0031320317304120>
- [24] G. Lin and W. Shen, “Research on convolutional neural network based on improved Relu piecewise activation function,” *Procedia Comput. Sci.*, vol. 131, pp. 977–984, 2018 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1877050918306197>
- [25] Y.-D. Zhang, C. Pan, J. Sun, and C. Tang, “Multiple sclerosis identification by convolutional neural network with dropout and parametric ReLU,” *J. Comput. Sci.*, vol. 28, pp. 1–10, Sep. 2018 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1877750318305763>
- [26] T. Rao, X. Li, H. Zhang, and M. Xu, “Multi-level region-based Convolutional Neural Network for image emotion classification,” *Neurocomputing*, vol. 333, pp. 429–439, Mar. 2019 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0925231218315145>
- [27] A. Savchenkov, A. Davis, and X. Zhao, “Generalized Convolutional Neural Networks for Point Cloud Data,” *2017 16th IEEE Int. Conf. Mach. Learn. Appl.*, pp. 930–935, 2017 [Online]. Available: <http://ieeexplore.ieee.org/document/8260757/>
- [28] N. Sharma, V. Jain, and A. Mishra, “An Analysis Of Convolutional Neural Networks For Image Classification,” *Procedia Comput. Sci.*, vol. 132, pp. 377–384, 2018 [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1877050918309335>