Hajj pilgrimage video analytics using CNN

Md Roman Bhuiyan[1], Junaidi Abdullah[2], Noramiza Hashim[3], Fahmid Al Farid[4], Mohd Ali Samsudin[5], Norra Abdullah[6], Jia Uddin[7]
[1,2,3,4]Faculty of Computing and Informatics, Multimedia University, Cyberjaya, Persiaran Multimedia, 63100, Cyberjaya, Malaysia
[5]University Sains Malaysia, Penang, Malaysia
[6]WSA VENTURE AUSTRALIA (M) SDN BHD, Malaysia
[7]Technology Studies Department, Endicott College, Woosong University, Daejeon, South Korea

Article Info

ABSTRACT

This paper advances video analytics with a focus on crowd analysis for Hajj and Umrah pilgrimages. In recent years, there has been an increased interest in the advancement of video analytics and visible surveillance to improve the safety and security of pilgrims during their stay in Makkah. It is mainly because Hajj is an entirely special event that involve hundreds of thousands of people being clustered in a small area. This paper proposed a convolutional neural network (CNN) system for performing multitude analysis, in particular for crowd counting. In addition, it also proposes a new algorithm for applications in Hajj and Umrah. We create a new dataset based on the Hajj pilgrimage scenario in order to address this challenge. The proposed algorithm outperforms the state-of-the-art approach with a significant reduction of the mean absolute error (MAE) result: 240.0 (177.5 improvement) and the mean square error (MSE) result: 260.5 (280.1 improvement) when used with the latest dataset (HAJJ-Crowd dataset). We present density map and prediction of traditional approach in our novel HAJJ-crowd dataset for the purpose of evaluation with our proposed method.

Corresponding Author:

Md Roman Bhuiyan
Faculty of Computing and Informatics
Multimedia University
Persiaran Multimedia, 63100 Cyberjaya, Selangor, Malaysia
Email: romanbhuiyanpv@gmail.com, 1181402508@student.mmu.edu.my

1. INTRODUCTION

One of the most popular and highly perceived occasion-timing rituals in religious circles all over the world is Hajj. About three million pilgrims come to Mecca to perform Hajj within 5-6 days every year, requiring them to tour several locations in mecca. The Hajj authorities need to plan the crowd movement by knowing the maximum capacity of every Hajj check points to manage the crowd. In addition, they have to foresee any possibility of dangerous condition which can lead to crowd stamping. Even though plan has been devised, the Hajj authorities still facing difficulties in executing the plan which can be reflected based on their outcomes, i.e. crowd stamping tragedies in September 2015 that results in 2000 pilgrim's death [1]-[3].

To monitor the crowd movement, video surveillance is installed around the kaabah, saei and jamaraat areas. This is due to tawaf and saei areas are the most crowded place during the rituals. However, the surveillance cameras do not have the intelligence to inform the Hajj authorities whether the condition of the crowd at a particular time is dangerous or not. Hence, it is important for a surveillance camera to be able to alert Hajj authorities, should dangerous scenario be about to happen, so that crowd stamping can be prevented.

Vinayakumar *et al.* [4], introduced the DSPNet, which covers multi-dimensional functions for vast numbers of crowd analysis. The current problem of counting numbers to estimate crowd density in extremely congested scenes is particularly to be addressed, since the method are not suitable for congested scenes. The DSPNet model was first used for an interface and meaning. The frontline is a default deep-neural, neural network, while the deep-neural core backend network uses a total knowledge integration ratio at various stages. The SCA cleaner module allows the multiscal functions to be integrated and image representations enhanced.

Analysis of crowd condition in surveillance camera is a difficult task due to excessive occlusions, inconsistencies in scene perception and multiple distributions of crowds. In normal crowded scenes, people detection and monitoring are difficult. This is even more difficult in Hajj scenario where there are too many moving people.

Other approaches to crowd analysis have been resulted in several semi-automated solutions for density estimation and crowd counting [5]-[7]. Effective implementation of the semi-automated solutions, however, is limited by two significant limitations; (1) lack of capacity to accommodate hundreds or thousands of crowds, rather than a few tens of people [5]; and (2) dependence on temporal limitations in crowd videos that do not extend to still images which are more prevalent [8]. In addition to arbitrary camera position and crowd density, there are still problem of erroneous crowd counts from an arbitrary still image [7].

This paper proposes moving-scene crowd analysis model. We aim at analyzing a mapping from motion pictures to crowd analysis and therefore applying the mapping to cross-scene crowd analysis within the target scenes. We propose a model for moving scene crowd analysis. One of the challenges in implementing our method is to utlized the dataset. Among available dataset related to crowd images are the 50 static images obtained from Flickr from different crowd scenes [9]. Other commonly used UCSD dataset were made up of movies collected from one or more scenes [10], [5]. In this work, we obtained our dataset from the YouTube and the detailed dataset is discussed in methodology section.

We propose a fully-fledged system for moving scene crowd video analysis based on the convolution neural network (CNN). We a method that leverage on CNN model particularly suit for Hajj applications. Also, this work introduces a system for counting and then estimating the crowd density. The CNN version is taught through an improved technique for crowd scenes to gain knowledge of targets, crowd density and crowd counts [11], [12]. Our edition of CNN learns crowd-specific capabilities that give better performance than handmade features. -The paper is organized as shown in; section 2 describes the details of the proposed system. Section 3 presents the Hajj-Crowd dataset. Section 4 presents the experimental setup and result analysis. Finally, section 5 concludes the paper and presents the future works.

2. RESEARCH METHOD

The proposed model is built based on a robust Hajj crowd counting detection design. The proposed model predicts precisely localized boxes in Hajj crowd images on people's heads. While it seems like a multi-stage process to find the head size for each person, we develop it as a one-stage end-to-end system. Figure 1 shows the proposed architecture of CNN based on 3 technical components; firstly, we have extraction of the frame.
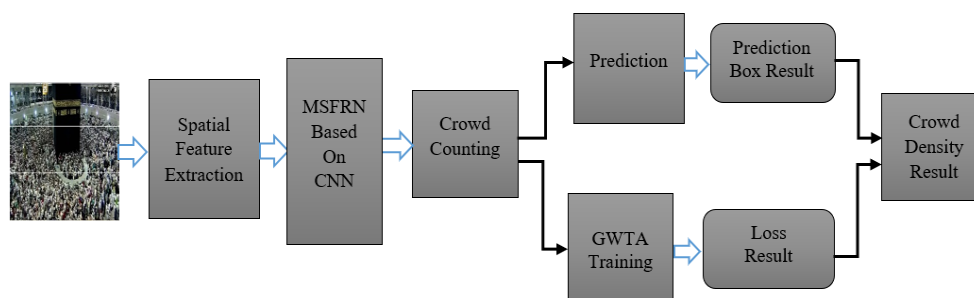

Figure 1. Proposed CNN architecture

For frame extraction we have collected few hajj crowds' videos and after that we did 30 frame extraction from those videos. Second is spatial feature extraction by the feature extractor at multiple resolutions. We have used the CNN based prediction map. This function maps are forwarded to a set of multi-scale feedback reasoning network (MSFRN) based on CNN where information is fused across the scales and predictions are made via boxes. The final output of crowd density is produced by using the non-

maximum suppression (NMS) which combines the correct detection from numerous resolutions. For model preparation, the final stage is replaced by the grid-winner-take-all (GWTA) section and for loss calculation we have used the back-propagation algorithm. It is considered in the next section as details of all the CNN functional components which include the algorithm for training.

2.1. CNN layer architecture

All current detectors of CNN objects run on top of deep backbone feature extractor network. Moreover, the functionality consistency may directly have an impact on detection accuracy. The CNN enabled networks are usually employed in crowd counting in various ways and offer near-stage performance [13]. In accordance with the pattern, multiple layers are also employed in CNN convolution to ensure improved crowd feature extraction. The backbone network is generated by the first five CNN convolution blocks initialized with ImageNet trained [14]. The network takes in a fixed size RGB crowd image (224 x 224) as input with down-sampling of the data for each block because of max-pooling. At each block the network branches, except for the final ones, which are duplicated using the subsequent block. These copied blocks are used for feature map build with a resolution of 0.5, 0.25, 0.125 and 0.166. It contrasts favorably with traditional features of the hyper board and helps to differentiate each branch of the scale by exchanging low-level features in a conflict free manner. Half the spatial scale's low-level features might theoretically catch and handle very large crowds from the data [15]. On the other hand, the minor resolution value divisions have a gradually higher-level receptive area which are ideal with regard to those data that have comparatively limited package. The dimension of diversity is taken care of by providing columns of various sizes for everyone to specialize in a different crowd type.

2.2. Box classification

We select a per pixel classification paradigm for the sizing. Essentially, a series of bounding boxes given prespecified sizes, in which case the model basically does classify each head into or as context to one of the boxes. It compares with the anchor box model commonly used in detector where the box parameters are regressed [16]. Model scale branches generate map set $\{D_n^s\}_h^{nB} = 0$, showing the confidence level for each pixel for classes of the box. Next, ground truth sizes for the heads are required for model training which is unavailable and hard for conventional large sized crowd databases to annotate. In this work, we are developing a system for approximating head sizes. We depend on the point annotations available with crowd data-sets for generating ground truth. Such annotations point defines the positions of people's heads. The position is in the middle of the head roughly however may vary greatly in the case of sparse crowds. Besides, recognizing each individual within the crowd, some information about scale can also be provided by the annotation points on scale. The gap between two adjacent individuals, based on the assumption of the uniform density of the crowd, may revealing the bordering size of the box as per the heads. Remember that only quadratic boxes are considered. Precisely, a given head size may be considered in simple terms as the length from the closest neighbor. Much as this is reasonable for the case of medium to large sized crowds, for the individual's sparsely populated crowds, with a distant nearest neighbor, this may result in incorrect box sizes. But this is considered as empirically producing fairly good head sizing across a large range of densities.

Illustrate here is a pseudo mathematical development of the ground truth. Consider P as set of all the individual's annotated (x, y) positions as per the provided image patch. Thus, the size of the box is specified as for every point (x, y) within P,

$$\beta[x,y] = \frac{min}{(x`,y`)\in P, (x`,y`)\neq(x,y)} \sqrt{(x-x`)^2 + (y-y`)^2} \tag{1}$$

The space right from the closest neighbor. Assuming there exist just one individual in the patch of the image, the box size will be viewed as an option ∞. Next, the values of β [x, y] discretion to the already defined bind that determine the box sizes. Supposing $\beta\frac{s}{1}, \beta\frac{s}{2}, \ldots\ldots, \beta\frac{s}{n_B}$ box sizes are defined for scale, the position classification box labels (x, y) becomes,

$$\beta\frac{-s}{b}[x,y] = \begin{cases} 1 \ if \ (x,y) \in P \ and \ b = 0 \\ 0 \ 1 \ if \ (x,y) \in P \ and \ b > 0 \\ 1 \ if \ (\beta\frac{s}{b} \leq [x,y] < \beta\frac{s}{b+1} \ and \ 0 < b < nB \\ 1 \ if \ (\beta\frac{s}{b} \leq [x,y] < \beta\frac{s}{b+1} \ and \ b = nB \\ 1 \ if \ (\beta\frac{s}{b} \leq [x,y] < \beta\frac{s}{b+1} \ and \ b = 1 \\ 0 \ Otherwise \end{cases} \tag{2}$$

A popular principle is pursued in selecting the box size $\beta \frac{-s}{b} S$ for each scale. The first box size (b=1) at the peak resolution scale (s=ns−1) is usually fixed to one, enhancing the capacity to address highly crowded density. For instance, the boxes that remain we choose larger sizes as per this scale with a continuous increase. The change is fine-grained in higher-resolution partitions however regarding resolutions that are low by definition there will be a steady increase in the coarseness. Precisely, assuming y^s indicates the increment in size for s, we set the box sizes as given next,

$$\beta \frac{s}{b} = \begin{cases} \beta \frac{s+1}{ns} & if \ s < ns - 1 \\ 1 + (b-1)y^s, & Otherwise \end{cases} \tag{3}$$

The standard size increment values for the various scales have the definition y={4, 2, 1, 1}. Observe that high-level resolution divisions (0.5 and 0.25) do contain boxes of better sizes compared to the ones of low resolution (0.16 and 0.25), in which coarse resolution capacity would be adequate (as depicted in Figure 2) [17].

Figure 2. The pseudo box produces ground truth using formula presented in (3)

2.3. Grid winner-take-all training

Calculation of loss: the CNN is trained using the back-propagation of entropy loss per pixel. Each pixel loss is defined,

$$l\left(\{d_i\}\frac{nb}{i} = 0, \{b_i\} \frac{nb}{i} = 0, \{a_i\} \frac{nb}{i} = 0\right) = - \sum_{i=0}^{nb} a_{i \ b_{i \ logd_i}} \tag{4}$$

in that $\{d_i\}\frac{nb}{i} = 0$ is taken as the set of nb+1 probability as per the already defined box classes and where $\{b_i\} \frac{nb}{i} = 0$ means the label of the appropriate ground truth. Next the loss of the total scale branch predictions

$$l(\{D_b^s\}, \{B_b^{-s}\}, \{a_b^s\}) = \sum_{x,y}^{W^s, H^s} \frac{l(\{D_b^s\}, \{B_b^{-s}\}, \{a_b^s\})}{W^s H^s}$$

in which case the inputs are a set of predictions $\{D_b^s\} \frac{nb}{b} = 0$ and pseudo ground truths (for convenience, defined limits may be dropped). Note that (wˢ, hˢ) are the maps of these for estimation spatial sizes and the lack of cross-entropy is summed up:

$$L_{comb} = \sum_{s=1}^{ns} l(\{D_b^s\}, \{B_b^{-s}\}, \{a_b^s\}) \tag{5}$$

2.4. Count of heads

The predictive fusion process is used instead of GWTA to test the model as illustrated in Figure 1. All branches evaluate the image input and result in predictions of multi-resolution. Based on these prediction charts, the box positions are obtained from and linearly scaled to input resolution. To avoid multi-threshold mixing, the current NMS is then applied. The boxes following the NMS form the model's final prediction,

and are listed to generate crowd count. The NMS threshold is defined to be 0.3 (30 percent area overlap) to encourage intermediate assessments during training. But a threshold search is run to determine the most appropriate model upon training to minimize the counting error over the validation set. Figure 3 shows the image prediction made by CNN. Here in Figure 3 (a) represent the original image, Figure 3 (b) CNN prediction with 3 x 3 boxes and Figure 3 (c) depection of the map prediction.
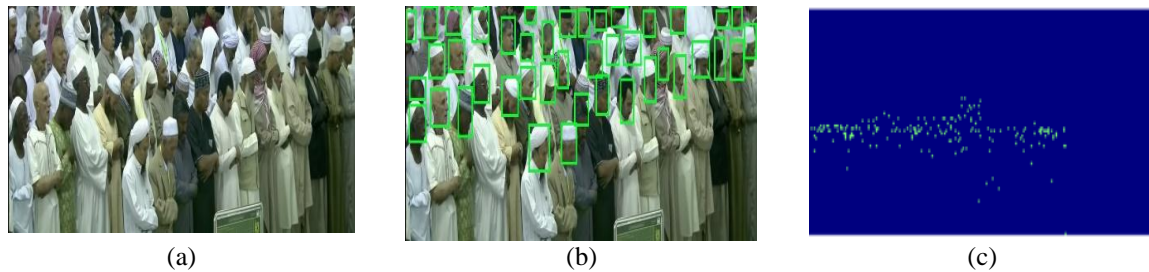

| (a) | (b) | (c) |

Figure 3. This prediction made by CNN on images from Hajj-crowd data-set, (a) instead of image prediction, (b) CNN prediction, (c) map prediction

3. HAJJ-CROWD DATA-SETS

This segment discusses the HAJJ-Crowd proposed from three perspectives: data capture and specification definition, method of annotation and experiment.

3.1. Data capture and specification definition

The collection of HAJJ-crowd data is done from the YOUTUBE'S live telecast in Mecca hajj 2019. Accordingly, in some populated surrounding Kabba (Tawaf area), 1000 images and 10 video sequences are recorded, containing some typical crowd scenes, including touching the black stone in the kabba area, throwing the stone into the mina area. In addision, we have collected 500 samples from Google through the typical crowd-related query keywords. At last, 1500 raw images are obtained by the two methods esribed above. Figure 4 shows the example of the proposed Hajj-crowd dataset.

Figure 4. Hajj-crowd dataset

3.2. Method of annotation (tools)

As an annotation tool, we have utilized on Python and open-cv for easy annotation of head points in the crowd photos. The method supports two label forms, namely point and bounding box. Every image is flexibly zoomed in/out during the annotation process to annotate head with different scales and is divided into a maximum of 3 x 3 small patches, enabling annotators to mark the head within five sizes: 2x (x=0,1,2,3,4) times the initial image dimensions.

4. EXPERIMENTAL SETUP AND RESULT ANALYSIS

4.1. Experimental setup

The CNN model aims to optimize the loss function using the back-propagation algorithm. First, we collected all the images and size of images 1280x720 resolution and the labels are generated under this size. Second, we used deep learning algorithm to enhance CNN and obtained the best results. The training and analysis are implemented on NVIDIA GEFORCE GTX 1660Ti GPU using the deep learning packages PyTorch framework and operating system ubuntu 18.4 LTE. Finally, we used python3 with the deep learning packages such as open cv2, NumPy, SciPy, matplotlib, torch vision, among others.

4.2. Experiment

The collection of HAJJ-crowd data is divided into three parts based on 1,500 images, namely the testing, validation, and training. Based on, two metrics are to measure the counting accuracy that is MAE and MSE. This can be equation as shown in:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |yi - y`i| \tag{6}$$

$$MSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} |y_i - y`_i|^2} \tag{7}$$

In which case N is taken to be the samples in the set of tests, y_i is considered to be the count mark whereas y'_i is the approximated count sample. Additionally, an examination of the model from various viewpoints. The previous has 5 groups by number of people: 0, (0, and 1000), (1000, 2000), (2000, 3000), and over 3000. Each image is assigned an attribute labels according to its annotated counting number and quality of the image. In the experimental set, MAE and MSE are applied in a specific perspective for each class to the corresponding samples. For instance, the luminescence attribute, the computed average figures of MSE and MAE as per the two categories which indicate the sensitivity of the counting models to the luminescence variance.

4.3. Result analysis

Figure 5 (a) and Figure 5 (b) clearly shows that, from 0 to 10 epochs there is no considerable change in pixel loss, whereas from 10 to 20 epochs we have 10-pixel loss. However, for 20 to 30 epochs to 40-52 epochs it continues to increase the pixel loss. Finally, pixel loss at 52 epochs becomes 16.0 the other hand. We can get Test valid loss while, minus Test loss from Training loss. More importantly, in Test valid loss at 40 epochs the pixel becomes 17 and at 52 epochs the pixel loss becomes 14 in number. At the same epochs, we calculated the test MAE based on above equation. After testing we have calculated the test valid loss and test valid MAE that you have seen the above Figure 5 (c) and Figure 5 (d) at the test MAE, we have seen that, when the epoch is 0 then error is above 600. After that when the epochs are increasing then error is going down. After completing 52 epochs, we have seen the error is coming down 240.0. In the test valid MAE, we have seen that, when the epoch is 0 then error is above 425. After that, when the epochs are increasing then error is going down. After completing 52 epochs, we have seen the error is coming down 255.0. Figure 5 shows the graphical representations of the result analysis.

4.4. Compare the proposed model with other state-of-the-art method

Hajj-crowd dataset is a large-scale crowd counting and density dataset. It includes 200 training images and 165 test pictures with the same 1280 x 720 resolution. Some results from the mainstream technique (Idrees et al. [18] Yu et al. [19], [20], FCN [21], Cascaded MTL [22], MCNN [23] and so on) on UCF_CC_50 datasets, are compared against those of non-pre-defined techniques, Idrees et al. attains the finest MAE of 419.5, followed by MSE of 541.6. Our method outperforms the state-of-the-art method, in the context of new dataset (named HAJJ-Crowd dataset), which attains a remarkable MAE result of: 240.0 (177.5-point improvement) and MSE of 260.5 (280.1-point improvement). Table 1 Estimation of errors on the UCF CC 50 dataset.

Table 1. Estimation of errors on the UCF CC 50 dataset

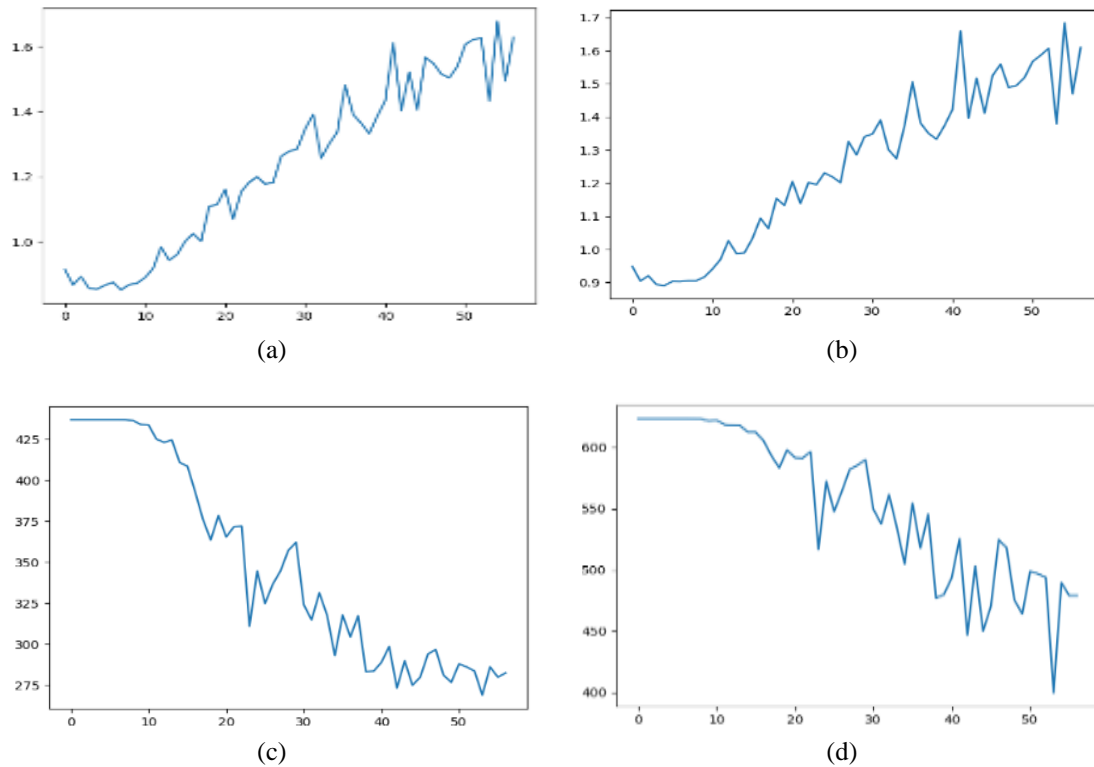| Method	Dataset Name	MAE	MSE
Idrees et al. [18]	UCF_CC_50	419.5	514.6
Yu et al. [19]	UCF_CC_50	467.0	498.5
MCNN [23]	UCF_CC_50	377.6	509.1
FCN [24]	UCF_CC_50	338.6	424.5
Cascaded-MTL [25]	UCF_CC_50	322.8	341.4
Ours	HAJJ-Crowd dataset (Developed by own)	240.0	260.5

(a)

(b)

(c)

(d)

Figure 5. Result analysis, (a) test loss, (b) valid loss, (c) test mae, (d) valid MAE

5. CONCLUSION

This research presents a new crowd density prediction model using convolutional neural network. The current model of convolutional neural network uses a multi-column structure of highest level-down processing of feedback to address the issues in massive crowds. Unlike the abovementioned model, the proposed model can detect moving crowds. The result of this experiment exhibits better performances compare to other methods. In particular, crowd analysis increases the counting efficiency for highly congested crowd scenes considerably. Upcoming research will consider better detection of crowds and would make head sizing more accurate annotation.

REFERENCES

[1] M. Sigala, A. Beer, L. Hodgson, and A. O'Connor, "Big Data for Measuring the Impact of Tourism Economic Development Programmes: A Process and Quality Criteria Framework for Using Big Data," *Big Data and Innovation in Tourism, Travel, and Hospitality*, pp. 57-73, 2019, doi: 10.1007/978-981-13-6339-9_4.

[2] G. Nguyen *et al.*, "Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey," *Artificial Intelligence Review*, vol. 52, no. 1, pp. 77-124, 2019, doi: 10.1007/s10462-018-09679-z.

[3] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *Journal of Big Data*, vol. 6, no. 1, p. 77, 2019, doi: 10.1186/s40537-019-0197-0.

[4] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat and S. Venkatraman, "Deep Learning Approach for Intelligent Intrusion Detection System," in *IEEE Access*, vol. 7, pp. 41525-41550, 2019, doi: 10.1109/ACCESS.2019.2895334.

[5] K. Sivaraman, R. M. V. Krishnan, B. Sundarraj, and S. Sri Gowthem, "Network failure detection and diagnosis by analyzing syslog and SNS data: Applying big data analysis to network operations," *International Journal of Pure and Applied Mathematics*, vol. 8, no. 12, pp. 883-887, 2019, doi: 10.35940/ijitee.I3187.0789S319.

[6] A. D. Dwivedi, G. Srivastava, S. Dhar, and R. Singh, "A decentralized privacy-preserving healthcare blockchain for IoT," *Sensors (Switzerland)*, vol. 19, no. 2, pp. 1-17, 2019, doi: 10.3390/s19020326.

[7] F. Al-Turjman, H. Zahmatkesh and L. Mostarda, "Quantifying Uncertainty in Internet of Medical Things and Big-Data Services Using Intelligence and Deep Learning," in *IEEE Access*, vol. 7, pp. 115749-115759, 2019, doi: 10.1109/ACCESS.2019.2931637.

[8] S. Kumar and M. Singh, "Big data analytics for healthcare industry: impact, applications, and tools," in *Big Data Mining and Analytics*, vol. 2, no. 1, pp. 48-57, March 2019, doi: 10.26599/BDMA.2018.9020031.

[9] L. Ang, K. P. Seng, G. K. Ijemaru and A. M. Zungeru, "Deployment of IoV for Smart Cities: Applications,

Architecture, and Challenges," in *IEEE Access*, vol. 7, pp. 6473-6492, 2019, doi: 10.1109/ACCESS.2018.2887076.

[10] B. P. L. Lau, *et al.*, "A survey of data fusion in smart city applications," *Information Fusion*, vol. 52, pp. 357-374, 2019, doi: 10.1016/j.inffus.2019.05.004.

[11] Y. Wu, *et al.*, "Large scale incremental learning," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, vol. 2019-June, pp. 374-382, 2019, doi: 10.1109/CVPR.2019.00046.

[12] A. Mosavi, S. Shamshirband, E. Salwana, K. wing Chau, and J. H. M. Tah, "Prediction of multi-inputs bubble column reactor using a novel hybrid model of computational fluid dynamics and machine learning," *Engineering Applications of Computational Fluid Mechanics*, vol. 13, no. 1, pp. 482-492, 2019, doi: 10.1080/19942060.2019.1613448.

[13] V. Palanisamy and R. Thirunavukarasu, "Implications of big data analytics in developing healthcare frameworks–A review," *J Journal of King Saud University-Computer and Information Sciences*, vol. 31, no. 4, pp. 415-425, 2019, doi: 10.1016/j.jksuci.2017.12.007.

[14] J. Sadowski, "When data is capital: Datafication, accumulation, and extraction," *Big Data Society*, vol. 6, no. 1, pp. 1-12, 2019, doi: 10.1177/2053951718820549.

[15] J. R. Saura, B. R. Herráez and A. Reyes-Menendez, "Comparing a Traditional Approach for Financial Brand Communication Analysis With a Big Data Analytics Technique," in *IEEE Access*, vol. 7, pp. 37100-37108, 2019, doi: 10.1109/ACCESS.2019.2905301.

[16] D. Nallaperuma, *et al.*, "Online Incremental Machine Learning Platform for Big Data-Driven Smart Traffic Management," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 12, pp. 4679-4690, Dec. 2019, doi: 10.1109/TITS.2019.2924883.

[17] S. Schulz, M. Becker, M. R. Groseclose, S. Schadt, and C. Hopf, "Advanced MALDI mass spectrometry imaging in pharmaceutical research and drug development," *Current Opinion in Biotechnology*, vol. 55, pp. 51-59, 2019, doi: 10.1016/j.copbio.2018.08.003.

[18] C. Shang and F. You, "Data Analytics and Machine Learning for Smart Process Manufacturing: Recent Advances and Perspectives in the Big Data Era," *Engineering*, vol. 5, no. 6, pp. 1010-1016, 2019, doi: 10.1016/j.eng.2019.01.019.

[19] Y. Yu, M. Li, L. Liu, Y. Li and J. Wang, "Clinical big data and deep learning: Applications, challenges, and future outlooks," in *Big Data Mining and Analytics*, vol. 2, no. 4, pp. 288-305, Dec. 2019, doi: 10.26599/BDMA.2019.9020007.

[20] M. Huang, W. Liu, T. Wang, H. Song, X. Li and A. Liu, "A Queuing Delay Utilization Scheme for On-Path Service Aggregation in Services-Oriented Computing Networks," in *IEEE Access*, vol. 7, pp. 23816-23833, 2019, doi: 10.1109/ACCESS.2019.2899402.

[21] G. Xu, Y. Shi, X. Sun, and W. Shen, "Internet of things in marine environment monitoring: A review," *Sensors (Switzerland)*, vol. 19, no. 7, pp. 1-21, 2019, doi: 10.3390/s19071711.

[22] M. Aqib, R. Mehmood, A. Alzahrani, I. Katib, A. Albeshri, and S. M. Altowaijri, "Smarter traffic prediction using big data, in-memory computing, deep learning and gpus," *Sensors,* vol. 19, no. 9, p. 2206, 2019, doi: 10.3390/s19092206.

[23] S. Leonelli and N. Tempini, "Data Journeys in the Sciences," Springer Nature, 2020, p. 412,

[24] N. Stylos and J. Zwiegelaar, "Big Data as a Game Changer: How Does It Shape Business Intelligence Within a Tourism and Hospitality Industry Context?," *Big Data and Innovation in Tourism, Travel, and Hospitality*, pp.163-181, 2019, doi: 10.1007/978-981-13-6339-9_11.

[25] Q. Song, H. Ge, J. Caverlee, and X. Hu, "Tensor completion algorithms in big data analytics," *ACM Transactions on Knowledge Discovery from Data*, vol. 13, no. 1, pp. 1-48,2017, doi: 10.1145/3278607.

BIOGRAPHIES OF AUTHORS

Md Roman Bhuiyan is currently pursuing his Doctor of Philosophy (Ph.D.) by research in Information Technology from Multimedia University, Cyberjaya, Malaysia. He received the B.Sc. (Eng.) in computer science and engineering from Uttara University, Bangladesh, in 2015, and the Masters of Software Engineering from the faculty of computer science and engineering, FTMS College Malaysia collaboration with Leeds Beckett University, UK, in 2017.

Junaidi Abdullah is a Senior Lecturer in the Faculty of Computing and Informatics, Multimedia University (MMU), Malaysia. He is currently the Chairperson of Assistive Technology Special Interest Group. He was the Deputy Dean (Academic) of FCI, MMU (2008 – 2011). He obtained a Bachelor degree in Engineering (BEng., First Class Honours), in University of Bristol, UK, and Ph.D. in Computer Science in the area of Computer Vision and Augmented Reality from the University of Southampton, UK in 2005. His research includes augmented reality, image and video processing, computer vision. He has published more than 50 internationally multi-disciplinary refereed conference papers, journal articles and books.

Noramiza Hashim graduated with a Diplôme d'Ingénieur (Master of Science in Engineering) from the Higher Institute for Advanced Technologies of Saint-Etienne (ISTASE) Université Jean Monnet, France in 2002. She also acquired a D.E.A (Diplôme d'Etude Approfondie / Master Degree in Research) from the same university. In 2008, she obtained her Ph.D. in Information Technology under a joint Ph.D. program between Multimedia University, Malaysia and Université de La Rochelle, France. She is currently lecturing at the Faculty of Computing and Informatics, Multimedia University. Her research interests are mainly on digital image processing and object recognition.

Fahmid Al Farid is currently pursuing his Doctor of Philosophy (Ph.D.) by research in Information Technology from Multimedia University, Cyberjaya, Malaysia. He received the B.S. degree in computer science and engineering from University of Chittagong, Bangladesh, in 2010, and the M.S. degree from the faculty of computer science and electrical engineering, University of Ulsan (UOU), South Korea, in 2015. From 2013 to 2014, he was a Research Assistant in Embedded System laboratory, University for Ulsan, South Korea. In 2015, he was a Research Assistant in Ubiquitous Computing Technology Research Institute (UTRI), Sungkyunkwan University, South Korea. He received Korean BK21 PLUS Scholarship, Supported by Korean Government in MS (2012-2014).

Dr Mohd Ali Samsudin is an Associate Professor in the School of Educational Studies, University Sains Malaysia. His expertise and research interests are Science and Physics Education, Psychometrics and Analysing Big Data using Social Network Analysis and R Statistics. He teaches several statistics courses for postgraduate students and conducts sworkshops with Rasch Model analysis, Structural Equation Modelling (SEM), Social Network Analysis (SNA) and R Language Programming Statistics. The Malaysian Ministry of Education has appointed him to lead a national project to develop and diagnose Science, Technology, Engineering and Math (STEM) competency among Science and Math teachers in Malaysia. Currently, he is a Master Trainer for Microsoft Innovative Educator programmed and has been certified as Microsoft Innovative Educator Expert. He also has been appointed by the Academy Science of Malaysia (ASM) as one of the national working committees (STI enculturation) for Science Outlook 2017-2018: Converging towards Progressive Malaysia 2050. He is also leading a national project to measure teachers' digital competency across 14 states in Malaysia. He is also has been appointed as one of the national experts for the Malaysia Smart School Quality Standard.

Norra Abdullah has graduated from the Northern University of Malaysia and obtained a degree in B.A (Hons) Accountancy. She is well exposed in the accountancy, human resource & administration for about 20 years. She is also exposed to Marketing & PR function during my working experience. The years of working had taken me through the various level of business management and work coordination. She started to earn working experience in accounting jobs in Ricomal Industries Sdn Bhd. She has been handling full set of accounts. Meanwhile, MayFirst Gold Sdn Bhd had exposed me into the world of marketing, customer service area and sales area. She was exposed into marketing aspects where she was trained to be a marketing person which involved advertising & promotions, brand management, products costing, PR and events besides dealing with multi categories of products and brands sourcing. She was also involved in doing marketing planning, long term strategic planning of the company and execute marketing strategies. She lead 20 associates under marketing department. She was then being promoted to be the Assistant general Manager cum finance manager in MayFirst Gold Sdn Bhd where my main task is to manage the finance and operation of the company. Currently She is the Senior Manager in Finance/Admin/Hr Department in WSA Venture Australia (M) Sdn Bhd where my main task is to handle and manage 3 departmental functions.

Jia Uddin received a BSc degree in Computer and Communication Engineering from International Islamic University Chittagong, Bangladesh in 2005, and a MSc. degree in Telecommunications from the Blekinge Institute of Technology, Sweden, in 2010. He did Ph.D. in Computer Engineering from the University of Ulsan, Korea, in January 2015. He is an Assistant Professor in Department of Technology Studies, Endicott College, Woosong University, South Korea and an Associate Professor (On Leave), Computer Science and Engineering Department at Brac University, Bangladesh. His research interests include fault diagnosis, computer vision, and multimedia signal processing.