



User Targeted Offline Advertising using Recognition Based Demographics and Queue Scheduling

Ruchika Malhotra, Samarth Gupta, Sarthak Katyal, Ronak Sakhuja

Abstract: Offline advertisements are static in nature. Advertising companies use billboards for advertising. These billboards display advertisements in a random fashion depending on the investment made by the advertiser. Advertisers pay a fixed amount of money for displaying their advertisements and not on the basis of relevant viewership. The technology proposed in the paper ensures that this disparity is handled wherein offline advertisements are targeted to the relevant audience. The technology has been named TARP which is an abbreviation for Target. Advertise. Revolutionise. Promote. TARP uses built in cameras on offline advertising platforms such as billboards & TV Screens in malls, restaurants, metro & airports to target advertisements based on gender, age and other relevant demographics. The technology is a boon for the advertising industry and benefits both advertisers and viewers. It displays what viewers want to see and who the advertisers want to reach out to. Convolutional neural networks are used to generate demographics of viewing population. Centroids of the viewing population are maintained for each billboard. Advertisements search for the most relevant billboard for display. Display of advertisements is monitored by a queue scheduling algorithm. The research paper proposes an algorithm to generate demographics, search most relevant billboard for each advertisement as well as generate priority queues.

Keywords: Offline, Advertisements, Targeting, Queue, Age, Gender.

I. INTRODUCTION

Advertising plays an important role in the times when the markets are expanding and the competition is increasing.

Businesses invest huge amount of money in advertising with the motive of reaching out to their target audience. The advertising industry is evolving day by day; from print media to digital advertising, and digital billboards. Advertisements running on the digital billboards are usually not user targeted, as desired by the businesses. Currently, billboards display advertisements in a queue-like fashion for a time quantum without taking into account the demographics of the viewing population. Our technology makes sure that the advertisements playing on the digital billboard are relevant to the audience, by collecting and analysing the demographics of the viewing population. This is accomplished by using cameras placed at various sites near the billboards. This ensures that the advertisements are displayed on the most appropriate billboard.

TARP stands for Target. Advertise. Revolutionise. Promote. It complies with data protection laws because no individual identifying information is being collected or stored. In the first step, cameras on offline advertising spots generate the demographics of the viewing audience using neural networks. In the second step, JSON data obtained is processed and centroid of viewing population for each billboard is estimated. These centroids are maintained in real time database. Advertisements and billboard centroids are plotted on 4-dimensional graph, each advertisement searches for the nearest billboard and enters its queue. Advertisements in queue of each billboard are then displayed depending on highest priority. The advertiser's portal is a one step solution for all the needs. Advertisers have the option of uploading new advertisements and to check statistics of existing advertisement. They can even top-up or make payments and check balance. New advertisement requires information like total spending, target audience, location preferences and zone selection. Based on the location preferences, advertiser can choose the zone on a map that shows the statistics from each zone like number of billboards, average males/females per billboard etc. Financial status of viewers in each zone is also showcased. Finally, advertisers check the base charge for the zone and upload the advertisement requesting further review. Review advertisement feature has the features like viewership so far, number of times advertisements are displayed, amount spent and relevant viewership wherein a targeted person views advertisement.

Revised Manuscript Received on February 15, 2020.

* Correspondence Author

Dr. Ruchika Malhotra, Department of Computer Science Engineering, Delhi Technological University, New Delhi, India. Email: ruchikamalhotra@dtu.ac.in

Samarth Gupta, Department of Computer Science Engineering, Department, Delhi Technological University, New Delhi, India. Email: samarthgupta1011@gmail.com

Sarthak Katyal, Department of Computer Science Engineering, Department, Delhi Technological University, New Delhi, India. Email: srthkkatyal@gmail.com

Ronak Sakhuja, Department of Computer Science Engineering, Department, Delhi Technological University, New Delhi, India. Email: ronaksakhuja@gmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

II. RELATED WORK

A. Behavioral Targeting

Behavioural targeting is used to target users for online advertising. It is used to collect data via users' website browsing habits depending on what the user searches for. Advertisements are more effective when they are targeted to a small user base. V. Bhatia and V. Hasija [1] used social data mining and behavioural data for targeted advertising. Many factors have been considered for targeting advertisements such as location, gender, search history, browsing history and so on.

B. Convolutional Neural Network

Artificial Neural networks have a problem in dealing with images as the number of arguments become very high. S. Albawi, T. A. Mohammed and S. Al-Zawi [2] suggest that Convolutional Neural Network can be a better alternative under such scenarios. The most important assumption about the issues that are solved by convolutional neural network is that it should not have any features that are spatially dependent. For example, in a face detection algorithm, we are only concerned if there are faces in an image not the location of the faces.

C. Age Detection

Haibin Liao, Yuchen Yan, Wenhua Dai et al. [3] explain the various advantages of convolution neural networks when working on images. The method involves extracting face features by using a deep learning model, meanwhile the robust features are extracted by using a statistical method. It finds the correlation between the observed features and produces a lower number of unobserved features. The deep learning model is then trained using rank based age estimation in addition to the divide and rule strategy. Our method uses a deep convolutional neural network to estimate the age class from the images, and the model is trained on the IMDB-WIKI face dataset, as used in [3].

D. Gender Detection

Humans can determine an individual's gender by recognising the facial features in most cases. Shan Sung Liew, Mohamed Khalil-Hani, Feeza Radzi et al. [4] propose a method for real time gender classification on face images. It uses a convolutional neural network with reduced complexity to reduce the computational load, meanwhile providing accurate results. The method fuses the convolutional and subsampling layers such that the processing layers in the CNN is reduced to only four. In addition to this, the convolution operation is replaced with cross-correlation, and the training is done using second order back propagation algorithm.

E. K Nearest Neighbours

P. Cunningham and S. J. Delany [5] explain nearest neighbour algorithm to classify an example based on the majority class of the K-nearest neighbours. The research provides an overview of various similarity measures as well as the computational issues involved. D. Adeniyi, Z. Wei and Y. Yongquan [6] use K-nearest neighbour algorithm to provide a user with the relevant products or information on a website based on the click stream. This real-time on-line

recommendation system inspires the model of our demographic based offline advertising model.

F. Transfer Learning

Machine learning and deep learning algorithms have achieved record breaking results in the fields like regression, classification and clustering. These models work because the training data set and test data set have been drawn from the same feature space and distribution. In cases where we need to rebuild similar model, it is very difficult to recollect the data and do the training process. In such cases, transfer learning would be necessary and ideal to use. Some common deep learning models have been made available by S. J. Pan and used by Q. Yangfor [7] to save time and computation power to attain new record breaking results.

G. Queue Based Scheduling

In priority-based scheduling, tasks with the same priority levels receive equal amounts of CPU time. Similarly, as per our use case queue for each billboard is created. The advertisements enter the queue depending upon the distance between advertisement coordinates and the billboard coordinates. The queue with the highest priority i.e. at the top is displayed. If there is a delay in new centroid calculation, or creation of new queues and the delay time exceeds the display time of advertisement with highest priority, then advertisement with subsequent priorities are displayed.

III. PROPOSED WORK

A. Overview

Figure 1 showcases the architecture of the technology used in TARP. As shown, a camera on the billboard sends the image frame to the first stage of processing. Image frame undergoes segmentation, rescaling and cropping. It is then passed to the network wherein class outputs are received (Refer Section 3B). There are six output classes: three for male class and corresponding age groups; three for female class and corresponding age groups. The JSON Data obtained as shown in Figure (1) is then processed to find the centroid of the viewing population (Refer Section 3C). The centroids of the viewing population are maintained in a real time database. The advertisements which are stationary on the coordinate plane search for the nearest billboard and join its queue (Refer Section 3D). Queue based scheduling algorithm controls the queue of each billboard and ensures that the highest priority advertisement is displayed. (Refer Section 3E)

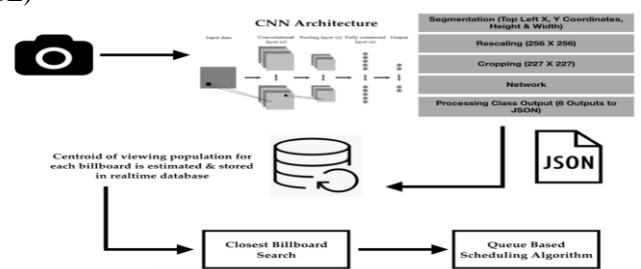


Fig. 1. Overall Architecture of TARP.

B. Generation of Demographics

It is easy to classify a person as male or female or the age group in online advertising because of the tracking nature of websites such as Google and Facebook which track your movement throughout the internet by injecting scripts into other websites using their advertisement portal. We want to take this approach of targeted advertising to offline channels, but since there is no way to track where a person goes and what a person does in his/her life, we focus on the two main attributes: Gender and Age [11].

The first step is Haar Cascade Face Segmentation. P. Viola and M. Jones [8] proposed the algorithm of Haar Cascade which is a machine learning object detection algorithm. In this algorithm, a cascade function is used to form a lot of positive and negative images where positive represents that the object to be found is present and negative where in it is not present. It is then used to detect objects in images other than the ones used to train it. We use pre-trained models for face detection in our model [13]. Pre-processing is then done. All three colour channels are processed directly by the network. Images are first rescaled to 256×256 and a crop of 227×227 is fed to the network. Network consists of three convolution layers and three fully connected layers as shown in Figure 3.

96 CNN filters of size $3 \times 7 \times 7$ pixels are used in the input layer in the first convolutional layer, after which we have a ReLU layer (rectified linear operator), and the last is a max pooling layer which takes the maximum value of 3×3 regions with a stride of two pixels and also a local response normalisation layer. The output of the size $96 \times 28 \times 28$ of the previous layers is later on processed by the next CNN layer, which has 256 filters of the size $96 \times 5 \times 5$ pixels.

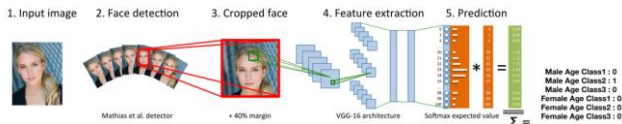


Fig. 2. Architecture for Generation of Demographics

```
conv1 = convolution2d(images, 96, [7,7], [4, 4], padding='VALID', biases_initializer=tf.constant_initializer(0.), scope='conv1')
pool1 = max_pool2d(conv1, 3, 2, padding='VALID', scope='pool1')
conv2 = convolution2d(pool1, 256, [5, 5], [1, 1], padding='SAME', scope='conv2')
pool2 = max_pool2d(conv2, 3, 2, padding='VALID', scope='pool2')
conv3 = convolution2d(pool2, 384, [3, 3], [1, 1], padding='SAME', biases_initializer=tf.constant_initializer(0.), scope='conv3')
pool3 = max_pool2d(conv3, 3, 2, padding='VALID', scope='pool3')
# can use tf.contrib.layers.flatten
flat = tf.reshape(pool3, [-1, 384*6*6], name='reshape')
full1 = fully_connected(flat, 512, scope='full1')
drop1 = tf.nn.dropout(full1, pkeep, name='drop1')
full2 = fully_connected(drop1, 512, scope='full2')
drop2 = tf.nn.dropout(full2, pkeep, name='drop2')
with tf.variable_scope('output') as scope:
    weights = tf.Variable(tf.random_normal([512, nlabels], mean=0.0, stddev=0.01), name='weights')
    biases = tf.Variable(tf.constant(0.0, shape=[nlabels], dtype=tf.float32), name='biases')
    output = tf.add(tf.matmul(drop2, weights), biases, name=scope.name)
```

Fig. 3. Layers in CNN

C. Calculation of centroids and storage in database

The JSON output yielded for an image frame is depicted in Figure 4. It can be used to mathematically characterise the viewing population of the billboard. To get an estimate of the population, centroid is calculated based on all the individuals. Our convolution neural network yields the age class as well as the gender of each of the individual. Therefore, for each individual a set of attributes is yielded (C1, C2, C3, G) where C1 represents Age Class 1, C2 represents Age Class 2, C3 represents Age Class 3 and G represents if the individual is a

Finally, this is followed by Rectified Linear Unit, a max pooling layer and a local response normalisation layer with the same hyper parameters as before. Now at last, the third CNN layer operates on the image size of $256 \times 14 \times 14$ by applying a total of 384 filters of size $256 \times 3 \times 3$ pixels, which is then followed by Rectifier Linear Unit and a max pooling layer.

First FCL (Fully Connected Layer) gets the output of the last convolutional layer and it contains a total 512 neurons, followed by a Rectified Linear unit and a dropout layer. The next fully connected layer which is the 2nd FCL gets the 512-dimensional output of the previous fully connected layer and this also contains 512 neurons, followed by a ReLU and a dropout layer. The last fully connected layer maps to the final classes for age or gender. Finally, the output of the last fully connected layer is fed to a soft-max layer that assigns a probability for each class. The prediction itself is made by taking the class with the maximal probability for the given test image.

We used the IMDB-WIKI 500K+ face images with age and gender labels of which we used 80-20 train test split to test and train our model. All three colour channels are processed directly by the network. Images are first rescaled to 256×256 and a crop of 227×227 is fed to the network. The segmented facial images in an image frame fed to the CNN network yields an output for each of the person in the population. The output for each of the face can be compiled into the following JSON format. 'Age Class 1' represents individuals below 18 years of age, 'Age Class 2' represents individuals below 60 years of age, 'Age Class 3' represents individuals of the age of 60 and above and 'isMale' represents if the individual is a male or not. Figure 4 shows the class outputs.

male or not. For each of the individual, C1, C2, C3 and G can take a value of either 0 or 1.

The calculated centroid for Billboard 'i' therefore is $C_i(\overline{C1}, \overline{C2}, \overline{C3}, \overline{G})$. In a population of N individuals, centroid calculation is done as per "(1)" to "(4)"

F. Advertiser's Portal

The advertiser's portal has been built on PHP and MySQL on Linux machine running Apache server hosted on Microsoft Azure. It has been built to provide an easier interface to the users who wish to upload their advertisement on the platform. It solves the problem of dealing with money by preloading credit into the account. The loaded money is then used for making expenditures on advertisement display. The advertiser can set their budget threshold for each advertisement so as to prevent overspending. Advertising is done by setting which locations to target depending on their needs. Advertiser is required to input the relevant viewership of the advertisement which includes gender and age group targeting. Once the advertisement is posted, the advertisement will be displayed on billboards where the relevant viewership is detected. Another feature of the portal is the dashboard, which visually displays the expenses and

advertisement statistics in graphical form for easier view-ability and understanding. Advertiser can track expenses in graphical view and also export data about their advertisements.

IV. RESULTS AND DISCUSSIONS

The centroid of the viewing population in front of a billboard have been plotted for successive frames of a video feed in figure 7. It can be seen that as the demographics of the population change, the centroid coordinates of the billboard change. Target demographics for an advertisement are predefined and hence the plot of advertisement coordinates is a stationary point. The four dimensions have been showcased using 3D Plot and using colour as the fourth dimension for easy visual representation. The four dimensions of a billboard correspond to the three age groups and the gender of the centroid of the viewing population.

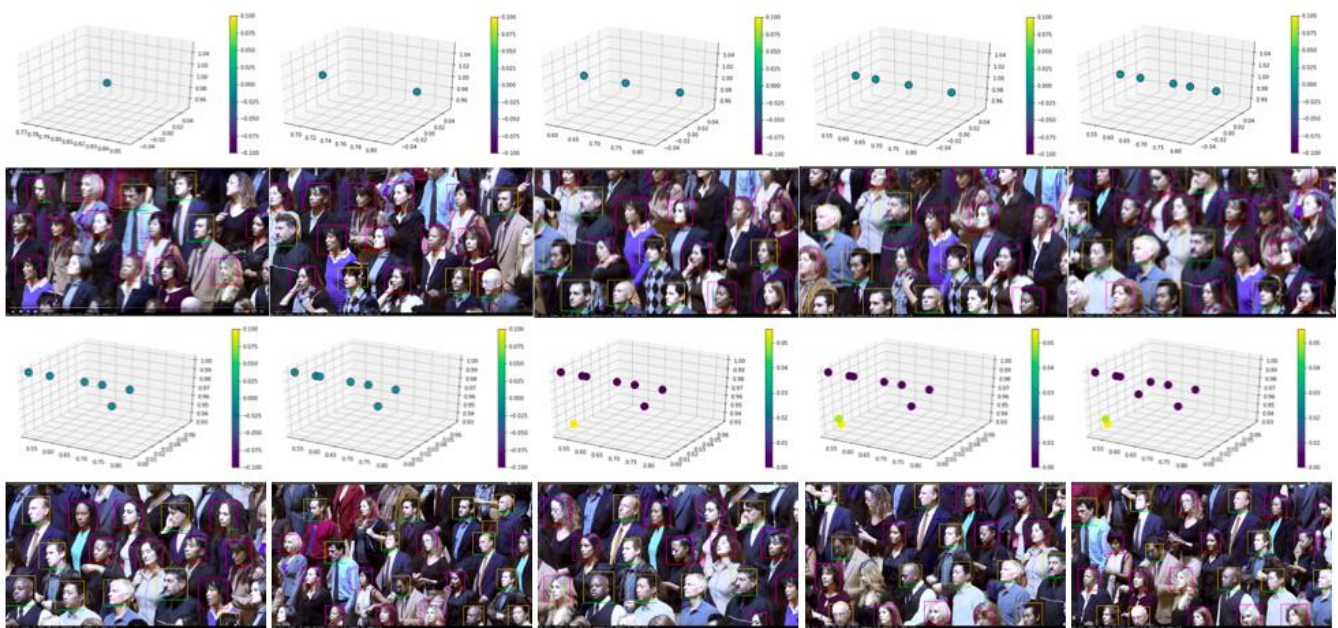


Fig. 7. 4D Plot showcasing changing billboard coordinates

Figure 8 shows how the queues are generated for each billboard. The advertisements with the highest priority in the priority queue are then displayed on the billboard. Table 1 shows the distance of each advertisement in queue from the billboard at an instance. The queue with smallest distance has the highest priority.

billboard1 has ['advertiser1', 'advertiser3', 'advertiser4', 'advertiser5'] in queue billboard2 has ['advertiser2'] in queue
billboard1 has ['advertiser3', 'advertiser4', 'advertiser5'] in queue billboard2 has ['advertiser1', 'advertiser2'] in queue

Fig. 8. Queue Generation of Billboards

Table- I: Distance between advertisements in queue and billboard

Billboard 1	Advertisement 3 (Distance : 0.432)	Advertisement 4 (Distance : 0.766)	Advertisement 5 (Distance : 0.952)
Billboard 2	Advertisement 1 (Distance : 0.372)	Advertisement 2 (Distance : 0.632)	

V. PERFORMANCE EVALUATION

A. Improvement from Random Display

Testing over 20 sample advertisements, mean of individual relevant percentages of all advertisements is calculated. Table 2 shows the results for each advertisement. Mean of relevant percentage of random display of advertisements is calculated to be 30.67 % and mean of relevant percentage for display of advertisements using TARP is 60.64%. Therefore, improvement percentage is 29.97%.

B. Delay Per Video Frame

The algorithm was tested on different images with different number of individuals and the response time was studied. The various sources because of which delay is incurred in response time are: delay involved in segmentation using Haar Cascades Face Detection, delay incurred in API call response from backend, time taken by neural network to process the image and create the JSON response,

delay incurred in receiving the response, delay in processing the response, finding out the centroid of the population, and finding the closest billboard for each advertisement.

Table 3 summarises the response time obtained as the number of people in an image frame increases. The mean response time is 4.34 seconds. This response time is well adjusted to the duration of the advertisement at the top of queue. In case the duration of advertisement is lower than the response time, the advertisement with next priority is displayed on the billboard.

Table- II: Improvement of relevant percentage

Advertisement No.	Random Display of Advertisement (Relevant Percentage)	Display using TARP
1	37.8%	62.7%
2	44.6%	74.8%
3	43.6%	68.7%
4	21%	64.8%
5	27.6%	56.8%
6	33%	54.8%
7	29.5%	78.8%
8	27.4%	65.7%
9	34.5%	91.8%
10	28.6%	47.8%
11	44.2%	44.8%
12	22.7%	71.25
13	18.6%	58.7%
14	69.4%	63.2%
15	35.3%	49.8%
16	26.8%	88.5%
17	12.5%	54.7%
18	13.9%	44.4%
19	19.5%	61.8%
20	24.7%	73.7%

Table- III: Response Time

Number of People in the Image	Response Time (in seconds)
18	3.17
27	5.34
11	2.62
28	5.52
18	3.81
16	3.47
24	4.83
28	5.51
26	5.17
18	3.18
Mean Number of People = 19.6	Mean Response Time = 4.34s

VI. CONCLUSION

Using demographic recognition techniques, searching the most relevant advertisement that should be displayed on a billboard ensures that the viewers view the advertisements of their interest. Advertisers can reach out to their target audience which in turn helps them sell their products efficiently. Advertisers have access to the portal to analyse the benefits of targeted advertising. They can pick across various locations based on the base charge and their budget. Advertiser's portal makes uploading a new advertisement easier and the simplifies the advertising process. Performance has enhanced as compared to random advertisement display. An overall improvement of 29.97% is observed. Advertisers can reach an exponentially wider target audience with a marginal increase in the cost. This ensures benefits for the advertisers as well as the billboard owners. Variable delay times can be handled using queuing model. The queue with the highest priority i.e. at the top is displayed. If there is a delay and the delay time exceeds the display time of advertisement with highest priority, then advertisement with subsequent priorities are displayed.

REFERENCES

1. V. Bhatia and V. Hasija, "Targeted advertising using behavioural data and social data mining," 2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN), Vienna, 2016, pp. 937-942.
2. S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," 2017 International Conference on Engineering and Technology (ICET), Antalya, 2017, pp. 1-6.
3. Haibin Liao, Yuchen Yan, Wenhua Dai and Ping Fan, "Age Estimation of Face Images Based on CNN and Divide-and-Rule Strategy," Mathematical Problems in Engineering, Hindawi, Volume 2018, Article ID 1712686.
4. Shan Sung Liew, Mohamed Khalil-Hani, Feeza Radzi and Rabia Bakhteri, "Gender Classification: A Convolutional Neural Network Approach", Turkish Journal of Electrical Engineering and Computer Sciences, Volume 24, pp. 1248-1264, 2016.
5. P. Cunningham and S. J. Delany, "k-nearest neighbour classifiers," Multiple Classifier Systems, vol. 34, no. 8, pp. 1-17, 2007.
6. D. Adeniyi, Z. Wei and Y. Yongquan, "Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method", Applied Computing and Informatics, Volume 12, pp. 90-108, 2016.

7. S. J. Pan and Q. Yang, "A Survey on Transfer Learning," in IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 10, pp. 1345-1359, Oct. 2010.
8. P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, Kauai, HI, USA, 2001, pp. 1-1.
9. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, pp. 886-893 vol. 1.
10. Asma Gul, Aris Perperoglou, Zardad Khan, Osama Mahmoud, Miftahuddin Miftahuddin, Werner Adler and Berthold Lausen, "Ensemble of a subset of kNN classifiers", Advances in Data Analysis and Classification, Volume 12, pp. 827-840, 2018.
11. X. Wang, A. Mohd Ali and P. Angelov, "Gender and Age Classification of Human Faces for Automatic Detection of Anomalous Human Behaviour," 2017 3rd IEEE International Conference on Cybernetics (CYBCONF), Exeter, 2017, pp. 1-6.
12. S. Kumar, S. Singh and J. Kumar, "A study on face recognition techniques with age and gender classification," 2017 International Conference on Computing, Communication and Automation (ICCCA), Greater Noida, 2017, pp. 1001-1006.
13. X. Liu, J. Li, C. Hu and J. Pan, "Deep convolutional neural networks-based age and gender classification with facial images," 2017 First International Conference on Electronics Instrumentation & Information Systems (EIIS), Harbin, 2017, pp. 1-4.
14. D. Tyas Purwa Hapsari, C. Gusti Berliana, P. Winda and M. Arief Soeleman, "Face Detection Using Haar Cascade in Difference Illumination," 2018 International Seminar on Application for Technology of Information and Communication, Semarang, 2018, pp. 555-559.
15. Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke and Andrew Rabinovich, "Going Deeper with Convolutions", arXiv, 1409.4842, 2014.



Ronak Sakhuja is pursuing B. Tech. in Computer Engineering from Delhi Technological University, Delhi, India. He has published Research Paper titled 'Object Detection in Foggy Conditions by Fusion of Saliency Map and YOLO' at International Conference of Sensing Technology held at University of Limerick, Ireland from 3-6 December 2018. He is currently the president of DTU Chapter of International Organization of Software Developers. He is also an active Microsoft Student Partner and incoming Software Engineer at Microsoft India Development Centre.

AUTHORS PROFILE



Dr. Ruchika Malhotra is Associate Head and Associate Professor at the Department of Software Engineering, Delhi Technological University (formerly Delhi College of Engineering), Delhi, India. She was awarded with prestigious Raman Fellowship for pursuing Post doctoral research in Indiana University Purdue University Indianapolis USA. She was an assistant professor at the University School of

Information Technology, Guru Gobind Singh Indraprastha University, Delhi, India. She received her master's and doctorate degree in software engineering from the University School of Information Technology, Guru Gobind Singh Indraprastha University, Delhi, India. She has received IBM Faculty Award 2013. Her h-index is 22 as reported by Google Scholar. She is author of book titled "Empirical Research in Software Engineering" published by CRC press and co-author of a book on "Object Oriented Software Engineering" published by PHI Learning. Her research interests are in software testing, improving software quality, statistical and adaptive prediction models, software metrics and the definition and validation of software metrics. She has published more than 140 research papers in international journals and conference



Samarth Gupta is pursuing B. Tech. in Computer Engineering from Delhi Technological University, Delhi, India. He has published Research Paper titled 'Object Detection in Foggy Conditions by Fusion of Saliency Map and YOLO' at International Conference of Sensing Technology held at University of Limerick, Ireland from 3-6 December 2018. He is an active member of IEEE and the recipient of IEEE Merwin Scholarship 2019. His fields of interest are Web Development, Machine Learning, Blockchain and Data Structures & Algorithms.



Sarthak Katyal is pursuing B. Tech. in Computer Engineering from Delhi Technological University, Delhi, India. He has published research papers titled 'Object Detection in Foggy Conditions by Fusion of Saliency Map and YOLO' and "Effective Analysis and Diagnosis of Liver Disorder by Data Mining" presented at International Conference of Sensing Technology held at University of Limerick, Ireland from 3-6 December 2018 and International Conference on Inventive Research in Computing Applications respectively.