JOIN (Jurnal Online Informatika)
p-ISSN: 2528-1682, e-ISSN: 2527-9165
Volume 9 Number 1 | June 2024: 10-17
DOI: 10.15575/join.v9i1.1203

Modeling Face Detection Application Using Convolutional Neural Network and Face-API for Effective and Efficient Online Attendance Tracking

Carles Juliandy[1], Ng Poi Wong[2], Darwin[3]
[1]Department of Information Technology, Universitas Mikroskil, Indonesia
[2,3]Department Computer Science, Universitas Mikroskil, Indonesia

Article Info

Article history:

Received October 12, 2023
Revised January 27, 2024
Accepted January 30, 2024
Available Online April 23, 2024

Keywords:

Convolutional Neural
Face-API
Network
Online Meeting Class
Tracking Attendance

ABSTRACT

The pandemic of Covid-19 emergency has ended, but it gives us a new lifestyle every aspect of life and also in the education aspect has changed. At that moment as one of the ways to prevent pandemic infection, many governments give the policy to close the offline class and continue with online classes. The online class system encountered several problems and one of those problems was to track the students' attendance to ensure all the students were attending the class. The teacher needed extra effort to track it because they needed to call the students one by one which is wasting time and sometimes would miss the presence of the students who attend the class. To make it effective efficient accurate and time-consuming when tracking attendance in online classes for teachers, we proposed the face detection model which combines face-api.js and CNN to detect and recognize the students' faces to help teachers track attendance by just uploading the screenshot image of the online meeting application. We tested our model with accuracy and speed testing. With 3 images of every student's face as training data, our model was able to recognize the face with 100% accuracy in just 41,65 seconds which is faster than calling students one by one that need almost 3 to 5 minutes if there are many students. Future research can be done by focusing research on improving the model to detect the students' faces with different brightness, contrast, and saturation because students may not have the same place and condition when joining an online meeting class.

Corresponding Author:

Carles Juliandy,
Information Technology Department, Faculty of Informatics, Universitas Mikroskil
Jl. M.H. Thamrin No. 140, Medan Kota, Kota Medan, Indonesia. 20212
Email: carles.juliandy@mikroskil.ac.id

1. INTRODUCTION

The novel coronavirus which was well-known as a global pandemic (2019-nCoV) changed every aspect of life [1]. Till 31st August 2021, it has almost more than 200 million positive cases [2]. This pandemic had no best-practice treatments after the infection, so prevention is the best way and crucial [3]. In the education field, the government closed down offline classes and continued with the online class method as a prevention step [1]. Even World Health Organization (WHO) has announced the end of COVID-19 as a Public Health Emergency of International Concern (PHEIC) [4], the online class and meeting method has still been adopted now because it brings us a better lifestyle so the improvement of online class method still needed.

The reality of online classes meets several problems and impacts the process and quality of learning [5]. This problem was caused by the lack of a measuring point to track the student performance because of postponed exams and also the lack of teachers to be aware of the student's attendance [6]. Several reasons such as systems like online meeting systems, methods, resources, management, and

students themselves caused the problems in recording the student's attendance, that's why the teachers have so lack awareness on this point [7]. There were several previous research was done to solve this problem such as creating e-learning systems and evaluating them [8], online attendance systems based on fingerprint and GPS on a smartphone [9], or tracking systems using Android Mobile Phone [10]. However, that solution may not apply to all schools and still needs more effort from students and teachers.

Based on previous research on to use of technology in the education field [11], in this research, we tried to use technology in our research, we proposed our model with a different approach to build an online attendance system using face detection on online system meetings using the screenshot image. So our proposed model, it hopes will lower teacher efforts to track students' attendance. Our proposed model was built using face-api.js [12] to detect face position because previously researched by Li in 2019 showed that face-api.js was a better tool to detect faces and used it to develop applications [13], meanwhile, a convolution neural network to classify the face.

Previous research about face recognition was done by Zhao in 2017 with improved LBPH that can detect face position and classify it accurately. Hence this research didn't propose anything about accuracy and time [14]. The research using face-api.js was done in 2020 by Vinod. In that research using face-api.js and combination with biometric features such as fingerprint and iris recognition, however, this innovation may not be suitable for all schools [15]. The success of Convolutional Neural Network (CNN) in segmentation was proved in 2020 by Sultana who proposed survey research that concluded that CNN was able to classify images based on the application task [16].

The contribution of this research was to combine face-api.js and CNN to provide attendance systems for online meetings or online class schools. Teachers just need to take a screenshot of the online class such as Zoom or Google Meet and then our proposed model would detect where the face using face-api.js and detect whose face was using CNN based on student datasets that are stored inside. After detection, our proposed model would be able to store the attendance in Excel file format to make it easy to record the attendance.

Because our proposed model was based on classification. So, for evaluating the result we use a confusion matrix because it was a well-known or fundamental role in evaluating the classification performance [17].

The remaining paper is structured as follows: Section 2 provides background research related to face detection, online attendance, and CNN. Section 3 presents the research methodology or our approach to record attendance. Section 4 provides the result and discussion and Section 5 provides the conclusion of the paper.

2. RELATED WORKS

Previous research was done to track students' attendance using the online method or another method, but no one was suitable in this pandemic situation especially in Indonesia because previous research needed huge infrastructure improvement that will be something hard to realize in Indonesia when in this pandemic situation.

Before the pandemic, in 2017, Milon researched online attendance systems using smartphone-based. A combination of the smartphone's local storage and web storage would make the synchronization well. With the ability to send emails and SMS to parents, it would keep the parents up to date about their children's attendance. But with the assumption, there were not all children have a smartphone, this model implementation would face some problems.

In 2018, Lia researched the use of fingerprint modules and GPS on a smartphone to build a real-time online attendance system. This model was built for office workers with great response time and a low error rate. It has worked greatly for workers but unfortunately, the model would cause some problems in realization for students, because not all students will have a smartphone with a fingerprint and GPS [9].

After the pandemic, in 2020, Asri researched a smart attendance system using the Quick Response (QR) code to track and record students' attendance. However, this research was not focused on solving the pandemic-era attendance system and the implementation would face several problems like other previous research. The main problem was how to make sure all children would have a phone especially a smartphone [18].

Based on previous research to detect and recognize a face from an image it was needed a computer vision algorithm was, because some problems with the quality of the image such as noise,

contrast changing, or blurred image would lead to lower image detection quality too [19]. One of the most used algorithms based on previous research was the Convolutional Neural Network [20].

Convolution Neural Network (CNN) was inspired by the human visual cortex, which means that CNN tried to make what humans saw as what computers saw too [21]. CNN is nowadays known as one of the best solutions for many computer vision problems [22]. From previous research, CNN was also used as a classification algorithm, such as classifying heartbeats [23], modeling sentences to find a related word to complete sentences [24], survey research in the deep convolutional neural network for image segmentation [16], computer vision for animal ecology [25] and facial expression recognition [26]. However previous research didn't focus anything on the online attendance area but just focused on the face detection and recognition area. With its ability as a classification algorithm and its ability to track and recognize the face, we are sure, that CNN will be able to recognize the face detected by face-api.js.

3. METHOD

In this research, we conducted the model of an online attendance application using the screenshot of an online meeting image. Previously the teachers needed to track their students' attendance manually and needed to take the screenshot of the online meeting class as evidence that the class was being held consistently, so the teachers must use more effort. but with our proposed model the teachers just needed to take screenshots of online meetings like usual and just upload the screenshots image to our model to track who attended the class.

The proposed model also prepared a report with spreadsheet extension format, so the teachers can also be used it as a report of attendant students to school. With this model, we sure would reduce the teacher's effort in tracking students' attendance and will reduce consumption time when the teachers take the attendance one by one student.

To simple understand our proposed model in this research, it can be seen in Figure 1 that describe clearly our proposed model.

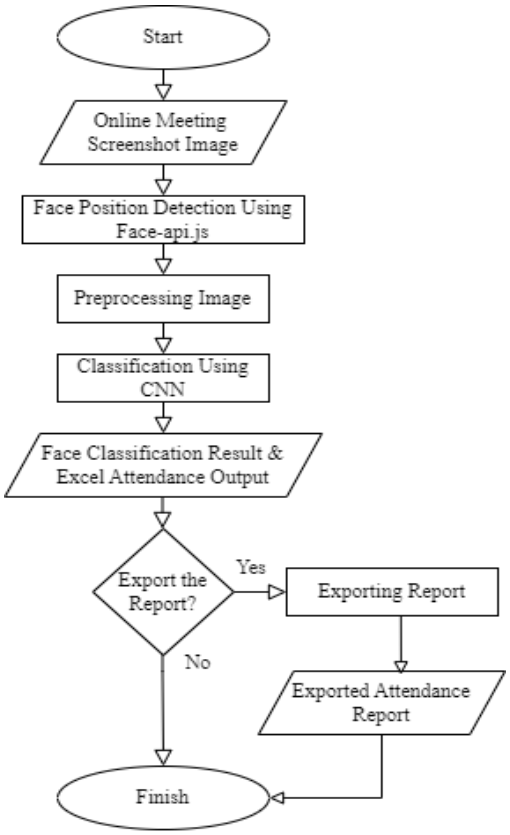

Figure 1. Flowchart of Proposed Model

Based on the flowchart in Figure 1, can be explained step by step the process in our proposed model was:

1) Screenshot image input
   Insert the screenshot image of an online class meeting that must clearly show the students' and teachers' faces.
2) Face position detection
   Detect where the face position from the screenshot image using face-api.js. Based on TF.JS, it provides a lot of out-of-box models and APIs for face recognition.
3) Preprocessing image
   We divided preprocessing into 2 different steps:
   a. Preprocessing for Training Dataset
      To train the dataset, the image on the dataset should follow these steps to completely train the dataset:
      a) Resizing
         Resize all the dataset images into 48 x 48 pixels to make the image size identically to others
      b) Grayscaling
      c) For grayscaling the image, we use the formula:
         $Grayscale = 0.299R + 0.587G + 0.114B$
         Grayscale = (0,299 x 155) + (0,587 x 120) + (0,114 x 135)
         Grayscale = 46,345 + 70,44 + 15,39
         Grayscale = 132,175
      d) Thresholding
         Convert a grayscale image into a binary image. For this process, we used the formula:
         C(T)=M1(T)·σ12.(T)+M2(T)·σ12 (T)
         Where:
         M1 was the amount of the pixel
         M2 was the amount of the remaining pixel
         σ12 dan σ22 was a variant value from the pixel
         C = ((16*16).0.4637) + ((16*16).0.5152)
         C = 118.7072 + 131.8912
         C = 250.5894
   b. Preprocessing for Testing Dataset
      To test our model, we use the screenshot image from an online meeting application. For preparing the testing image to be ready for face classification we need to preprocess the screenshot image. The preprocessing steps were the same as the data training process from steps a) to c).
4) Classification using CNN
   After the preprocessing steps, we extracted the feature from the image using the landmark point method. With this landmark point, we took the position of nose, jaw, temples, and mouth. After getting the landmark point of the testing image, we need to train the CNN to classify the testing image. The details of the training process can be seen in Figure 2 below

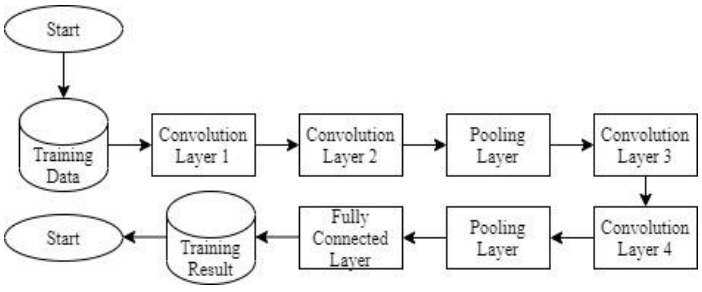

Figure 2. Convolutional Neural Network Data Training

For the training process, we need to determine filter, pool size, stride, padding, optimizer, batch size, epoch, and activation function. The detail of each feature we use was:

a. Filter = 3 x 3
b. Pool Size = 2 x 2

c. Stride and Pading we set to default 1 and 0
d. The padding we used was Adam, Stochastic Gradient Descent (SGD), and Root Mean Square Propagation (RMSprop)
e. Batch Size = 32 and 128
f. Epoch = 50, 100, 200, and 300
g. The activation Function we use was softmax on the output layer.

The detail of each layer training can be seen in Table 1 below

Table 1. Detail of Each Layer Training

Layer	Size	Parameter
Input	48x48x1	0
Conv 1 + ReLu	46x46x64	(3 x 3 x 1 + 1) x 64 = 640
Pool 1	23x23x64	0
Norm	23x23x64	64 x 4 = 256
Conv 2 + ReLu	21x21x128	(3 x 3 x 64 + 1) x 128 = 73.856
Pool 2	10x10x128	0
Norm	10x10x128	128 x 4 = 512
Conv 3 + ReLu	8x8x256	(3 x 3 x 128 + 1) x 256 = 295.168
Norm	4x4x256	256 x 4 = 1.024
Conv 4 + ReLu	2x2x512	(3 x 3 x 256 + 1) x 512 = 1.180.160
Norm	1x1x512	512 x 4 = 2.048
Dropout	1x1x512	0
Flatten	512	0
FC + Softmax	256	(512 + 1) x 256 = 131.328
Dropout	4	(256 + 1) x 4 = 1028

5) Face recognition output based on data training and show the name of the students based on 1 and 0 CNN matching.
6) An optional feature for the teacher to export the attendance of the students into a spreadsheet file or not. The report would consist of student ID number, student name, date, and attendance or absence.

4. RESULT AND DISCUSSION

In this research, the training dataset we used was the image of students' faces from any angle. In this model, we collected 4 images of 12 students and the teacher, so there are 13 persons in that class. The image was a self-image of the student that closed up the face, so it could extract the features of the face better without any noise. We have a total of 52 images that combine students' and teachers' faces.

We conducted the testing in two big scenarios, the first testing was accuracy testing, and the second was time consumption testing to complete the face detection. For accuracy testing, we tried the face recognition ability by changing the number of data training for each person. We tried to change the data training from 4 images of each person to 1 image of each person and then we compared the accuracy. The online meeting application we used in this research was Zoom.

For the first testing scheme, we started by using 4 images of each person that show the face as the training data. We finished the accuracy testing by reducing the training data 1 by 1 for each person until it would remain just 1 image in the training data. We used the screenshot of the online class. That class has 12 students and 1 teacher, but in the screenshot, there are only 9 students and 1 teacher. The Face-API.js testing result for face recognition is successful in detecting 10 people in the image, and the CNN also works well and can recognize and classify it as the right people, we got 100% accuracy when there are 4 and 3 images of each people in the training data. The details of the face recognizing can be seen in Figure 3.

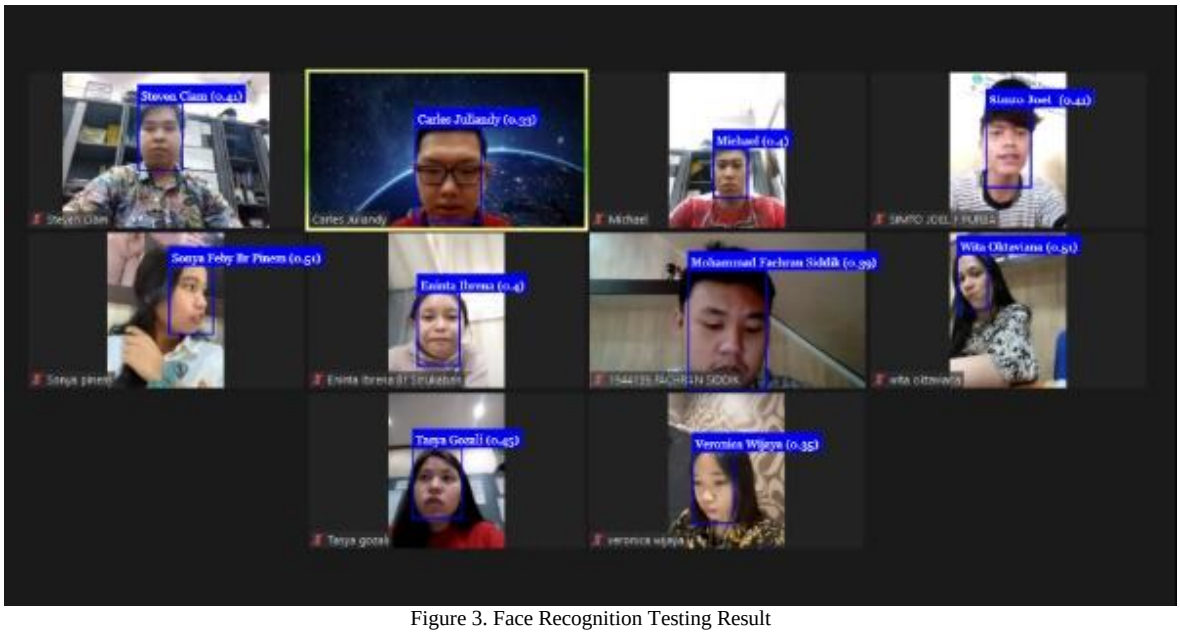

Figure 3. Face Recognition Testing Result

To evaluate the accuracy of our proposed model, we used a confusion matrix. The confusion matrix operators we used in this testing were:

a. True Positive (TP) for face-api.js can detect the face and CNN classifies it as the right people
b. True Negative (TN) for face-api.js did not detect the face and CNN did not classify it
c. False Positive (FP) for face-api.js can detect the face and CNN classified it as the wrong people
d. False Negative (FN) for face-api.js did not detect the face and CNN classified it as the wrong people.

The details of our accuracy testing can be seen in Table 2 below.

Table 2. Confusion Matrix for Accuracy Testing

Training Dataset/People	People to Recognized	Result			
		TP	TN	FP	FN
4	10	10	0	0	0
3	10	10	0	0	0
2	10	7	0	3	0
1	10	4	0	6	0

From the confusion matrix, we got that when the training dataset with the 3 or 4 images for each person will give us 100% accuracy recognition. On the other side, the 1 and 2 images for data training for each person just get 70% and 40% accuracy. So, for our proposed model, we were sure our model was robust enough to recognize the people's faces with only 3 to 4 images in the dataset as data training for each person.

After testing the accuracy of the model, the next step we would test the speed of our proposed model by testing the consumption time in recognizing the face. We run the test on the processor Intel ® CoreTM i5 @ 2.7 GHz with graphic card NVIDIA GeForce GTX 1050. To compare the speed, we used the different numbers of the image in the training dataset. The speed test results in detail can be seen in Table 3 below.

Table 3. Time Consumption Testing

Number of Images in Dataset for Each People (13 people)	Time to Recognized (s)
4	49,21
3	41.65
2	35.37
1	30.94

Table 3 showed that more images in the dataset as data training would inflict the speed for recognizing the face. This can happen because our proposed model would get the convolution data of the

image one by one to increase the accuracy. It can be seen that our proposed model able to recognize the face with only 3 images in the dataset for each person to get 100% accuracy. So, it can be concluded that the longer time needed would inflict higher accuracy too. So, in our proposed model, we should find the balance point between accuracy and time consumption. The best time in our testing model was 3 images for each people in the classroom and it will get 100% accuracy.

For the final report, we also prepared the spreadsheet file format to help the teachers easily track student attendance. The teacher just needs to input the class name and our proposed model will get the data that was consisted of student ID, and student name based on the class name. The complete report would include the date and attendance status. The details of the attendance report can be seen in Figure 4 below

Student_ID	Student_Name	Date	Attendance
2144001	Steven Ciam	9/10/2021	A
2144002	Michael	9/10/2021	A
2144003	Simto Joel	9/10/2021	A
2144004	Evan Rahmat	9/10/2021	NA
2144005	Sonya Febi Br. Pinem	9/10/2021	A
2144006	Eninta Ibrena	9/10/2021	A
2144007	Mohammad Fachran Siddik	9/10/2021	A
2144008	Wita Oktaviana	9/10/2021	A
2144009	Johan	9/10/2021	NA
2144010	Tasya Gozali	9/10/2021	A
2144011	Veronica Wijaya	9/10/2021	A
2144012	Naomi Sulastrini	9/10/2021	NA
T_0001	Carles Juliandy (Teacher)	9/10/2021	A

Figure 4. Final Attendance Report

As we can see in Figure 4, there is a list of the person in that class which are 12 students and 1 teacher. When the faces are successfully recognized by our model, it will give the “A” status in the attendance column which means attending the class, and for the person which is fails to be detected by our model, it will give the status “NA” which is mean not attending the class.

From the testing result, we can conclude that our proposed model works well in detecting and recognizing the student and teacher faces. The model can recognize the person in the screenshot with 100% accuracy in only 41,65 seconds when there are 3 images of each person in the training data. With less than 1 minute for the teacher to track the attendance of the whole class student, we are sure that our model is not only effective but also efficient because it is more time-saving in tracking student attendance manually or only using the report from the online meeting application and input it again to the attendance report.

5. CONCLUSION

The pandemic Covid-19 emergency has been ended by the World Health Organization (WHO), but the changed lifestyles in several aspects such as online school carry over now. However, in the implementation, the online school meets some problems to make sure the learning quality is maximal. For the teacher online school is a waste of time when tracking the students' attendance because it needs to be tracked one-by-one and there is a high risk of human error while tracking. We proposed a model combining face-api.js and CNN for face recognition to track online class attendance. The proposed model works well in recognizing the attendance of people who join the online class with 100% accuracy and needs only 3 images as training data for each person in 41,65 seconds. The teacher is also available to get the report in spreadsheet format to quickly make the attendance list, so the teacher will save more time in tracking the student attendance while taking online classes. Future research can be done by focusing on the improvement to recognizing the face at different brightness, contrast, and saturation because every student may have not the same place and condition when attending the online meeting class another improvement was how to increase the speed to recognize it, so the teacher can tracking the attendance faster.

REFERENCES

[1] A. A. Oloyede, N. Faruk, and W. O. Raji, “COVID-19 lockdown and remote attendance teaching in developing countries: A review of some online pedagogical resources,” *African Journal of Science, Technology, Innovation and Development*, vol. 0, no. 0, pp. 1–19, 2021, doi: 10.1080/20421338.2021.1889768.

[2] "Coronavirus Outbreak," Website page. Accessed: Sep. 01, 2021. [Online]. Available: https://www.worldometers.info/coronavirus/
[3] T. Singhal, "A Review of Coronavirus Disease-2019 (COVID-19)," *Indian J Pediatr*, vol. 87, no. 4, pp. 281–286, 2020, doi: 10.1007/s12098-020-03263-6.
[4] World Health Organization (WHO), "Statement on the fifteenth meeting of the IHR (2005) Emergency Committee on the COVID-19 pandemic," https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic.
[5] A. Supriyanto, S. Hartini, W. N. Irdasari, A. Miftahul, S. Oktapiana, and S. D. Mumpuni, "Teacher professional quality: Counselling services with technology in Pandemic Covid-19," *Counsellia: Jurnal Bimbingan dan Konseling*, vol. 10, no. 2, p. 176, 2020, doi: 10.25273/counsellia.v10i2.7768.
[6] H. H. S. Simon Burgess, "Schools, skills, and learning: The impact of COVID-19 on education." Accessed: Sep. 01, 2021. [Online]. Available: https://voxeu.org/article/impact-covid-19-education
[7] R. Taufiq, M. Baharun, B. Sunaryo, B. Pudjoatmodjo, and W. M. Utomo, "Indonesia: Covid-19 and E-Learning in Student Attendance Method," *SciTech Framework*, vol. 2, no. 1, pp. 12–22, 2020.
[8] B. Setiawan, Sofyan Rofi, and Tri Endang Jatmikowati, "The Student Learning Activity Levels on the Online Learning During the Covid-19 Pandemic," *Jurnal Pendidikan Islam Indonesia*, vol. 5, no. 2, pp. 186–197, 2021, doi: 10.35316/jpii.v5i2.289.
[9] L. Kamelia, E. A. D. Hamidi, W. Darmalaksana, and A. Nugraha, "Real-Time Online Attendance System Based on Fingerprint and GPS in the Smartphone," *Proceeding of 2018 4th International Conference on Wireless and Telematics, ICWT 2018*, pp. 2018–2021, 2018, doi: 10.1109/ICWT.2018.8527837.
[10] P. S. Helode, Dr. K. H. Walse, and Karande M.U., "An Online Secure Social Networking with Friend Discovery System," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 5, no. 4, pp. 8198–8205, 2017, doi: 10.15680/IJIRCCE.2017.
[11] Andi, R. Purba, and R. Yunis, "Application of Blockchain Technology to Prevent The Potential Of Plagiarism in Scientific Publication," *Proceedings of 2019 4th International Conference on Informatics and Computing, ICIC 2019*, 2019, doi: 10.1109/ICIC47613.2019.8985920.
[12] "Face-API JS." [Online]. Available: https://justadudewhohacks.github.io/face-api.js
[13] C. Li and C. Li, "Web Front-End Realtime Face Recognition Based on TFJS," *Proceedings - 2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, CISP-BMEI 2019*, 2019, doi: 10.1109/CISP-BMEI48845.2019.8965963.
[14] X. M. Zhao and C. B. Wei, "A real-time face recognition system based on the improved LBPH algorithm," *2017 IEEE 2nd International Conference on Signal and Image Processing, ICSIP 2017*, vol. 2017-Janua, pp. 72–76, 2017, doi: 10.1109/SIPROCESS.2017.8124508.
[15] V. Vibin Mammen, S. Thokaiandal, C. S. Sindhuja, V. Mekala, M. Manimegalai, and N. Prabhuram, "A comprehensive study on academic and industry authentication and attendance systems," *International Journal of Scientific and Technology Research*, vol. 9, no. 3, pp. 5426–5432, 2020.
[16] F. Sultana, A. Sufian, and P. Dutta, "Evolution of Image Segmentation using Deep Convolutional Neural Network: A Survey," *Knowl Based Syst*, vol. 201–202, pp. 1–38, 2020, doi: 10.1016/j.knosys.2020.106062.
[17] J. Xu, Y. Zhang, and D. Miao, "Three-way confusion matrix for classification: A measure driven view," *Inf Sci (N Y)*, vol. 507, pp. 772–794, 2020, doi: 10.1016/j.ins.2019.06.064.
[18] A. Nuhi, A. Memeti, F. Imeri, and B. Cico, "Smart Attendance System using QR Code," *2020 9th Mediterranean Conference on Embedded Computing, MECO 2020*, pp. 8–11, 2020, doi: 10.1109/MECO49872.2020.9134225.
[19] D. Yin, R. G. Lopes, J. Shlens, E. D. Cubuk, and J. Gilmer, "A Fourier perspective on model robustness in computer vision," *Adv Neural Inf Process Syst*, vol. 32, no. NeurIPS, 2019.
[20] M. Yu, L. Gong, and S. Kollias, "Computer vision based fall detection by a convolutional neural network," *ICMI 2017 - Proceedings of the 19th ACM International Conference on Multimodal Interaction*, vol. 2017-Janua, pp. 416–420, 2017, doi: 10.1145/3136755.3136802.
[21] Y. Jia *et al.*, "Caffe," pp. 675–678, 2014, doi: 10.1145/2647868.2654889.
[22] Z. Qin, F. Yu, C. Liu, and X. Chen, "How convolutional neural networks see the world --- A survey of convolutional neural network visualization methods," *Mathematical Foundations of Computing*, vol. 1, no. 2, pp. 149–180, 2018, doi: 10.3934/mfc.2018008.
[23] U. R. Acharya *et al.*, "A deep convolutional neural network model to classify heartbeats," *Comput Biol Med*, vol. 89, no. September, pp. 389–396, 2017, doi: 10.1016/j.compbiomed.2017.08.022.
[24] N. Kalchbrenner, E. Grefenstette, and P. Blunsom, "A convolutional neural network for modelling sentences," *52nd Annual Meeting of the Association for Computational Linguistics, ACL 2014 - Proceedings of the Conference*, vol. 1, pp. 655–665, 2014, doi: 10.3115/v1/p14-1062.
[25] B. G. Weinstein, "A computer vision for animal ecology," *Journal of Animal Ecology*, vol. 87, no. 3, pp. 533–545, 2018, doi: 10.1111/1365-2656.12780.
[26] K. Shan, J. Guo, W. You, D. Lu, and R. Bie, "Automatic facial expression recognition based on a deep convolutional-neural-network structure," *Proceedings - 2017 15th IEEE/ACIS International Conference on Software Engineering Research, Management and Applications, SERA 2017*, pp. 123–128, 2017, doi: 10.1109/SERA.2017.7965717.