

COVID-19 SURVEILLANCE SYSTEM USING COMPUTER VISION

INIYAN M¹, ASST.PROF TAMIZHARASI T²

¹*M.Tech in Artificial Intelligence and Machine Learning,School of Computer Science,VIT University,Vellore,India*

² *Asst. Prof- School of Computer Science ,VIT University, Vellore, India*

Abstract - A Highly contagious disease named COVID-19 that is created by a newly discovered virus. Without any need for Special treatment people who are affected by this virus can recover. Elderly and people with autism can develop serious diseases such as cardiovascular disease, diabetes, respiratory kind of disease, and cancer. The most well-known is about COVID-19, the way how it affects the people and spread all around the world. This disease is spread through the mucus and is expelled by the person when he/she sneezes. The best approach to prevent ourselves from this disease is to wash our hands regularly and using the alcohol based liquid. There are no specific treatments or vaccines discovered at a particular time. But we can prevent from spreading by following several measures. It includes like wearing masks, using alcohol based liquid, frequent temperature checking. The main idea of this project work is to detect the existence of masks in videos and images directly. To develop this model, deep learning technique is used. To make this system work fast in a rapid speed and efficiently Deep learning technique is used. Finally, the outcome will indicate the existence or absence of the mask in the given input video or image.

Key Words: Corona Virus, object detection, Computer Vision, Face Mask, Object Tracking

1.INTRODUCTION

2020 presents a list of things that will amaze the humanity. Among those, the Corona virus epidemic is the most life-changing event that has shaken the world since the beginning of the year. Related to the health and well-being of the poor, COVID-19 must take effective measures to prevent the spread of the disease. From basic hygiene to all health care, people all around the world are doing everything they can to keep themselves and the community healthy. Public Places like many Super Markets, Grocery Stores, Public Transport Stations are crowded as the local economy is unlocked. We need a tremendous number of staff to offer the best service in hotels, markets and trucks. Therefore, Human Rights Management not only uses implementation plans to implement this project through the right amount of customer data as a measure of market popularity in department stores and supermarkets. The quantity of clients

is vital data for settling on the business choices. It not exclusively is utilized as a record for shopping fame of retail chain or hypermarkets, yet additionally causes chiefs to orchestrate the arrangement of work force. The quantity of clients of a year is frequently used to foresee the group top periods. The number of entrants is typically used in a year to predict the maximum period of the season. At exhibitions and playgrounds, you can count the number of visitors by the door and provide information to the staff campus. It not exclusively is utilized as a record for shopping fame of retail chain or hypermarkets, yet additionally causes chiefs to orchestrate the arrangement of work force. The quantity of clients of a year is frequently used to foresee the group top periods. In the presentation or arena, we can check the quantity of guests through the passageway and give the data to the premises security faculty. Powerful control the quantity of guests is vital. Accordingly, the data of the quantity of individuals through the passageway has high reference esteem.

These include:

- Contact Free Access Control, a contactless section and exit of work force like representatives and guests inside any association and naturally produces a guest pass
- Face Mask Detection is a CCTV-based reconnaissance that utilizes semantic division for distinguishing individuals without veils
- Social Distance Detection guarantees that there is no group arrangement past indicated restricts in broad daylight spots and government premises
- Thermal Scanning utilizes insightful warm innovation and checks the temperature of everybody entering the premises and triggers vital cautions in the event of any abnormality in the temperature
- Zone Monitoring is a keen coordinated video investigation framework that screens unapproved

developments in confined zones or regions forced by the specialists

- The whole framework is an ideal fit in assorted applications like astute observing of vehicles, swarms, foundation, objects for programmed location and examination of occasions which can be conveyed in working environments, shopping centers, air terminals, rail route stations, lodging social orders, building destinations, producing locales among others.
- CCTV outside places of individuals who have tried positive and monitor their action.

Strategies need to be developed to monitor how the community meets these safety standards. We can use the mask detection system to check it. Identifying a mask means finding out if someone is wearing a mask. The initial step to perceive the presence of a veil on the face is to distinguish the face, which makes the procedure isolated into two sections: to recognize faces and to distinguish covers on those countenances. Face discovery is one of the uses of article recognition and can be utilized in numerous zones like security, biometrics, law implementation and that's only the tip of the iceberg. There are numerous indicator frameworks created around the planet and being executed. Notwithstanding, all this science requires advancement; a superior, more exact locator, in light of the fact that the world can't manage the cost of any more expansion in crown cases. In this task, we will be building up a face cover locator that can recognize faces with covers and faces with no masks. In this report, we have proposed a locator which utilizes DL technique to detect the face covering mask. The execution of the calculation is on pictures, recordings and live video transfers.

2. RELATED WORK

Article ID is one of the moving subjects in the field of picture taking care of and Computer vision. Going from restricted extension singular applications to gigantic degree mechanical applications, object acknowledgment likewise, affirmation is used in a wide extent of undertakings. A couple of models fuse picture recuperation, security and knowledge, OCR, clinical imaging and cultivating checking.

In object recognizable proof, an image is scrutinized and in any event one articles in that image are organized. The territory of those things is similarly shown by a cutoff called as the bouncing box. For the most part, authors used model affirmation to expect faces reliant on prior face models. A

forward jump face distinguishing proof development by then was made that was an improved system of using Haar, electronic picture features used in recognizing the object. Regardless, it failed since it didn't perform well on faces in dull districts and non-front facing appearances. Starting now and into the foreseeable future, researchers are restless to develop new estimations reliant on significant sorting out some way to improve the models. DL grants us to learn features with beginning to end way and wiping out the need to utilize prior data for outlining feature extractors.

A few papers in regards to picture preparing been concentrated to remove the idea of utilizing PC vision that will be utilized for live reconnaissance for Coronavirus avoidance Detection the peculiarity identified with infringement of lawful direction by a person on foot or public[1]. This method can use an unsupervised learning algorithm to understand the exact path to the type of object. The operator (human) does not contribute to this script. The author have shown that it is more efficient to use a gray image to perform this task. They suggested ways to avoid many things and move things around to find out what's missing. They used this method to automatically detect this intrusion. People are automatically monitored based on how their faces are identified. Even multiple passengers at once pass facial expressions to the front door to determine if a candidate is a face. Finally, track faces of people who are moving and count people who are not moving. To tackle the issue of the light, we applied casing contrast calculation and NCC (Normalized Color Coordinates) shading space for distinguishing the face competitor locale. At that point, we apply the component of face to decide if the competitor is face or not. At long last, they track the essence of the moving individuals, and tally the moving individuals. For the consolidation/split issues, we additionally give a compelling strategy. The examinations results show that the proposed strategy can accomplish a high precision and can serve the continuous undertaking. The exactness is over 80 percent. The blunders perhaps caused if the human face covered with the veils or the hair, and so forth Later on, we can think about more cases, and adjust our proposed calculation to get higher and exact accuracy[2]. They suggest how to move objects and follow the steps to identify their bad condition. Use this method to identify automated events as traffic accidents. Or use the wrong road with a car or pedestrian. In the future, people will identify some shortcomings, such as double lanes and the use of special public transport. An article introduces new ways to adjust the accuracy of 3D routes for tourists. If the direction of the event doesn't have a spot with any of the pre-taken care of classes, it is set in the

game plan of dark tracks, that can dynamically be identified with another class or the specific action reliant on the amount of occasions, making the system self-configurable and adaptable[5].

A paper that centers around street mishaps and street irregularities. The methodology for location the peculiarity identified with infringement of legitimate direction by a person on foot or a vehicle. This methodology can perceive the right direction for the specific class of items with utilizing solo learning calculation. The administrator (human) doesn't help this marking. The author showed that utilizing grayscale pictures for tackling this assignment is more reasonable. They recommended the path for extraction continuous directions for the moving articles and for inconsistency discovery from them. They utilized this methodology for naturally location such abnormalities as vehicle episodes or utilizing incorrectly direction via vehicle or people on foot. Later on, They tried to attempt to identify different inconsistencies, for example, twofold line convergence or utilizing extraordinary public vehicle lines[3].

The objections of the paper was fuse of the face and bystander disclosure was productive and executed the structure to the traffic light system to change the sign advancement of the traffic light. In joining of the two disclosures, there were two polygons present in perceiving a walker. Considering the enlisted worth using the Chi-square, it contemplated that the preliminary or the system can recognize a bystander's body and face and the accomplishment speed of 87.3 percent. It also construed that there were a couple of segments to consider that impacts to recognize an individual by walking, one of it was the lighting of the environment and second was the stuff that a bystander was wearing like umbrellas, covers, eyeglasses and besides the shade of the pieces of clothing that a walker wearing, thus forth[6].

3. EXISTING WORK

Numerous organizations and free lancers have given comparable answer for the issue. The vast majority of it are intended to deal with machines with high necessities. Business utilization of this is limited to just to organizations that can bear the cost of these machines.

This makes it monetarily difficult to reach to limited scope organizations. These models work on prior datasets and no live information is gathered and prepared intermittently which may limit us from utilizing it in various country. The

preparation models are gigantic and sets aside huge effort to measure also, it is deployed. In Existing Work, they attempted to prepare the model for epochs of ten what's more, in the wake of finding ideal age, they saved the model for almost eight epochs to test on the genuine information. To get free of the issue of impediments of the face which causes inconvenience face finders to distinguish veils in the pictures, they utilized an underlying OpenCV profound learning face identification model. For example, the Haar-Cascade model could be utilized yet the issue of the Haar-Cascade model is that the identification outline is a square shape, not a square. That is the reason, without catching the segment of the foundation, the face casing can fit the sum of the face, which can meddle with the face cover model expectations.

4. PROPOSED SYSTEM

Every one of the three issues are tended to in the proposed framework like covers, separating and cleaning. It Can be run on negligible equipment and programming necessities. Pictures are gathered occasionally and ordered and put away in dataset to prepare the model occasionally. This model is precise, and since we utilized the MobileNetV2 design, it's likewise computationally productive and consequently making it simpler to send the model to implanted frameworks. This task can be incorporated with implanted frameworks for application in air terminals, rail route stations, workplaces, schools, and public spots to guarantee that public security rules are followed. we have fabricated an ongoing face veil discovery model utilizing profound discovering that distinguishes faces with or without cover and is sent with TensorFlow and Keras. This face veil identification model has accomplished accuracy of 94 percent. This specific model can be conveyed at different spots like the clinics, workplaces, air terminals to give some examples, as this model distinguishes if an individual is wearing the face cover, to stop the disease it is considered as on the method to prevent one people from each other during this pandemic and by observing the position of the face cover on the face, we can ensure that an individual wears it the correct way.

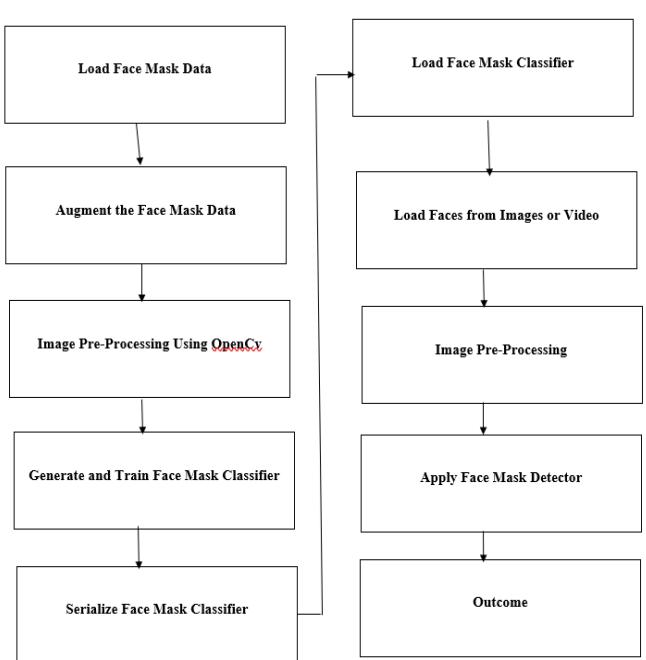


Fig 1 - Architecture of the Proposed System

5. METHODOLOGY

5.1 DATASET

The dataset which we have used contains 3830 complete pictures out of which 1916 are of covered with mask appearances and 1916 are of uncovered faces. Every one of the photos are real pictures isolated from Bing Search Programming interface, Kaggle datasets. From every one of the sources, the degree of the pictures is same. The photos cover various races that is Asian, Caucasian, etc. The degree of covered to uncovered appearances find that the dataset is changed. We need to part our dataset into three sections namely train and test dataset. The motivation behind parting information is to keep away from overfitting which is focusing on minor subtleties/commotion which isn't essential and just advances the preparation dataset exactness. We need a model that performs well on a dataset that it has never seen (test information), which is called speculation. The preparing set is the genuine subset of the dataset that we use to prepare the model. The model notices also, gains from this information and afterward streamlines its boundaries. The approval dataset is utilized to choose hyperparameters (learning rate, regularization boundaries). At the point when the model is performing great enough on our approval dataset, we can quit picking up utilizing a preparation dataset. The test set is the remaining subset of information used to give a fair-minded assessment of a last

model fit on the preparation dataset. Data is part as indicated by a split extent which is dependent on such a model we are building and the dataset itself. If our dataset and model are with the ultimate objective that a lot of getting ready is needed, by then we use a greater piece of the data just for planning which is our case. In case the model has a lot of hyperparameters that can be tuned, by then we need to take a higher proportion of endorsement dataset. Models with fewer hyperparameters are not difficult to tune and refresh, thus we can take a more modest approval dataset.

5.2 PROCESS

Face Mask discovery has been refined by receiving Deep Learning strategies. In this paper proposed model has been essentially planned into two stages and these are preparing the model and executing that model. Proposed Facemask locator model targets identifying if an individual is wearing a veil. The results are tried utilizing a live video transfer or when a picture has been given as a contribution to the face veil locator By observing the arrangement of the face cover on the face, it very well may be guaranteed that an individual wears it the privilege route assisting with forestalling the spread of Coronavirus illness. The proposed face veil identifier is least unpredictable in construction and gives speedy outcomes and in this way can be conveyed at different spots. Python programming language with jupyter journal has been utilized to build up this model.

After the model was prepared, a picture was stacked as a contribution to recognize whether an individual is wearing a cover or not. The yield so accomplished from the past advance is then stacked and pre-handled. To restrict wherein the picture all countenances are found, face recognition happens and it likewise distinguishes the ROI. Then, face covers are perceived continuously video transfers. Face recognition happens a similar path as talked about before. A specific casing from the stream is gotten and resized. After that discovery of face veil happens and whatever outcome is so acquired is shown on the specific screen.

6. ALGORITHM AND DISCUSSION

6.1 MOBILEV2NET

MobileNetV2 is a design of bottleneck profundity distinguishable convolution working of essential blocks with residuals. It has two sorts of squares. The first is a leftover square with step of 1. Second one is additionally lingering block with step 2 and it is for downsizing. In proposed model

MobileNetV2 classifier has been utilized. It is a convolutional neural organization engineering that tries to perform well on cell phones. Where MobileNet is a kind of convolutional neural organization which was intended for portable and installed vision applications. They depend on a smoothed out design that employs depthwise divisible convolutions to assemble lightweight profound neural organizations that can have low inactivity for versatile and installed devices. The Diagram underneath show the essential design of MobileNetV2. Where a step is a piece of CNN(convolutional neural network) used for tuning the pressure of pictures and video data. This specific model depends on reversed remaining construction, where lingering network is a neural organization that contains a "skip association," additionally called a "remaining association," that sidesteps a few layers of a network. It likewise utilizes the profound divisible convolution, and furthermore utilizes light weight depthwise convolutions to channel highlights as a wellspring of non-linearity.

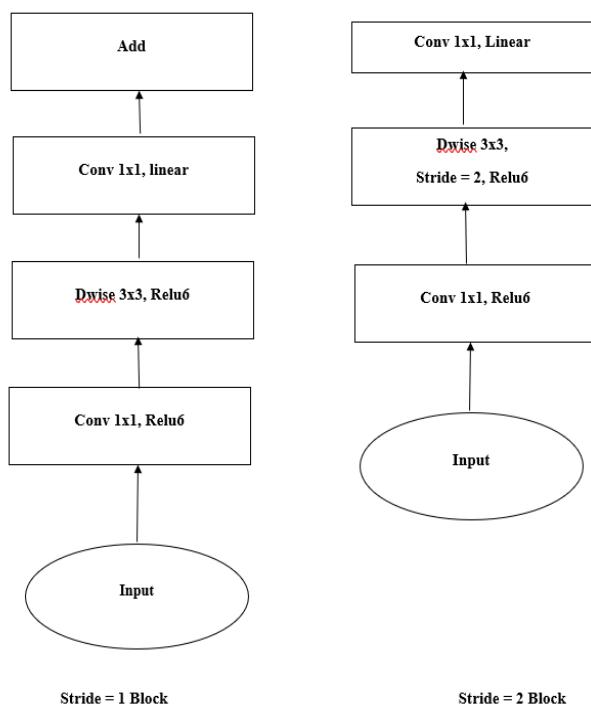


Fig 2 - Architecture of MobileV2Net

The issue can be settled in two sections: first distinguishing the presence of a few countenances in a given picture or then again transfer of video and afterward in the subsequent part, recognize the presence or nonappearance of face veil on face. To recognize the face, we have utilized the OpenCV library. The most recent OpenCV incorporates a DL technique, which accompanies a pre-prepared face recognition convolutional

neural organization (CNN). The new model upgrades the face recognition execution contrasted with the customary models. At whatever point another test picture is given, it is initially changed over into BLOBS (Binary Huge Object alludes to a gathering of associated pixels in a parallel picture) and afterward sent into the pretrained model which yields the quantity of distinguished countenances. Each face identified comes out with a level of certainty which is then contrasted with a limit an incentive with sift through the superfluous recognitions.

After we have the faces, we need to assess the jumping box around it and send it to the second piece of the model to check if the face has a cover or not. The second stage of the model is prepared by us utilizing a dataset comprising of pictures with veil and without cover. We have utilized Keras alongside Tensorflow to prepare our model. Initial segment of the preparing remembers putting away all marks of the pictures for a Numpy cluster and the relating pictures are additionally reshaped (224, 244, 3) for the base model. Picture expansion is a helpful method since it expands our dataset with pictures with an entirely different point of view. Prior to contributing, we played out the accompanying picture increases arbitrarily: revolutions up to 20 degrees, zooming in and out up to 15 percent, width or stature move up to 20%, up to 15 degrees shear point in the counterclockwise bearing, flip data sources on a level plane and focuses beyond the sources of info are filled from the closest accessible pixel of the information. For the picture characterization, it is currently a typical practice to use move realizing which means utilizing a model which has been prepared on large number of names previously and it has been tried that this strategy brings about huge expansion in exactness. Clearly, the presumption here is that both the issues have adequate closeness. It utilizes an all around organized and neural network that has been prepared on a lot of informational index. Because of to some degree same nature of the issue, we can utilize similar loads which have the ability to separate highlights and later in the profound layers, convert those highlights to objects.

The base model that we have utilized here is MobileNetV2 with the given 'ImageNet' loads. ImageNet is a picture data set that has been prepared on countless pictures subsequently it makes a difference a great deal in Image order. For the base model, we shorten the head and utilize a progression of our selfdefined layers. We utilized a normal pooling layer, a smooth layer, a thick layer with yield shape, and initiation ReLU, a half dropout layer for enhancement, at

long last another thick layer with yield shape , and Sigmoid initiation is utilized.

```
aug = ImageDataGenerator(
    rotation_range=20,
    zoom_range=0.15,
    width_shift_range=0.2,
    height_shift_range=0.2,
    shear_range=0.15,
    horizontal_flip=True,
    fill_mode="nearest")
```

Fig 3 - Snapshot of Image Augmentation

6.2 TRAINING

At the training time, for every pixel, we look at the default jumping boxes having various sizes furthermore, perspective proportions with ground truth boxes lastly use Intersection over Union technique to choose the best coordinating box. The technique assesses how much piece of our anticipated box coordinate with the ground reality. The qualities range from 0 to 1 and expanding estimations of technique decide the exactnesses in the expectation; the best worth being the most elevated estimation of the technique.

```
baseModel = MobileNetV2(weights="imagenet", include_top=False,
                        input_tensor=Input(shape=(224, 224, 3)))

headModel = baseModel.output
headModel = AveragePooling2D(pool_size=(7, 7))(headModel)
headModel = Flatten(name="flatten")(headModel)
headModel = Dense(128, activation="relu")(headModel)
headModel = Dropout(0.5)(headModel)
headModel = Dense(2, activation="softmax")(headModel)

model = Model(inputs=baseModel.input, outputs=headModel)

for layer in baseModel.layers:
    layer.trainable = False

WARNING:tensorflow:input_shape is undefined or non-square, or `rows` is not in [96, 128, 160, 192, 224]. Weights for input shape (224, 224) will be loaded as the default.
```

```
112]: # compile our model
print("[INFO] compiling model...")
opt = Adam(lr=INIT_LR, decay=INIT_LR / EPOCHS)
model.compile(loss="binary_crossentropy", optimizer=opt,
              metrics=metrics)
```

Fig 4 - Snapshot of Neural Network and Optimizer

6.3 LOSS FUNCTION

The deficiency of the general recognition issue can be broken into two primary subsets: loss of confidence and loss of localization. The restriction misfortune is only the distinction between the default anticipated bouncing box and ground truth jumping box .

```
plt.plot(np.arange(0, N), H.history["val_accuracy"], label="val_acc")
plt.title("Training Loss and Accuracy")
plt.xlabel("Epoch #")
plt.ylabel("Loss/Accuracy")
plt.legend(loc="lower left")
```

```
117]: <matplotlib.legend.Legend at 0x205b9cc1250>
```

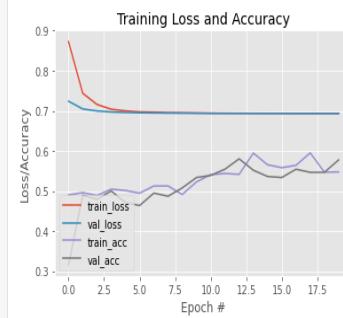


Fig 5 - Snapshot of Training loss and accuracy

7. PERFORMANCE MEASURES

7.1 ACCURACY

Accurate or predictive is the number of samples that are placed correctly and indicates how close the predicted value is to the calculation of principles and accuracy. Perfect accuracy is the percentage closest to or near the terms and actual values.

7.2 F1 SCORE

The F1 score is shown as a measure that combines precision and the recall, with the effort to convey harmony between them. The result of F1 shows how good the dealer is in terms of precision and recall. As you can see, getting one is a great strategy. Both repetitions and numerical calculations are greater to increase the F1 value. It is good if the recall, precision and accuracy value are lower than the F1 value.

	precision	recall	f1-score	support
with_mask	0.99	0.86	0.92	383
without_mask	0.88	0.99	0.93	384
accuracy			0.93	767
macro avg	0.93	0.93	0.93	767
weighted avg	0.93	0.93	0.93	767

[INFO] saving mask detector model...
dict_keys(['loss', 'accuracy', 'val_loss', 'val_accuracy'])

Fig 6 - Snapshot of Accuracy

8. RESULT AND ANALYSIS

After effectively actualizing and testing the code following yield was accomplished. Screen shots of the yields have been shown ahead. The outcome shows if the individual is wearing the face cover or not. The model so created is identified with the recognition of face covers being put on by individuals, from the pictures just as from live video transfers. On preparing the model utilizing OpenCV, Keras, Tensorflow a precision of 93% was accomplished. Post preparing, the classifier was exposed to pictures and live video transfers. The countenances were perceived in the pictures and recordings and the appearances were extricated. Thus, the face cover classifier was applied and the necessary outcomes were acquired. Likewise the model so created is least confounded in design and gives quick outcomes and henceforth can be conveyed or utilized at different spots recorded previously.

9. CONCLUSION AND FUTURE WORK

This model using DL techniques obtaines accuracy of 93 percent ad identifies whether a particular person wears the face mask or not. In excess of fifty nations around the planet have as of late started wearing face covers necessary. Individuals need to cover their countenances in broad daylight, grocery stores, public vehicles, workplaces, and stores. Retail organizations frequently use programming to check the quantity of individuals entering their stores. They may likewise like to gauge impacts on advanced presentations and limited time screens. We are intending to improve our Face Mask Detection instrument and delivery it as an open-source project. Our product can be likened to any current USB, IP cameras, and CCTV cameras to distinguish individuals without a cover. This location live video feed can be executed in web and work area applications so the administrator can see notice messages. Programming administrators can likewise get a picture in the event that somebody isn't wearing a cover.

Besides, an alert framework can likewise be executed to sound a signal when somebody without a veil enters the territory. This product can likewise be associated with the passageway doors and just individuals wearing face veils can come in.

REFERENCES

[1] Cyrel O. Manlises ; Jesus M. Martinez ; Jackson L. Belenzo ; Czarleine K. Perez ; Maria Khristina Theresa A. Postrero. "Real-time integrated CCTV using face and

pedestrian detection image processing algorithm for automatic traffic light transitions.

[2] Boris A.Alpotav, Paval B.Babayan, Maskim D.Ershov " Embedded Image Processing and Video Analysis in Intelligent Camera-based Vision System

[3] M. D. Ershov and N. Yu. Shubin, "Road situation analysis algorithms", *Proc. of Digital Signal Processing and its Applications*

[4] R. Sanchez-Iborra, J.F. Ingles-Romero, G. Domenech-Asensi, J.L. Moreno-Cegarra, Maria-Dolores Cano, "Proactive Intelligent System for Optimizing Traffic Signaling", Dependable Autonomic and Secure Computing 14th Intl Conf on Pervasive Intelligence and Computing 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech) 2016 IEEE 14th Intl C.

[5] Jennifer C. Dela Cruz, Ramon G. Garcia, Sarah Mae A. Hernandez, John Melliton B. Medina, Mary Rose, T. Ostique, Angela Mae A. Santos, Marvin Verdadero, "Multiple Face Recognition Surveillance System with Real-Time Alert Notification using 3D Recognition Pattern"

[6] Seungwon Jung, Yongsung Kim, Eenjun Hwang, "Real-time car tracking system based on surveillance videos", EURASIP Journal on Image and Video Processing

[7] P. Babayan, S. Buiko, L. Vdovkin, M. Ershov, V. Muraviev, A. Sirenko, et al., "Real-time pyramidal Lukas-Kanade tracker performance estimation", *Proc. of SPIE Real-Time Image Processing and Deep Learning*

[8] Carlo Pane, Marco Gasparini, Andrea Prati, Giovanni Gualdi, Rita Cucchiara, "A people counting system for business analytics", Advanced Video and Signal Based Surveillance

[9] David Ryan, Simon Denman, Sridha Sridharan, Clinton Fookes, "An evaluation of crowd counting methods features and regression models", Computer Vision and Image Understanding

[10] Z. Shao, J. Cai and Z. Wang, "Smart monitoring cameras driven intelligent processing to big surveillance video data"

- [11] N. Haering, P.L. Venetianer and A. Lipton, "The evolution of video surveillance: an overview", *Machine Vision and Applications*, vol. 19, no. 5–6, pp. 279-90, 2008.
- [12] J. Chung and K. Sohn, "Image-based learning to measure traffic density using a deep convolutional neural network", *IEEE Transactions On Intelligent Transportation Systems*,
- [13] Y. Xu, X. Ouyang, Y. Cheng, S. Yu, L. Ng and C-C Xiong, "Dual-mode vehicle motion pattern learning for high performance road traffic anomaly detection", *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*
- [14] Y-K. Wang, C-T. Fan and J-F. Chen, "Traffic camera anomaly detection", *22nd International Conference on Pattern Recognition*, 2014.
- [15] C. Shih, S.C. Chen, C.F. Hung, K.W. Chen, S.Y. Lin, C.W. Lin, et al., "Real-time camera tampering detection using two-stage scene Matching", *IEEE International Conference on Multimedia and Expo*, pp. 1-6, 2013.
- [16] T. Theodore, L. Christensen, P. Fihl and T.B. Moeslund, "Tamper detection for active surveillance systems", *IEEE International Conference on 10th Advanced Video and Signal Based Surveillance*,
- [17] Z.O. Zhao, P. Zheng, S. Xu and X. Wu, "Object Detection with Deep Learning: A Review", *IEEE Transactions On Neural Networks And Learning Systems*.
- [18] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: unified real-time object detection", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*,
- [19] B. Babenko, M-H. Yang and S. Belongie, "Visual Tracking with Online Multiple Instance Learning", *IEEE Conference on Computer Vision and Pattern Recognition*, June 2009.
- [20] Z Kalal, K. Mikolajczyk and J. Matas, "Tracking-Learning-Detection", *Pattern Analysis and Machine Intelligence*
- [21] P. Dierckx, "Algorithms for smoothing data with periodic and parametric splines", *Computer Graphics and Image Processing*
- [22] W. Hu, T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *IEEE Trans. on Syst., Man and Cyber*,
- [23] A. Mecocci and M. Pannozzo, "A completely autonomous system that learns anomalous movements in advanced video-surveillance applications," *IEEE Int'l Conf. on Image Processing*
- [24] W. F. Abaya et al., "Low cost smart security camera with night vision capability using Raspberry Pi and OpenCV", *IEEE International Conference on Humanoid Nanotechnology Information Technology Communication and Control Environment and Management (HNICEM)*, 2014.
- [25] S. Prasad, P. Mahalakshmi, A. J. C. Sunder and R. Swathi, "Smart surveillance monitoring system using Raspberry Pi and PIR sensor", *International Journal of Computer Science and Information Technologies*, vol. 5, no. 6, pp. 7107-7109, 2011.
- [26] Z. Yin and R. T. Collins, "Object tracking and detection after occlusion via numerical hybrid local and global-mode seeking", *IEEE Conference on Computer Vision and Pattern Recognition*, 2008.
- [27] Vishwanath Bharadwaja, R. Ananmy, Sarraf Nikhil, K. V. Vineetha, Jalpa Shah, Dhanesh G. Kurup, "Implementation of Artificial Neural Network on Raspberry Pi for Signal Processing Applications", *Advances in Computing Communications and Informatics (ICACCI) 2018 International Conference on*, pp. 1488-1491, 2018.
- [28] Janewit Wittayaprapakorn, Thongchai Yooyativong, "Vehicle Seat Vacancy Identification using Image Processing Technique", *Digital Arts Media and Technology (ICDAMT) International Conference on*, pp. 152-155, 2017.
- [29] R. S. Davies, I. D. Wilson, J. A. Ware, "Stereoscopic mobile Natural Feature Tracking", *Computer & Information Technology (GSCIT) 2014 Global Summit on*, pp. 1-5, 2014.
- [30] Sathishkumar Sivagurunathan, Aditya Piratla, Jayalakshmi Surendran, Monotosh Das, "Automatic detection of entry into a restricted area", *Computing Communication and Networking Technologies (ICCCNT) 2014 International Conference on*, pp. 1-5, 2014.
- [31] J. Amat, M. Frigola, A. Casals, "Human robot interaction from visual perception", *Intelligent Robots and Systems 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, vol. 2, pp. 1997-2002 vol.2, 2004.
- [32] Arnaldo Camara Lara, Roberto Jr. Hirata, "Motion Segmentation using Mathematical Morphology", *Computer*

Graphics and Image Processing 2006. SIBGRAPI '06. 19th Brazilian Symposium on, pp. 315-322, 2006.
[33] Thomas B. Moeslund, Erik Granum, "A Survey of Computer Vision-Based Human Motion Capture", Computer Vision and Image Understanding, vol. 81, pp. 231, 2001.

[34] Ross Davies, Ian Wilson, Andrew Ware, "Stereoscopic Human Detection in a Natural Environment", Annals of Emerging Technologies in Computing, vol. 2, pp. 15, 2018.