Passenger Counter Based on Random Forest Regressor Using Drive Recorder and Sensors in Buses

Hayato Nakashima
Graduate School of Information Science
Nara Institute of Science and Technology
Ikoma, Japan
nakashima.hayato.nz4@is.naist.jp

Ismail Arai
Information Initiative Center
Nara Institute of Science and Technology
Ikoma, Japan
ismail@itc.naist.jp

Kazutoshi Fujikawa
Information Initiative Center
Nara Institute of Science and Technology
Ikoma, Japan
fujikawa@itc.naist.jp

***Abstract*—In recent years, some bus companies have raised revenue by reviewing the route plan using the number of passengers. The company has a system that can automatically counts the number of passengers on an ongoing basis. But they are costly because they use cameras and sensors those are dedicated for counting. It is too expensive for bus companies that really need to reconsider their route planning to introduce the system. In order to solve this problem and realize efficient operation, we propose a method to count passengers by using a drive recorder and sensors those are already equipped with buses. Drive recorders and various sensors will be obliged by the government to be set up by bus operators in the future. We constructed a model using Random Forest Regression with the position of the bus from the GPS module in the buses, the position of the bus stop used for operation management, and the number of passengers estimated from the image processing method combining YOLOv3 and Deep SORT. As a result, the average correct answer rate when the passengers get on and off are 96.2% and 70.1% respectively. Our method which utilized non-dedicated camera achieved higher correct answer rate than the conventional method which utilizes dedicated camera for counting passenger.**

***Index Terms*—Machine learning, Image Processing, Public Transportation**

I. Introduction

In recent years, the route bus business has been in a difficult management situation due to the motorization mainly in local cities, the aging population, and the progress of depopulation in Japan. The management of such deficit routes has been compensated by the surplus of the other routes. However, due to the amendment of the Road Transport Law which relax a new bus companies' barrier to entry, excessive competition causes, then many bus companies are unable to maintain the deficit routes. Therefore, the demand for bus companies to know the number of passengers for reviewing the route plan is increasing. Eagle Bus Corporation in Saitama Prefecture Japan has installed sensors to actually count passengers and succeeded in improving the route plan such as the relocating of bus stops using these data [1]. Bus companies have a demand to utilize data on the number of passengers at each bus stop. Also, there is a constant demand for services that allow customers to know the crowdedness on the bus which is scheduled for boarding. In the bus location system provided by Kobe Traffic Advancement Corporation in Hyogo Prefecture Japan, the smartphone application provides crowdedness on the bus to the users [2].

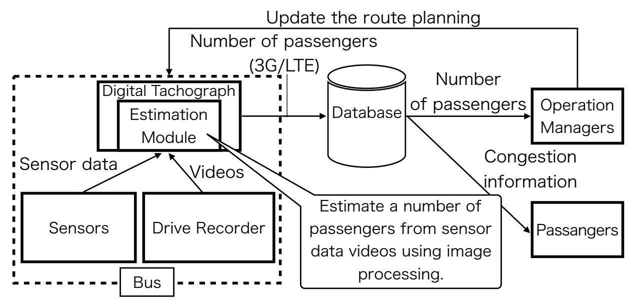


Fig. 1: Use of the passenger counter and the proposed method

The traditional method for counting passengers is manually counting by investigators. This method has two problems. Firstly, it is costly for hiring investigators for several days and entering into a spreadsheet software. The other is less information because a bus company cannot hire investigators every day. Generally, only one or two days of data can be collected a year. Since living traffic such as visiting a hospital and shopping is not done steadily every day, it is impossible to accurately grasp the actual condition of passengers. Some bus companies introduce passenger counting system using bus cards transaction, sensors, or GPS [1]. However, most of the bus companies which depend on some subside from the government cannot install the above systems because it is not subject to subsides. Therefore, it is necessary to construct a system that informs the operation manager of the number of passengers on the bus in real time with a low cost that can be introduced by a bus company at the community bus level.

On the other hand, in recent years, various devices have been installed in buses. After the Karuizawa ski bus fall accident that occurred in 2016, the Ministry of Land, Infrastructure, Transport and Tourism Japan announced the "recording and preservation of video by drive recorder" as "strengthening safety measures on the hardware side" in order to secure the safety of the transportation of the charter bus [3] by obligating "guidance and supervision using the video from the drive recorder". The government is promoting the installation of a drive recorder with subsidies and it will spread in the future [4]. In many cases, it is equipped with an external camera used for investigation and analysis of general traffic accidents, and an inward camera used for studying troubles inside the car. Furthermore, the government is promoting the installation of a digital tachograph with subsidies and it will spread in the future [4]. The digital tachograph is a device that automatically records the speed, time, position, and etc. of the vehicle on a memory card or a cloud. By analyzing them, it is possible to easily check whether the driver complies with legal speed, break time, and etc. Recently, a system capable of automatically counting the number of passengers continuously by attaching a dedicated camera or sensor to the ceiling has appeared. But they are costly because they use cameras and sensors those are dedicated for counting. In this paper, we constructed a model using Random Forest Regression with the position of the bus from the GPS module in the bus, the position of the bus stop used for operation management, and the number of passengers estimated from the image processing method combining YOLOv3 and Deep SORT. As a result, the average correct answer rate for only image processing (YOLOv3 + Deep SORT) are 93.5% (when the passengers gets on) and 36.1%(when the passengers gets off). When applying correction using sensor data to YOLOv 3 + Deep SORT are 96.2%(when the passengers gets on) and 70.8%(when the passengers gets off). In the estimation at the time of riding, even if the image of the drive recorder whose installation purpose is different is used, it can be estimated with higher accuracy than the existing method by correction of the proposed method.

II. Related work and issues

A. Related work

Eagle Bus actually installed sensors to count passengers and succeeded in improving the route plan by relocating the bus stops according to the data such as the number of passengers getting on/off at each bus stop. Traditionally, Eagle Bus has been operated on the basis of experience and intuition of the driver, and formulation of a route plan using data has not been done so much. Therefore, the Eagle Bus installed an infrared ride sensor inside the bus, and grasped the number of passengers at each bus stop, the number of passengers (crowdedness) between bus stops, the position on the line, and the time it took to operate. From such data, the company graphed the route section where is no passenger, and tried to optimize the operation schedule. The unit for judging the cost from the data acquired by the infrared sensor has been changed from "one unit" or "one diagram (traveling)" to "one minute" or "one kilometer" respectively. As a result, even within a single diagram, "which section is in deficit" can be highlighted.

The bus location system of Kobe Traffic Advancement Corporation provides the crowdedness in the bus [2]. By using the smartphone application, the user can know the crowdedness of the bus scheduled to board into three stages of "Crowded", "Normal", or "Vacant".

B. Issues

Currently, as a method of automatically counting passengers who get on/off a bus, a method to count with high precision by installing a dedicated camera on the ceiling of an entrance gate has been proposed [5]. In this paper, the captured frame is firstly divided into many blocks and each block will be classified according to its motion vector. If the block quantity of similar motion vectors is more than a threshold, those blocks are regarded as the same moving object. Then it is counted as getting on/off the bus according to its direction. Experimental results show that the algorithm can provide a count accuracy of 92% on average.

There is also an effort to estimate the number of passengers in an indoor passage by infrared sensors [6]. Experiments using an infrared sensor installed on the ceiling show that it is possible to measure the number of passengers traveling with an accuracy of 95%.

However, these methods are costly because they use cameras and sensors those are dedicated for counting. Therefore, it is difficult to introduce it in a bus business where the business situation is severe. In this research, the number of passengers at each bus stop is estimated by using sensor data of existing drive recorder and digital tachograph. Therefore, the cost for estimating the number of passengers can be reduced.

III. Proposed Method

We constructed a model to count the passengers using Random Forest Regression with the position of the bus from the GPS module in the buses, the position of the bus stop used for operation management, and the number of passengers estimated from image processing. The drive recorder mounted on the bus generally employs four cameras. Two of these cameras shoot inside the bus, one shoots a driver, and another shoots passengers(Fig. 2). We first estimate the number of passengers near the front door by processing the image of the inward camera that was installed. By using the method, the numbers of passengers getting on and off for each bus stop are estimated. After that, the proposed method corrects the image processing output by using Random Forest Regression with the sensor data. As the equipment used in the method has already been introduced, automatic counting can be carried out with a reasonable cost. The details of the method are as follows.

A. Image Processing method

In the past, we proposed a counting method combining a background subtraction method using video from a drive recorder(Fig. 3) and a Random Forest Regressor using sensor data. However, the drive recorder handled in this paper can not use that method. Compared with the video in this paper, this

Fig. 2: The sample video of drive recorder

Fig. 3: The sample video of drive recorder of the last paper

Fig. 4: Example of execution with Background Subtraction

video has vibrations from the bus engine, and also the angle of view is different(Fig. 4). In the last paper, we counted using passenger's sideways movement, but in the video of this paper, there are few sideways movements and there are many moves of depth. Occlusion tends to occur in the movement of the depth, and it is difficult to correctly detect each passengers by using the background subtraction method.

We also tried detecting human beings using OpenPose(Fig. 5). However, since OpenPose presumes parts of a body,

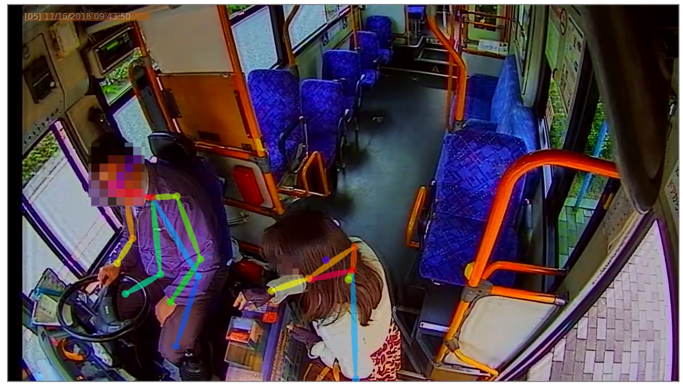


Fig. 5: Example of execution with OpenPose

Fig. 6: The sample frame with YOLOv3+Deep SORT

counting precision was not good because it can not be detected as a human when each part of a body can not be estimated well. Especially during boarding, due to the position of the camera, it often shoots the back of the passengers, so the detection accuracy is not high.

1) Detecting the human using YOLOv3: We use YOLOv3, a method for object detection. The method regards object detection as a regression problem to spatially separated bounding boxes and associated class probabilities. It uses a single neural network, and the network predicts bounding boxes and class probabilities directly from full images in one evaluation. Because it solves as a simple regression problem, processing speed is fast. In other algorithms, we tried to detect a candidate region of an object using techniques such as ”sliding window” and ”Region Proposal”, so we often erroneously detect the background as an object. In YOLOv 3, such erroneous detection is half of ”Fast R-CNN” [8] [9]. According to the detected object, it is possible to obtain the size and the position of the rectangular outline, the type of the object (person, car, food, etc.) and the score of the judged object. The score indicates the accuracy of the judged object in the range of 0 to 1. For these reasons, we used YOLOv3 with pre-trained model to detect humans. When the score of human beings is over 0.2, this method detects humans as the blue box in Fig. 6.

2) Tracking the human using Deep SORT: The proposed method tracks humans detected by YOLOv3 with Deep SORT.

The method overcomes degradation of tracking accuracy due to occlusion by replacing the association metric with a more informed metric that combines motion and appearance information. In particular, the method applies a convolutional neural network (CNN) that has been trained to distinguish pedestrians on a large-scale person re-identification dataset [10]. Therefore, Deep SORT, which is strong against occlusion, is superior to Simple On-line Real-time Tracking (SORT) [11] to track passengers traveling continuously in the bus. Because, SORT and Deep SORT are simpler to process than the Multiple Hypothesis Tracking (MHT) [12] and the Joint Probabilistic Data Association Filter (JPDAF) [13], it is suitable to processing in the bus. The system tracks passengers by inputting human information detected by YOLOv3 into Deep SORT. An example at that time is shown in the white box in Fig. 6. After that, the outline is approximated the tracking number(green character) and the center of gravity(red point).

3) Counting: The proposed method set the blue line as the boundary between the entrance area and the passenger area as shown in Fig. 6. Passengers' getting on/off are determined by whether they cut across this line. The optimal position of the line depends on the vehicle and the angle of view, so that it can be specified arbitrarily. Next, the red points of the last frame and the current frame judged by the III-A2 are connected by a line segment. The method counts when the line segment intersects the border line. At this time, the getting on/off is determined based on whether the center of the gravity point of the current frame is on the upper or lower side of the border line. Finally, the number of passengers who get on/off the bus are output.

B. Correction method using sensor data

When estimating the number of passengers only by processing images of a drive recorder, a large error is caused by occlusion and passengers' staying in the vicinity of the border line. Therefore, we propose a correction method using sensor data obtained from digital tachograph. We define three features obtained when the bus stops at the bus stop. Table I shows the features used to construct the model. The number of passengers estimated by image processing using a drive recorder is defined as a feature. Since there is a correlation between the number of passengers and the time when the bus stops at the bus stop, the time is defined as a feature. The time is calculated from the position coordinates of the bus stop in the database used for navigation management and the position coordinates transmitted from the bus vehicle. We use the time when the distance between two coordinates is 50 m or less and the speed of the bus is 0 km/h. In this study, in order to separately estimate the number of passengers who get on/off the bus, we created each model using Random Forest Regressor. When estimating the number of passengers who get on, also use "the number of passengers (getting off)" estimated by image processing. This is because there is a correlation between the number of passengers who get on / off when the number of passengers staying in the vicinity of the boundary becomes double count.

TABLE I: Features used for construction of model

Feature	Data source
Stopping time at the bus stop	Speed, GPS, DB
# of passengers (getting on)	Image processing
# of passengers (getting off)	Image processing

TABLE II: Data sets of model

Bus stop ID	Stopping times	Total time [second]	Total people (Getting on)	Total people (Getting off)
0	46	2348	132	17
1	15	1786	20	0
2	55	3662	380	54
3	60	2611	217	78
4	18	301	0	17
5	0	0	0	0
6	12	722	20	3
7	12	301	37	0
8	138	13908	32	23
Total	356	4064	838	192

TABLE III: Feature Importances

Feature	Getting on	Getting off
Stopping time at the bus stop	0.02329249	0.22383565
# of passengers (getting on)	0.9665134	0.01019412
# of passengers (getting off)	0.06729245	0.7088719

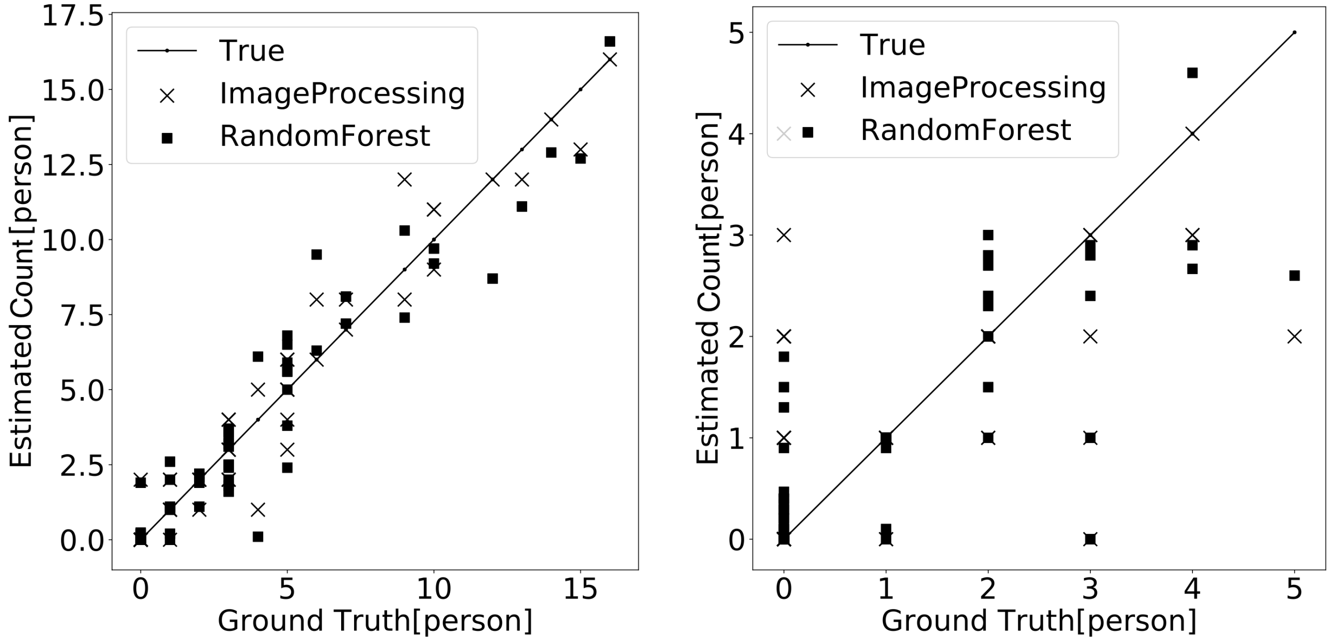
IV. Experiments and discussions

A. Experiments environment

The data set used for the evaluation of this research are the in-vehicle videos and sensor data obtained for four days from one bus of Minato Tourist Bus Corporation in Hyogo Prefecture Japan. The time of the getting the data are as follows.

- Friday, November 16th, 2018
  - 07:06:09 - 14:09:57
  - 16:07:34 - 23:03:18
- Saturday, November 17th, 2018
  - 07:05:24 - 14:01:51
  - 16:04:07 - 21:32:51
- Sunday, November 18th, 2018
  - 07:06:55 - 14:03:49
  - 16:06:10 - 21:26:29
- Monday, November 19th, 2018
  - 07:04:04 - 14:03:30
  - 18:04:16 - 23:01:11
- Tuesday, November 20th, 2018
  - 07:03:26 - 15:59:06

Based on these data, we extract 356 videos(68 seconds in average) and sensor data at each bus stop. We prepare the correct label by manually observing the video. Table II shows the ID of the bus stop, the number of times of the bus



(a) Getting on

(b) Getting off

Fig. 7: Comparison of estimation results

stopped at each bus stop, the total time, and the total number of passengers' getting on/off. On this route, passengers get on from the front door and can get off from any front and rear doors. 70% of these data randomly extracted are used as learning data and the remaining 30% are used as test data to evaluate the accuracy.

B. Experimental results

We estimated the number of passengers on the front door by using the data set and proposed method and evaluated. The evaluation results of passengers who get on/off are shown in Fig. 7(a) and Fig.7(b) respectively. Table III shows feature importances.

The accuracy of ”getting on” which performed image processing only is 93.5%. The accuracy of the estimated value corrected by Random Forest Regressor is 96.2%. Comparing the result using Random Forest Regressor with the result of only image processing, it can be seen that the correction of Random Forest Regressor is effective. In the feature importances, ”# of passengers (getting on)” is the highest. The second is ”# of passengers (getting off)”. Since the two estimated values of getting on/off are correlated, these features are important.

The accuracy of ”getting off” which performed image processing only is 36.1%. The accuracy of the estimated value corrected by Random Forest Regressor is 70.8%. In the feature importances, ”# of passengers (getting off)” is the highest. The second is ”Stopping time at the bus stop”. When getting off, there is a correlation between the number of people getting off and the time. In the case of estimation using only image processing, the estimated value of passengers who got off are large. The method misjudged the passengers who are staying near the boarder line as the passengers who get off. Therefore, they can be corrected by features such as estimated value of passengers by image processing and the times when opening the door. In addition, when comparing the passenger estimation when getting on and off, the accuracy of getting off is low. This is because the number of passengers who get off from the front door is small. On the route used in the data set of this paper, passengers get on from the only front door, but they can get off from either the front or the back door when getting off the bus. In this paper, we have not done image processing to estimate passengers getting off the rear door yet, the absolute value of passengers who got off is small, and one person's error is large. Therefore, it is necessary to estimate the getting on and off from the back door in future. As for the rear door, a camera like that shown in the Fig.8 is installed and it is possible to use it. Once that is possible, we can add the feature which is the number of getting on and off from the rear door to the Random Forest Regressor.

V. Conclusion

In this research, we proposed a passenger counting method using a drive recorder and a digital tachograph. Since these devices are installed under legal obligation, it is possible to estimate passengers without additional cost.

We use YOLOv3, a method for object detection. Compared to the background subtraction used in the past, we were able to

Fig. 8: The sample frame of rear camera

get the result that it is hard to receive restrictions on vibration and field angle. Thus, we track the passengers detected by YOLOv3 with Deep SORT that can track with less occlusion. After this, we constructed a model using Random Forest Regression with the position of the bus from the GPS module in the buses, the position of the bus stop used for operation management, and the number of passengers estimated from the image processing. As a result, the average correct answer rate (getting on) when using image processing is 93.5%, and the average correct answer rate corrected by Random Forest Regressor is 96.2%. As a result of correction using sensor data, the accuracy of the estimation by image processing is improved by 2.7% in the proposed method.

From this result, we confirmed that by combining existing equipment we can count passengers with higher accuracy and affordable price than existing methods using counting-only cameras and sensors.

VI. Future work

For future work, we would like to estimate using the rear camera. Later, we are planning to do the following.

The method of this paper counts the number of passengers only by checking whether the passenger crosses the boundary line. In order to prevent multiple count by passengers staying in the vicinity of the boundary line, it should be counted not only by the boundary line but also by the area rectangle. In this paper, we aim to find out the deficit section to improve the route planning, so we don't need accurate count at congestion. However, there are a few demands that require an accurate counter during congestion, so it is also necessary to consider how often such problems occur and how they will be improved by the method using the area rectangle.

Also, in an environment experimented with this method, a lot of machine power is required. In future it is necessary to consider to operate on a single board computer used in digital tachograph. In recent years, lightweight models that can work even on them have been proposed [14]. It can be operated on a digital tachograph by using the model, analyzing with lower resolution and frame rate, and so on.

References

[1] Nikkei BigData, "Eagle Bus regenerates deficit route by utilizing sensor, improves balance in 1 minute and 1 km unit (In Japanese)" 2014.
Available: http://business.nikkeibp.co.jp/article/bigdata/20140217/259878 (accessed 2018-11-22)

[2] Kobe City, "Japan's first! Bass location system corresponding to congestion situation in bus started to be provided (In Japanese)" 2016.
Available: http://www.city.kobe.lg.jp/information/press/2016/04/2016042 8142001.html (accessed 2018-11-22)

[3] Ministry of Land, Infrastructure, Transport and Tourism (MLIT), "About countermeasures that received Karuizawa ski bus accident (In Japanese)" 2016.
Available: http://www.mlit.go.jp/jidosha/jidosha_tk2_000050.html (accessed 2018-11-22)

[4] Ministry of Land, Infrastructure, Transport and Tourism (MLIT), "Support for advanced navigation management (In Japanese)" 2016.
Available: http://www.mlit.go.jp/jidosha/anzen/subcontents/jikoboushi1.html (accessed 2018-11-22)

[5] Chao-Ho Chen, Yin-Chan Chang and Tsong-Yi Chen., "People Counting System for Getting In/Out of a Bus Based on Video Processing." *ISDA '08. Eighth International Conference*,565–569, 2008.

[6] Gwang-Gook Lee, Hyeong-ki Kim and Ja-Young Yoon., "Pedestrian Counting Using an IR Line Laser" *Convergence and Hybrid Information Technology, 2008*,482-485,2 008.

[7] Hayato Nakashima, Ismail Arai, Kazutoshi Fujikawa., "Proposal of a Method for Estimating the Number of Passengers with Using Drive Recorder and Sensors Equipped in Buses" *2018 IEEE International Conference on Big Data, pp.5379–5381, Dec, 2018*

[8] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi., "You Only Look Once: Unified, Real-Time Object Detection" *Computer Vision and Pattern Recognition (cs.CV).*

[9] Joseph Redmon, Ali Farhadi., "YOLOv3: An Incremental Improvement" *Computer Vision and Pattern Recognition (cs.CV).*

[10] Nicolai Wojke, Alex Bewley, Dietrich Paulus., "Simple Online and Realtime Tracking with a Deep Association Metric" *Computer Vision and Pattern Recognition (cs.CV).*

[11] Alex Bewley, Zongyuan Ge, Lionel Ott, Fabio Ramos, Ben Upcroft., "Simple Online and Realtime Tracking" *2016 IEEE International Conference on Image Processing, 2016, pp.3464–3468, 2016.*

[12] Chanho Kim, Fuxin Li, Arridhana Ciptadi, ames M. Rehg., "Multiple Hypothesis Tracking Revisited" *2015 IEEE International Conference on Computer Vision, pp. 4696-4704, 2015.*

[13] Joseph Redmon, Ali Farhadi., "Seyed Hamid Rezatofighi, Anton Milan, Zhen Zhang, Qinfeng Shi, Anthony Dick, Ian Reid" *2015 IEEE International Conference on Computer Vision, pp. 3047-3055, 2015.*

[14] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, Liang-Chieh Chen., "MobileNetV2: Inverted Residuals and Linear Bottlenecks" *Computer Vision and Pattern Recognition (cs.CV).*