Artificial Intelligence based Camera Calibration

Syed Navid Raza
Department of Electrical Engineering
Yeungnam University
Gyeongsan, South Korea
rizvi.navid@hotmail.com

Hafiz Raza ur Rehman
Department of Information and Communication Engineering
Yeungnam University
Gyeongsan, South Korea
mrazaurrehman@ynu.ac.kr

Suk Gyu Lee
Department of Electrical Engineering
Yeungnam University
Gyeongsan, South Korea
sglee@ynu.ac.kr

Gyu Sang Choi
Department of Information and Communication Engineering
Yeungnam University
Gyeongsan, South Korea
castchoi@ynu.ac.kr

***Abstract*— Camera calibration technique plays a vital role in three dimensional computer vision systems. The aim of this technique is to calibrate the camera in order to collect more precise three dimensional data from the images to utilize for robot navigation, three dimensional reconstruction, biomedical, virtual reality and visual surveillance. In camera calibration one of the major issue is to search out the set of image parameters describing the mapping between three dimensional images reference coordinates and two-dimensional images reference coordinates. Currently, MATLAB toolbox and OpenCv are the most popular tools used by the researchers for camera calibration. We utilize the concept of deep learning to recognize the chessboard corners. Our presented technique is a convolutional neural network (CNN) trained on a huge number of chessboard images. The network is trained on different datasets: noisy and with high lens malformation images. The proposed scheme is more accurate than the conventional MATLAB algorithm technique. The presented technique is more accurate against the different sort of ruination present in the training set. Results reaffirmed the correctness and effectiveness of our proposed CNN technique.**

Keywords—convolutional neural network, camera calibration, checkboard detection

I. Introduction

The design tenacity of three dimensional(3-D) computer vision system is to achieve image statistics from the camera. It is for the assessment of location, shape information of 3-D object and thus to become aware of object in the space. The brightness of image at each point, presents the intensity of reflected light from that point on the surface of 3-D object. The geometric vicinity of this point in image links with the geometric vicinity of space object's surface. The geometric model for camera imaging determines the association among these locations. The geometric parameters of the model are referred as the camera parameters, these parameters must be determined by experiment and calculation. The process of experiment and calculation is called as camera calibration.

There are multiple techniques of camera calibration. These technique are categorized into three classes [1]: traditional calibration technique, self-calibration and active vision. Traditional techniques [2] are portrayal as direct linear transformation (DLT), two-step technique and dual-plane technique. If, there is no calibration object is present then it brings a type of calibration which is called as self-calibration [3]. Currently, Active Vision camera self-calibration techniques are used in taking account the non-linear distortion of camera. There are many other techniques for the recognition of checkboard features and for creating checkboard from them. Early approaches used general corner detector like Harris [4], SUSAN (Smallest Univalve Segment Assimilating Nucleus) [5] and Moravec[6] corners. Due to specific composition of the checkboard pattern, it is viable to do a prior filtering for the sections of image liable to have a checkboard [7]. Another complex coalition of general image corners is presented by Placht et al. with ROCHADE (Robust Checkboard Advanced Detection) [8]. The centerlines of Scharr-filtering of the input image are determined and utilized to assess the saddle-points which are detected features. After subpixel refinement, the checkboards are build.

There are numerous types of calibration techniques, like calibration method based on three dimensional calibration reference proposed by Tsai [9], in which calibration of reference object is made up of the two planar pattern which are at right angles to each other, three dimensional space coordinate of the reference object surface's calibration point should be known before the calibration. Consequently, this technique needs a costly precision calibration system. Zhang [10] contributes a calibration technique depends on 2-D pattern, Zhang calibration technique merely needs shooting the same calibration pattern for more than two times from different angles. Hence, it can obtain the intrinsic and extrinsic camera parameters. This technique doesn't want to recognize the particular position and displacement of planar pattern movement. Thus, planar pattern is easy to create and in contrast is simple and easier than other techniques.

Wang also puts forward another technique of camera calibration, using OpenCV. OpenCV is open computer vision library, it consists of a sequence of C functions and a small amount of C++ classes, and contains many general algorithm about image processing and computer vision. Therefore, it possess powerful capability to image and matrix operations. This library uses Zhang's calibration method in camera calibration. The camera calibration module of OpenCV offers a good interface for user, and supports Windows and Linux platform. It improves the development efficiency, the execution speed, and has good cross-platform portability [11]. Another checkboard recognition technique is an algorithm by Vezhnevets [12]. It works on the basis of black quadrangles recognition in the image and coalite them into checkboard. This approach was enhanced in OCamcalib [13] with a well checkboard formation algorithm and pre-processing to deal with noisy images.

With the advancement in technologies, people think neural network will be more reliable for resolving the nonlinear models. The FAST (Features from Accelerated Segment Test) [14] image corner detector is based on deep learning(DL), as its descendant, FAST-ER (FAST-Enhanced Repeatability) [15]. DL is very effective in in segmentation of different type of Images like MRI images [16] and electron-microscopy images [17]. The well trained layers of convolutional neural network(CNN) are very fruitful for pixel-wise processing. DL is very helpful to get the features from the images as well as from the hand written digits.

Memon and khan [18] have done the camera calibration that is focused on stereo setup. The major flaw in their system is that, their system have to retrain for every new setup. Jun and Kim [19] proposed the similar technique but that is not confined to stereo setups. They proposed two level perceptron: one for the center portion of the image and the other for the boundary area of image.

MATLAB Calibration Toolbox provides a diversity of calibration exercises and several old calibration techniques. Similarly, it offers calibration goal, user interface with appropriate and stretchy, and the toolbox C source code in the computer vision library. With the aid of calibration tool in the MATLAB toolbox, the camera can be calibrated. The calibration tool also can be used to calibrate the camera parameters rapidly, efficaciously, and optimize the intrinsic parameters [20]. In calibration, there is analysis of the intrinsic and extrinsic parameters [21]. The consequence of rise in images and distortion is very difficult to describe in such a non-open environment. In this paper, we proposed the camera calibration by writing an algorithm in MATLAB and CNN to analyze the statistics of mean square error(MSE) of pixels per image. We also studied effect of such factors such as rise of images and distortion in calibration process.

This paper is organized as follows: Section 2 describes the generic architecture of CNN, modified version of basic CNN system and datasets used for experiment. Section3 based on experiments and obtained results, the effect of different dataset on MSE is discussed. The last section concludes this paper.

II. Calibration with Convolutional Neural Network:

A. *Convolutional Network Topology:*

Convolutional neural network is a mluti-layer system. It consists of multiple layers like convolutional layers, maxpooling and Flatten layer. The generic model of a convolutional neural network is, as shown in Fig. 1.

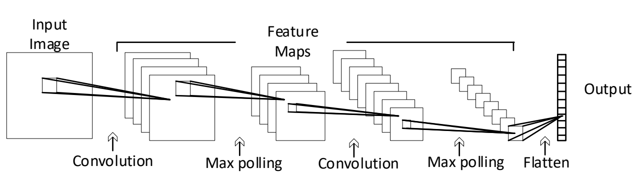


Fig. 1. Generic CNN model [24].

We propose a convolutional neural network for recognition of checkboard corners and to get the mean square error per image. The proposed network is comprises of four convolutional layers, two dense layers and one flatten layer as shown in Figure. 2. First layers extract the features from the pixels, while the descendant layers combine these feature into a chessboard corner weight as shown in Fig. 2.

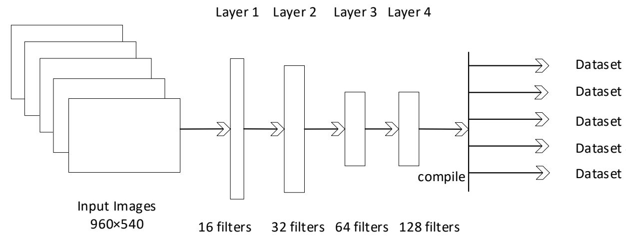


Fig. 2. The proposed CNN system

The first layer of the network have a convolutional filter with the large number of neurons. The activation function used for first layer is ReLU (Rectified Linear Unit) [25]. The spatial size of the filter is large to avoid the blur effect of the images. We can write general recursive equation of convolutional neural network in which input image is X, filter kernels$W_{1,n,m}$, biasb_n, m is the number of neurons and i and n are the numbers as shown in Equation(1).

$$L_{1,n}(X)(x,y)=\max\left(\sum_{m=1}^{m}\left(W_{1,n,m}\times L_{n-1}(X)\right)(x,y)+b_n,0\right) \quad (1)$$

There are 16, 32, 64 and 128 neurons. Stride, maxpooling is 2 and 3*3 convolutions in four convolutional layers respectively. The activation function in each layer are ReLUs.

B. *Datasets*

The network is trained on Adam (Adaptive moment Estimation). It is an extension to stochastic gradient descent and having the properties of two extensions of stochastic gradient descent: AdaGrad (Adaptive Gradient Algorithm) and RMSProp (Root Mean Square Propagation). It is very fast in convergence and used widely for optimization. The datasets for training comprises of two parts: images as captured by phone (using Huawei honor 8 lite) and distorted versions of these images. The dataset $D_{captured}$ contains total 80 images which covered the whole chessboard as shown in Fig.3.

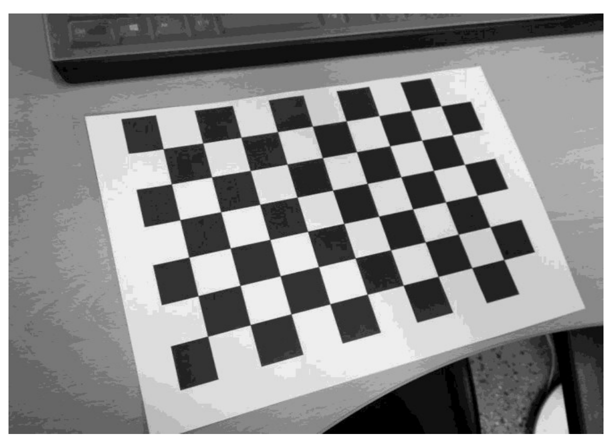


Fig. 3. Captured data set image

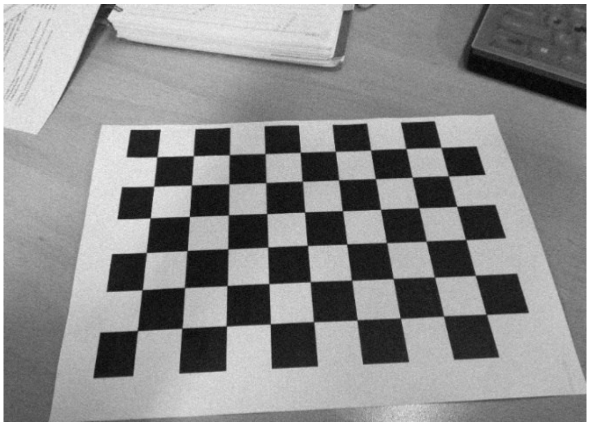


Fig. 4. Noise dataset image

It is resized to the resolution (960×540). The other dataset $D_{rotated}$ consists of images which are rotated [23] at (90-180-270) - degree of the input images, half of images have inverted intensities (black become white and vice versa). We increased dataset by adding Gaussian noise in the captured image and made a new datasetD_{noise}. The noise is uniformly distributed over all the image having standard deviation of 0.1. The noisy dataset is shown in Fig. 4.

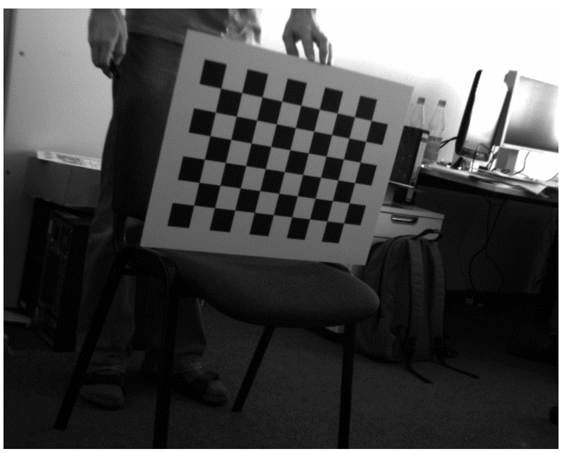


Fig. 5. ueye dataset image [8]

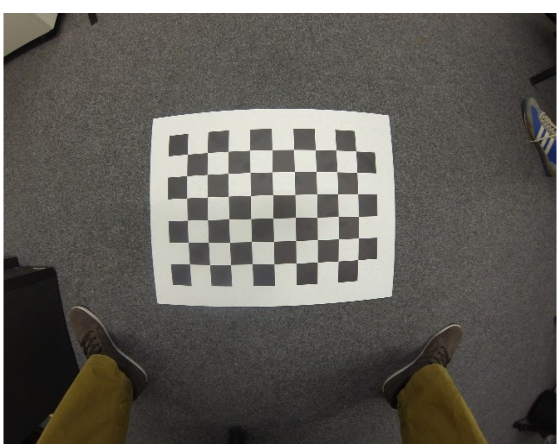


Fig. 6. GoPro dataset image [8]

We used two data training datasets from [8]. The ueye dataset is captured by two IDS UI-1241LE cameras (Imaging Development Systems, Obersulm, Germany). The resolution of images is 1280×1024 as shown in Fig.5.The other dataset is GoPro dataset which has a very high resolution (4000×3000) with a fine quality: as shown in Fig. 6.

III. Experiments and Results

We perform comparison of corners detection with the MATLAB algorithm and CNN algorithm in python. We assess both methods on our training sets as well as the dataset from [8], ueye and GoPro. Results for the all training sets are given in the Table 1 and 2.

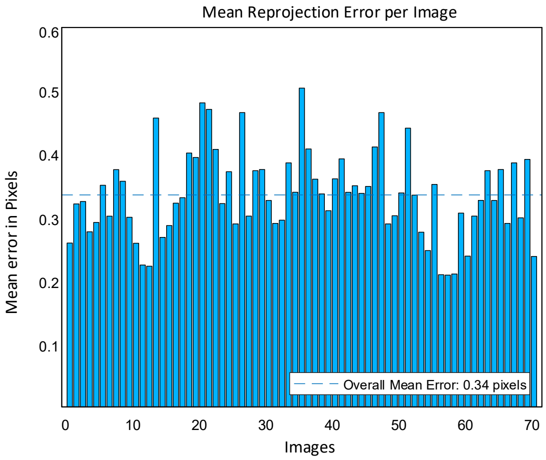
Table 1. Results of Datasets sets with MATLAB algorithm

Datasets	$D_{captured}$	$D_{rotated}$	D_{noise}	ueye	Gopro
MSE	0.34	2.09	0.35	0.45	0.30
Time()	195	190	28	8600	540

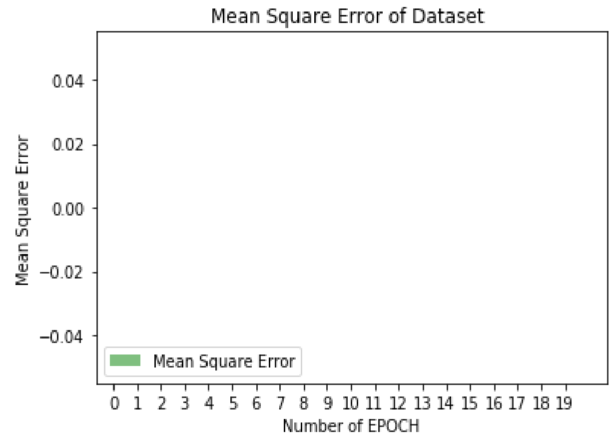
Table 2. Results of Data sets with CNN algorithm

Datasets	$D_{captured}$	$D_{rotated}$	D_{noise}	ueye	Gopro
MSE	0.000016	0.0024	0.0413	0.0063	0.000184
Time(ms)	200	260	250	300	500
Accuracy	1.0	0.9960	0.959	0.985	1.0

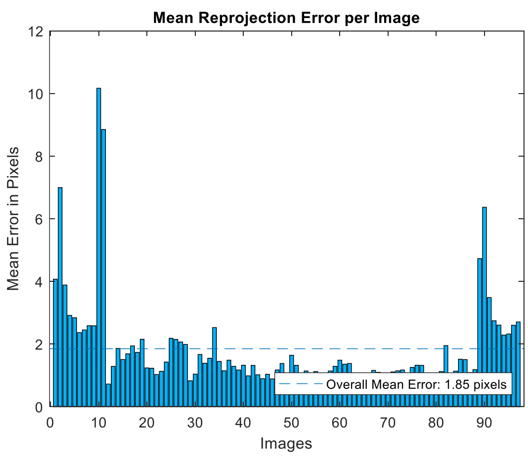
We trained the network on different datasets. The number of epochs are 20 and the batch size is 10 for every dataset. We trained every dataset separately through the CNN. The difference of performance between both the algorithms is shown in Fig.7. For captured dataset, MATLAB algorithm recognized 71/80 checkboards. It is notable in Fig. 7(a) that the MSE for $D_{captured}$ is 0.34 pixels per image in MATLAB. On the contrary, CNN recognized all 80 checkboards. There is no error in any epoch in CNN as shown Fig. 7(b). The reason is that, all images are noiseless in this dataset. Therefore, we can say MSE is zero in CNN for noiseless data. For $D_{rotated}$ dataset recognized 97/240 checkboards in MATLAB. It shows more error than the simple captured images as shown in Fig. 7(c). This dataset has rotated and inverted images as well as data set is three times more than the captured images dataset. Thus, this dataset has more noise. However, same dataset in CNN is shown in Fig. 7(d). It detects 240/240 checkboards.It has some error at the first epoch and goes down to zero on the next epochs. After first epoch, CNN used processed data. Therefore, MSE goes to its minimum value. The dataset D_{noise} describes the MSE is 0.35 pixels per image as shown Fig. 7(e), in MATLAB. The dataset shows the same behavior in CNN like the previous datasets as shown in Fig. 7(f). The checkboard detection remained 71/80 in MATLAB and 80/80 in CNN. The comparison of MSE per image for dataset ueye with MATLAB and CNN is shown in the Fig. 7(g) and Fig. 7(h) respectively. The checkboard detection is 205/206 in MATLAB and 200/206 in CNN. These images conflict each other because images were from two different views. Hence, it gives some error. When we apply the both algorithms on GoPro dataset. Both techniques recognized the 99/100 and 100/100 checkboards. It gives 0.45 MSE per image in MATLAB and in CNN it approaches to zero after first epoch as shown in Fig. 7(i) and in Fig.7 (j) respectively. If we combine all datasets and pass through our proposed MATLAB algorithm. It increased the MSE to a higher value. Hence, it is preferred to pass the data set separately through the algorithm. When we pass all datasets jointly through the CNN system. It shows high value of MSE. Its minimum value will reach to 0.3. It is due to conflicting behavior of all the datasets. Similarly, in this case accuracy reduced to approximately 0.6.



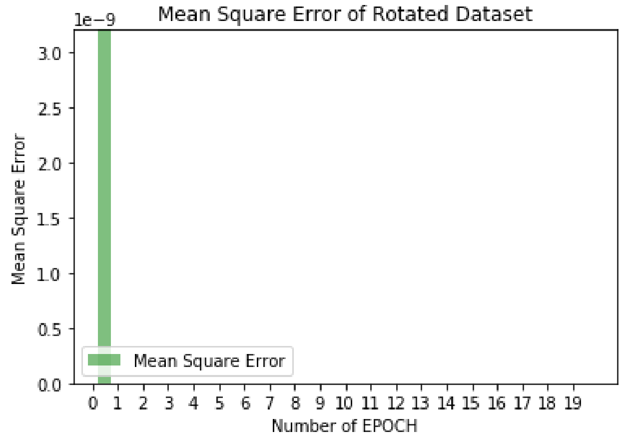
(a)



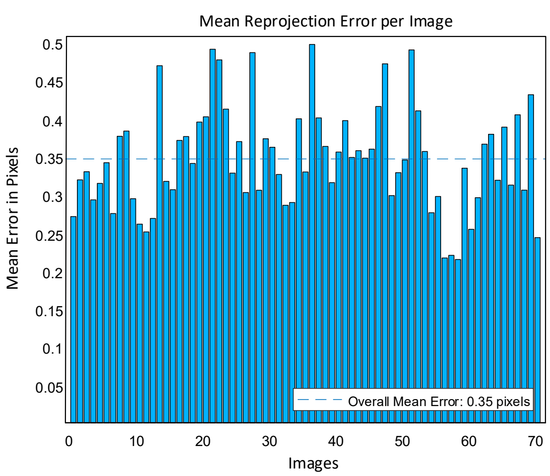
(b)



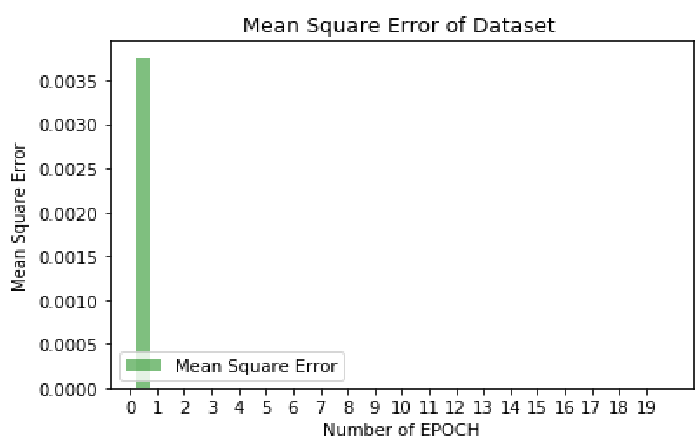
(c)



(d)



(e)



(f)

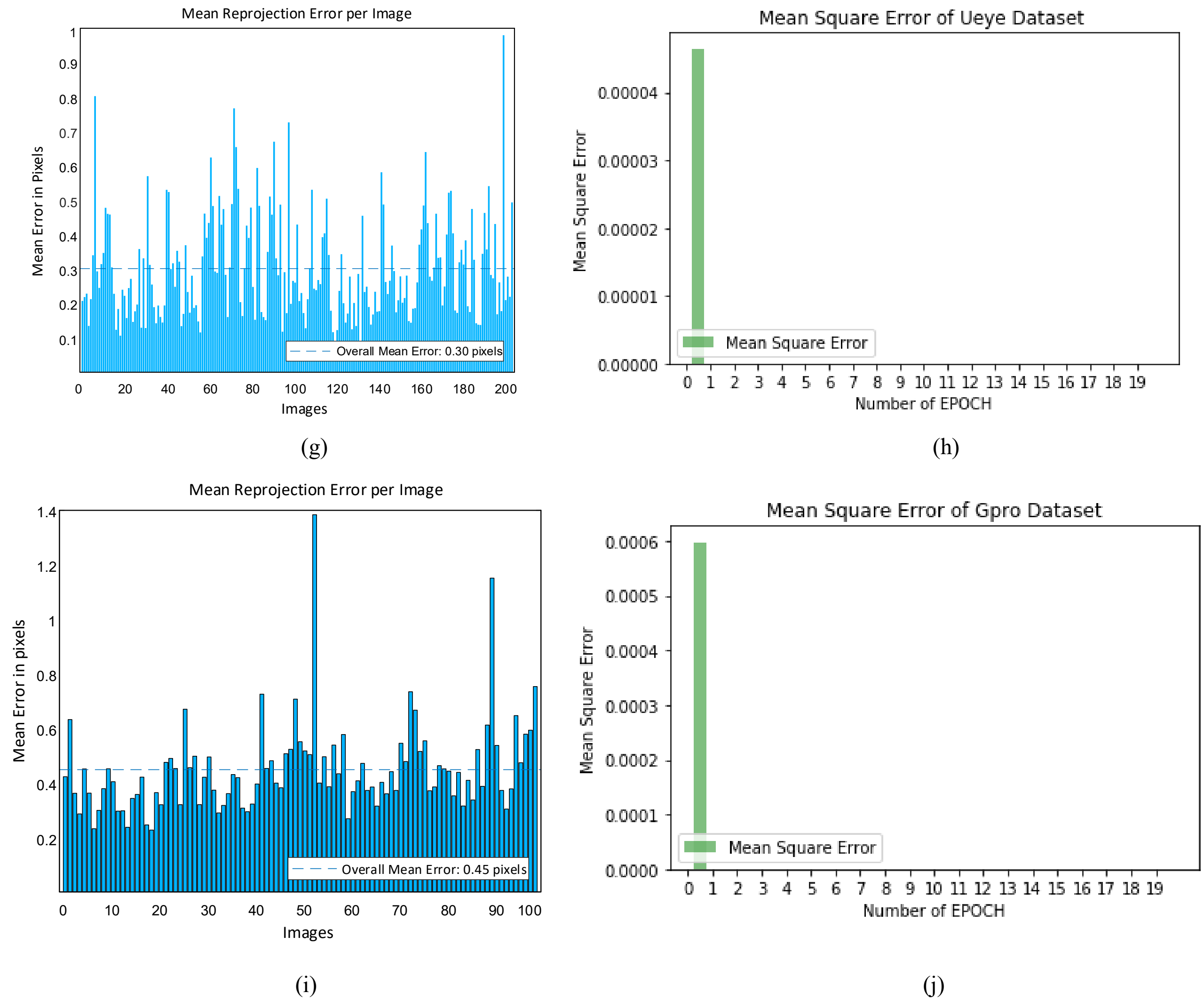


Fig.7. It shows the comparative results of different datasets after applying the both algorithm.

IV. CONCLUSION

Inthis paper, we introduced a convolutional neural technique to solve the issue of camera calibration. The CNN system is trained with real images. We described how the different errors of camera calibration can be remitted and minimized using this technique. Additionally, this technique can be applied on different types of cameras. It is able to detect checkboard corners under the worse noise situations. It is notable that this technique showed same behavior under the enormous noise situations. This approach refined all the training data and decreased the error to zero. The neural network can be trained easily to adopt new situations and practices. As future work, we plan to apply this technique in autonomous multi robots. Henceforth, robots can detect objects for their path planning and exploration.

ACKNOWLEDGMENT

This research was supported by the Ministry of Trade, Industry & Energy (MOTIE, Korea) under Industrial Technology Innovation Program. No.10063130, Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2019R1A2C1006159), and MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2019-2016-0-00313) supervised by the IITP(Institute for Information & communications Technology Promotion).

REFERENCES

[1] Qi, Wang, Fu Li, and Liu Zhenzhong. "Review on camera calibration." Control and Decision Conference (CCDC), 2010 Chinese. IEEE, 2010.

[2] Hu, Zhan-yi, and Fu-chao Wu. "A review on some active vision based camera calibration techniques." CHINESE JOURNAL OF COMPUTERS-CHINESE EDITION- 25.11 (2002): 1149-1156.

[3] HU, Pei-cheng, Ning Li, and Jian-jiang ZHOU. "Improved camera self-calibration method based on circular points." Opto-Electronic Engineering 34.12 (2007): 54-60.

[4] Harris, Chris, and Mike Stephens. "A combined corner and edge detector." Alvey vision conference. Vol. 15. No. 50. 1988.

[5] Smith, Stephen M., and J. Michael Brady. "SUSAN—a new approach to low level image processing." International journal of computer vision 23.1 (1997): 45-78.

[6] Moravec, Hans P. "Techniques towards automatic visual obstacle avoidance." (1977).

[7] Arca, Stefano, Elena Casiraghi, and Gabriele Lombardi. "Corner localization in chessboards for camera calibration." Proceedings of International Conference on Multimedia, Image Processing and Computer Vision (IADAT-micv2005). 2005.

[8] Placht, Simon, et al. "Rochade: Robust checkerboard advanced detection for camera calibration." European Conference on Computer Vision. Springer, Cham, 2014.

[9] Tsai, Roger. "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses." IEEE Journal on Robotics and Automation 3.4 (1987): 323-344.

[10] Zhang, Zhengyou. "A flexible new technique for camera calibration." IEEE Transactions on pattern analysis and machine intelligence 22 (2000).

[11] Wang, Y. M., Y. Li, and J. B. Zheng. "A camera calibration technique based on OpenCV." Information Sciences and Interaction Sciences (ICIS), 2010 3rd International Conference on. IEEE, 2010.

[12] Vezhnevets, V. "OpenCV calibration object detection." Part of the Free Open-Source OpenCV Image Processing Library.[(accessed on 20 December 2016)] (2005).

[13] Scaramuzza, D., and R. Siegwart. "Ocamcalib toolbox: Omnidirectional camera calibration toolbox for matlab." Google for "ocamcalib". (2006).

[14] Cai, Sheng, Qing-an Li, and Yan-feng Qiao. "Camera calibration of attitude measurement system based on BP neural network." Journal of Optoelectronics Laser 18.7 (2007): 832.

[15] Fetić, Azra, Davor Jurić, and Dinko Osmanković. "The procedure of a camera calibration using Camera Calibration Toolbox for MATLAB." MIPRO, 2012 Proceedings of the 35th International Convention. IEEE, 2012.

[16] Rosten, Edward, and Tom Drummond. "Machine learning for high-speed corner detection." European conference on computer vision. Springer, Berlin, Heidelberg, 2006.

[17] Rosten, Edward, Reid Porter, and Tom Drummond. "Faster and better: A machine learning approach to corner detection." IEEE transactions on pattern analysis and machine intelligence 32.1 (2010): 105-119.

[18] Powell, Stephanie, et al. "Registration and machine learning-based automated segmentation of subcortical and cerebellar brain structures." Neuroimage 39.1 (2008): 238-247.

[19] Roels, Joris, et al. "Automated membrane detection in electron microscopy using convolutional neural networks." Belgian-Dutch Conference on Machine Learning. 2016.

[20] Memon, Qurban, and Sohaib Khan. "Camera calibration and three-dimensional world reconstruction of stereo-vision using neural networks." International Journal of Systems Science32.9 (2001): 1155-1159.

[21] Jun, Junghee, and Choongwon Kim. "Robust camera calibration using neural network." TENCON 99. Proceedings of the IEEE Region 10 Conference. Vol. 1. IEEE, 1999.

[22] da Silva, Suane Pires P., et al. "Localization of Mobile Robots with Topological Maps and Classification with Reject Option using Convolutional Neural Networks in Omnidirectional Images." 2018 International Joint Conference on Neural Networks (IJCNN). IEEE, 2018.

[23] Nair, Vinod, and Geoffrey E. Hinton. "Rectified linear units improve restricted boltzmann machines." Proceedings of the 27th international conference on machine learning (ICML-10). 2010.

[24] Donné, Simon, et al. "MATE: Machine learning for adaptive calibration template detection." Sensors 16.11 (2016): 1858..