VOL. 12, NO. 16, AUGUST 2017 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

# AN AUTOMATIC HYBRID APPROACH TO DETECT CONCEALED WEAPONS USING DEEP LEARNING

Suhad Al-Shoukry
Al-Furat Al-Awsat Technical University, Iraq
E-Mail: suhadalzhra2010@yahoo.com

**ABSTRACT**

Detecting concealed weapon underneath a person's clothing is important for public safety in places such as airports. The success rate of our proposed approach in detecting concealed weapons based on fused images is higher than that of the bag-of-features approach. Our approach uses a traditional discrete wavelet transform with hybrid bag-of-words to obtain fused imagery. We then utilized convolutional neural network (CNN) with a pre-trained CNN model using the CNN features of the fused image to train a multiclass SVM classifier. Our approach works well with X-ray images. The experimental results indicate the efficiency of the proposed hybrid approach.

**Keywords:** convolutional neural network, bag-of-words, SVM classifier.

## INTRODUCTION

Security has become a critical issue as a result of the terrorist attacks worldwide (1). Airports, buses, and train stations should employ methods for detecting weapons, such as guns, knives, and chemical explosives concealed beneath clothing. This issue facilitated the emergence of security as a business enterprise (2). Security professionals around the world are paid to watch video streams from cameras to detect criminal activities. Enormous savings per year will be achieved if computer vision will be employed to take over a fraction of the billions of hours spent on these endeavors. Computer vision can detect incidents that could go unnoticed by human viewers. Most surveillance systems are equipped with multiple cameras that can easily track, detect, and recognize objects(3). The rest of this study is describes the types of detection utilized in concealed weapon detection (CWD).

## RELATED WORK

Weapons, such as knives and guns, have flat surfaces that could scatter the beam (4). Thus, these objects can be detected if the illuminated surface is perpendicular to the line of sight of the radar. Millimeter wave imaging, which has a frequency range of 30-300 GHz, is used in scanning systems that are currently used to screen people in sensitive areas such as airports (5).

The current trend in CWD is the development of high-frequency devices with terahertz (>1 THz) imaging; these devices offer enhanced image quality and facilitate the detection and discrimination of concealed objects at stand-off ranges higher than the range of millimeter wave imaging(6). The location of the transmission image is the total of X. Existing image sensor technologies for CWD applications include thermal/infrared (IR), millimeter wave (MMW), and visual sensors. Image fusion has been identified as a key solution to improve CWD procedures because single-sensor technologies cannot provide acceptable performance for CWD applications (7). Image fusion is a process of combining complementary information from multiple sensor images to generate a single image that provides a more accurate description of the scene than individual images(8). MMW sensors offers several advantages, but the availability of low-cost IR technology promotes the fusion of visual and IR images (9).

Visible light cameras (VLC) employ sensors with the ability to communicate; these sensors are embedded with tiny VLC hardware (10). VLC receivers use photo-sensitive systems or a camera system for light sensing (11). These receivers extract data with the aid of light beams (12). IR cameras, such as long-wavelength IR (LWIR) cameras, generate images based on slight temperature variations (13). The images produced by IR cameras are possibly slow or may vary because of the temperature ranges required by different applications. IR sensors should consider dynamic range and temperature variations to produce optimal results (14). The ever-present threat of terrorism subjects the public to enhanced screening systems that use full body scanners. These scanners have raised concerns over privacy invasion (15). Microwave frequencies can effectively detect concealed objects,; however, these devices raise embarrassment and privacy concerns (16) during projection of 2D magnitude images, such as scanners that use microwave frequencies, to reveal body images in full graphic detail.

Cooperative signal processing techniques have been proposed in combination with computer vision algorithms as a ground-breaking technology for immediate detection of concealed weapons through scanning (17). Extensive research and the ongoing development of imaging devices coupled with multispectral band systems increased the popularity of these devices as surveillance screening systems. The exploration of multispectral bands invariably led to safe imaging scanners that can prevent possible terrorist attacks (18). The rapid digitization of semiconductor technology paved the way for miniature digital imaging sensors with superior performance (19). These imaging sensors can be easily interfaced with modern consumer electronics to capture images that can explore specific electromagnetic spectrum to provide uniform image output (14). The mechanism for obtaining an image involves substantial internal processing. When

VOL. 12, NO. 16, AUGUST 2017 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

light source strikes the sensor, an electrical charge is stored in the sensors, which is then converted to digital information using digital signal processing techniques. These techniques include quantization and filtering or interpolation algorithms. Encoded data are subsequently stored in memory and further decoded to generate an analogue image signal (20).

Thermal infrared imaging is commonly used in military night vision, target detection, and surveillance applications (21) as shown in Figure-1.

**Figure-1.** Example of thermal infrared imaging.

## PROPOSED APPROACH

Training uses bag-of-features (BOF) and convolutional neural network for non-weapon and weapon pictures as dataset. Figure-2 shownhybrid approach.

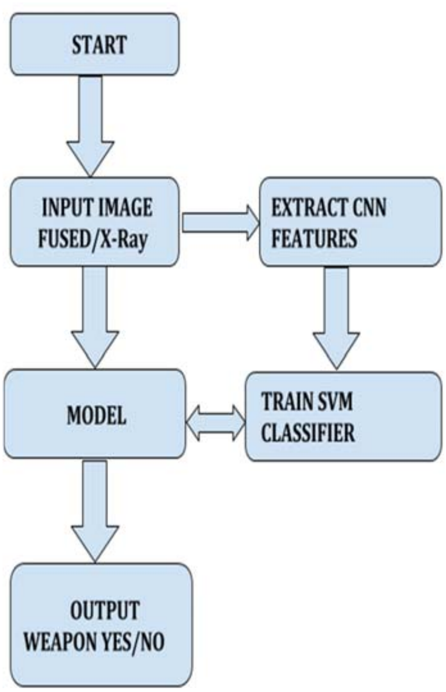


**Figure-2.** Hybrid approach.

### A. BOF approach

Bag of words (BoW) is a technique obtained from the world of natural language processing and adapted to computer vision (22). Figure-3 shows the BoW process.

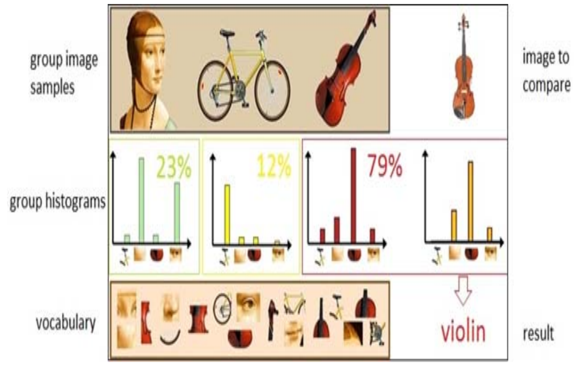


**Figure-3.** BoW process.

### B. CNNs

CNNs are several layers of convolutions with nonlinear activation functions, such as ReLU or tanh. These functions are applied to the results. Each input neuron in traditional feed forward neural network is connected to the output neuron in the next layer(23). This method is called a fully connected layer or affine layer. CNNsdo not have this feature. Instead, convolutions are used over the input layer to compute the output. This condition results in local connections, wherein each region of the input is connected to a neuron in the output(24). Each layer applies different filters, typically in hundreds or thousands, and their results are combined.

We propose an approach of fully leveraging existing image-fusion techniques to create fused images. These techniques can provide several classified concealed images. These fused images are then fed into the CNN to extract features that will be used to train a multi-class support vector machine [16]. Figures 4 and 5 explain our approach. The BOF approach is demonstrated as follows:

**Step 1:**

This step involves reading the input image from the given path.

**Step 2:**

In this step, we create the dataset and label the dataset according to their folder. For example, the dataset will divided into two classes, namely, weapon and nonweapon. We divide our dataset to 70% training and 30% testing.

**Step 3:**

Several internal procedures are run in this step. One of these procedures is the extraction of SURF features, where k means clustering is used to create BOF.

**Step 4:**

In this step, the features from Step 3 are used to train a SVM classifier.

VOL. 12, NO. 16, AUGUST 2017 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

**Step 5:**

Five performance evaluation metrics are run based on the result of test set on a model built by the SVM classifier.

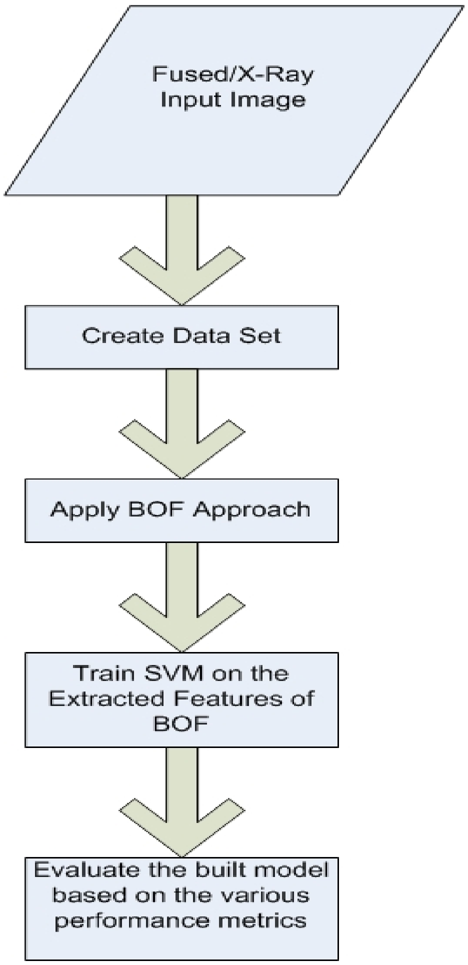


**Figure-4.** BoW through SVM classifier.

The proposed hybrid approach using CNN shows the entire whole CNN, which is composed of 23 layers. However, only 19 layers are considered in extraction.

**Step 1:**

This step involves reading the input image from a given path.

**Step 2:**

In this step, we create and label the dataset according to their folder. For example, the dataset is divided into two classes, namely, weapon and non-weapon. We divide the dataset to 70% training and 30% testing.

**Step 3:**

CNN is run in this step, but its output layer is not considered. We will consider the output of the FC7 layer, which actually contains the extracted features.

**Step 4:**

In this step, the features from Step 3 are used to train a SVM classifier.

**Step 5:**

Five performance evaluation metrics are run based on the result of the test set on the model built by CNN.

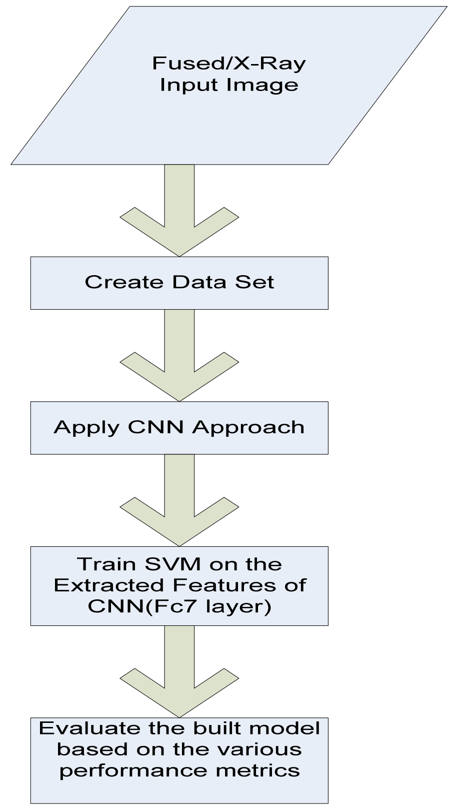


**Figure-5.** CNN through SVM classifier.

## SIMULATION ENVIRONMENT

Several evaluations were conducted to assess the proposed approach. These approaches are precision, recall, and F1-score for each class, and the mean F1 score for overall classes. These methods are summarized as follows:

a) **Accuracy:** This feature refers to normal percentage.

b) **Precision:** Low precision indicates large number of false positives.

c) **Recall:** Low Recall indicates large number of false negatives. In F1Scores, F1 conveys the balance between precision and recallMeanF1 pertains to the mean of every successful CNN, especially when increased CNN data work well.

Table-1 shows the simulation parameters.

VOL. 12, NO. 16, AUGUST 2017 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences



**Table-1.** Simulation parameters.

| Parameter | Value |
|---|---|
| Platform | Windows 10 |
| Ram | 16 GB |
| GPU | YES (NVDIA) |
| Language | Matlab |
| Tuned Model | Yes. Used Alexnet pre-trained model in the feature extraction |
| Data Size | 3000–5000 |
| Methods | Bag-of-Features Approach and Proposed Hybrid Approach |
| Duration | 48 hours total to train CNN |
| Accuracy | Mean of Confusion Matrix |

## RESULTS AND DISCUSSIONS

Performance analysis was conducted after simulating the proposed method. Obtaining a labeled dataset of fused images is difficult because of security reasons. Thus, we obtained images from publicly available domains. A quality dataset for X-ray images was crucial in testing our approach. X-ray image data from security agencies, such as TSA, are not readily available to the public; thus, we obtained a large dataset of gun and non-gun X-ray images from researchers working on the same problem [23]. Mery *et al*. [24] detected dangerous objects from multiple views and created a large dataset of X-ray images from multiple angles. We were given permission to use this dataset. We conducted experiments on several datasets, which are discussed separately in the next section.

Figure-6 shows that CNN produces higher results than BoW in terms of accuracy vs. data size. Accuracy varies based on the approach. However, the proposed CNN-based hybrid approach performs better when data size is increased.

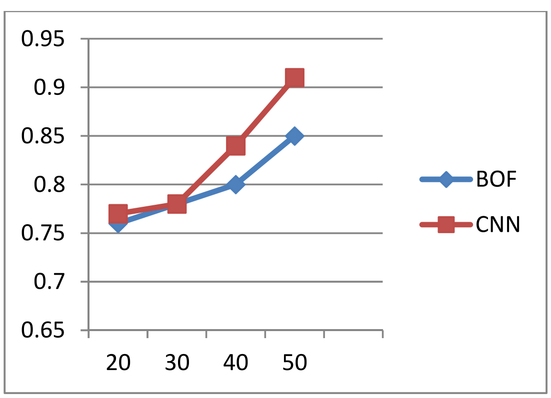


**Figure-6.** Accuracy vs. data size.

Figures7 and show8 that the precision of CNN approach for weapon and non-weapon classes is higher than that of BOF. This finding entails low false positives.

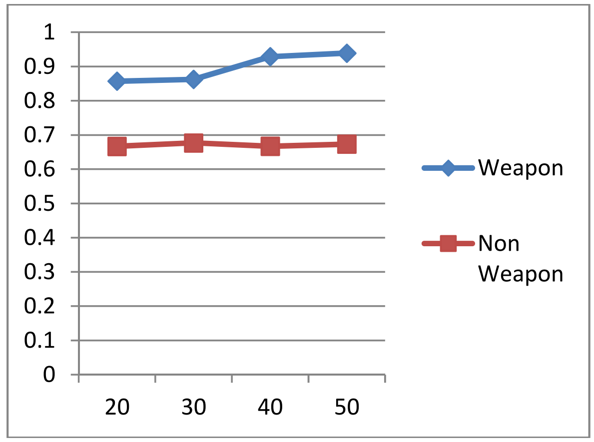


**Figure-7.** Precision for weapon and non-weapons using BOF approach

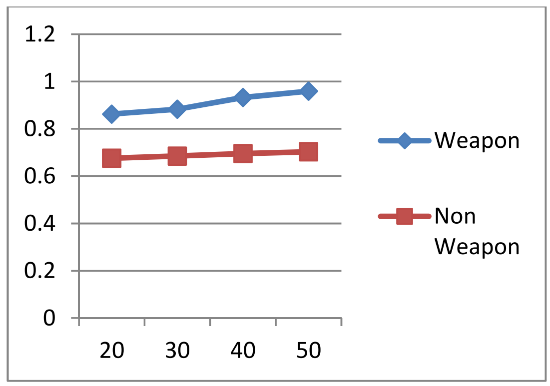


**Figure-8.** Precision for weapon and non-weapons using CNN approach.

Figures 9 and 10 show that the recall of CNN for weapon and non-weapon classes is higher than that of the BOF approach. This finding suggests few false negative results.

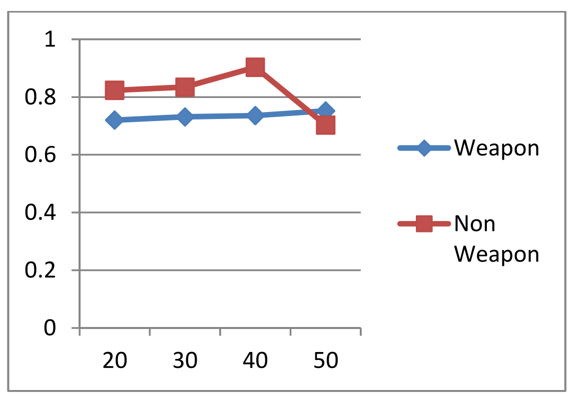


**Figure-9.** Recall for weapon and non-weapons using BOF approach.

Figures 11 and 12 show that the F1Scores of CNN approach for weapon and non-weapon classes are

VOL. 12, NO. 16, AUGUST 2017 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences



more balanced than that of the BOF approach. Hence, we choose the CNN approach.

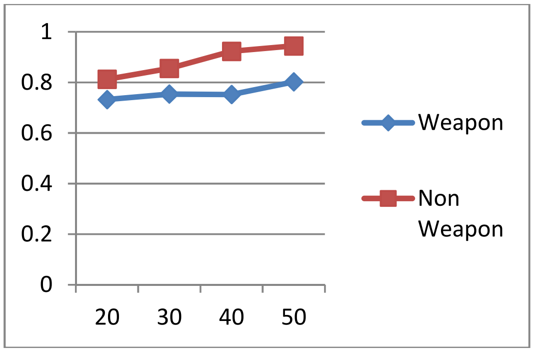


**Figure-10.** Recall for weapon and non-weapons using CNN approach.

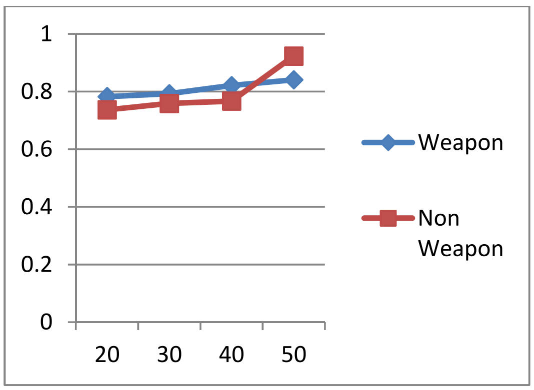


**Figure-11.** F1-scores for weapon and non-weapons using BOF approach.

Figures 12 and 13 show that the score of CNN is better than that of the BOF approach if we take mean F1 Scores.

We use five evaluation metrics to assess the performance of the hybrid CNN approach. This approach shows that CNN performed better than BOF. Performance improves with dataset increase. This finding is theoretically feasible based on the concept of CNN.

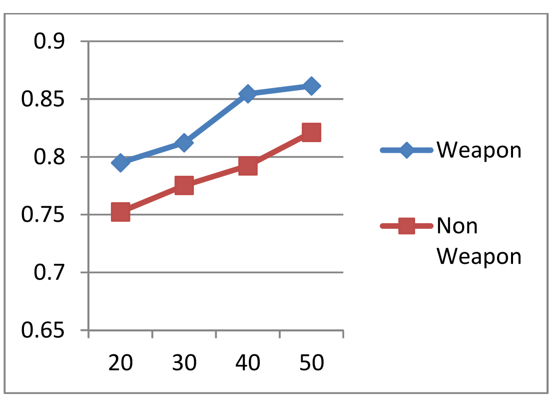


**Figure-12.** F1-scores for weapon and non-weapons using CNN approach.

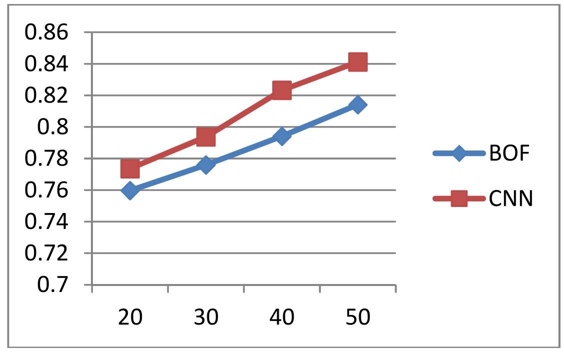


**Figure-13.** Mean F1-scores for weapon and non-weapons.

## CONCLUSIONS

We proposed an automatic hybrid approach that outperforms the current BOF approach. The results are shown in our accuracy graph. The use of additional images during increased the accuracy of model outputs. This hybrid approach was used to detect concealed weapons and offers a new perspective for using robust deep learning set in CWD application.

## REFERENCES

[1] Dlamini MT, Eloff JH, Eloff MM. 2009. Information security: The moving target. Computers & Security. 28(3): 189-98.

[2] Marchese LE, Terroux M, Dufour D, Bolduc M, Chevalier C, Généreux F, *et al.* 2014. Editors. Case study of concealed weapons detection at stand-off distances using a compact, large field-of-view THz camera. SPIE Defense+ Security; 2014: International Society for Optics and Photonics.

[3] Ishigaki T, Yokota S, Li X, Takahashi S. 2016. Image processing apparatus, solid object detection method, solid object detection program, and moving object control system. US Patent 20, 160, 019, 429.

[4] Pratihar P, Yadav AK. 2014. Detection techniques for human safety from concealed weapon and harmful EDS. International Review of Applied Eng Research. 4(1): 71-6.

[5] Harmer SW, Cole SE, Bowring NJ, Rezgui ND, Andrews D. 2012. On body concealed weapon detection using a phased antenna array. Progress in Electromagnetics Research. 124: 187-210.

[6] Cooper KB, Dengler RJ, Llombart N, Thomas B, Chattopadhyay G, Siegel PH. 2011. THz imaging radar for standoff personnel screening. IEEE Transactions on Terahertz Science and Technology. 1(1): 169-82.

[7] Gooch RW, Black SH, Kocian TA, Kennedy AM, Diep BQ. 2016. Wafer level packaged infrared (IR) focal plane array (FPA) with evanescent wave coupling. Google Patents.

[8] Ancuti CO, Ancuti C, De Vleeschouwer C, Bovik AC. 2017. Single-Scale Fusion: An Effective Approach to Merging Images. IEEE Transactions on Image Processing. 26(1): 65-78.

[9] Hervieu A, Le Bris A, Mallet C. 2016. Fusion of Hyperspectral and Vhr Multispectral Image Classifications in Urban Areas. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences. 457-64.

[10] Grobe L, Paraskevopoulos A, Hilt J, Schulz D, Lassak F, Hartlieb F, *et al.* 2013. High-speed visible light communication systems. IEEE Communications Magazine. 51(12): 60-6.

[11] Haruyama S, Yamazato T. 2015. 9 Image sensors based visible light communication. Visible Light Communication. 181.

[12] Tapia G, Elwany A. 2014. A review on process monitoring and control in metal-based additive manufacturing. Journal of Manufacturing Science and Engineering. 136(6): 060801.

[13] Meola C, Boccardi S, Carlomagno GM. 2015. Measurements of very small temperature variations with LWIR QWIP infrared camera. Infrared Physics & Technology. 72: 195-203.

[14] Russ JC. 2016. The image processing handbook: CRC press.

[15] Dillon TW, Thomas DS. 2015. Airport body scanning: will the American public finally accept? Journal of Transportation Security. 8(1-2): 1-16.

[16] Sheen DM, Fernandes JL, Tedeschi JR, McMakin DL, Jones AM, Lechelt WM, *et al.*, editors. 2013. Wide-bandwidth, wide-beamwidth, high-resolution, millimeter-wave imaging for concealed weapon detection. SPIE Defense, Security, and Sensing; 2013: International Society for Optics and Photonics.

[17] Brown KW, Sar DR, Gallivan JR, Phillips WM. 2014. Infrared concealed object detection enhanced with closed-loop control of illumination by. mmw energy. Google Patents.

[18] Ewing KJ, Sanghera JS. 2016. Extended infrared imaging system. US Patent 20, 160, 061, 666.

[19] Bryen SD. 2015. Technology Security and National Power: Winners and Losers: Transaction Publishers.

[20] Campbell SP, Mobbs P, Adsumilli BC, Chawla S. 2016. Image sensor alignment in a multi-camera system accelerator architecture. US Patent 20, 160, 173, 785.

[21] Ring E. 2013. Beyond human vision: the development and applications of infrared thermal imaging. The Imaging Science Journal.

[22] Wu L, Hoi SC, Yu N. 2010. Semantics-preserving bag-of-words models and applications. IEEE Transactions on Image Processing. 19(7): 1908-20.

[23] Chaudhuri TD, Ghosh I. 2016. Artificial Neural Network and Time Series Modeling Based Approach to Forecasting the Exchange Rate in a Multivariate Framework. arXiv preprint arXiv:160702093.

[24] Truong L, Barik R, Totoni E, Liu H, Markley C, Fox A, *et al*., editors. 2016. Latte: a language, compiler, and runtime for elegant and efficient deep neural networks. Proceedings of the 37th ACM SIGPLAN Conference on Programming Language Design and Implementation; ACM.