Neuromorphic convolutional recurrent neural network for road safety or safety near the road

WOO-SUP HAN [1], IL SONG HAN [2]
[1]ODIGA, London, U.K.
[2]Korea Advanced Institute of Science and Technology, Daejeon, REPUBLIC OF KOREA
[1]phil.han@odiga.co.uk, [2]i.s.han@kaist.ac.kr

Abstract: - Neuromorphic visual processing inspired by the biological vision system of brain offers an intelligent vision system of new machine vision in everyday environment. With the growing interest for detecting moving objects on the road or roadside for enhancing the safety and security, the proposed neuromorphic visual processing was tested on vehicle's blind spot cyclist or ramming vehicle terror attack. The neuromorphic convolutional-recurrent neural network has been proposed to detect target objects and demonstrated successfully, based on the saliency of neuromorphic visual processing without complex optimization of template matching. Neuromorphic features were processed by Autoencoder and simplified Gabor filter, and detected either blind spot cyclist on the bridge or ramming vehicle in CCTV video footage of terror attack. The consistent performance of either Gabor-like filters or the small Autoencoder filters demonstrated the feasibility of real-time and robust neuromorphic vision implemented by the small embedded system.

Key-Words: Neuromorphic visual processing, Deep neural networks, Vision, Cyclist detection, Car ramming detection, Crowd movement detection

1 Introduction

The neuromorphic visual processing (NVP) has begun to offer a viable alternative for robust object recognition method in challenging environmental conditions. Although the computer vision algorithms are effective in their condition of usage, in certain situations they lack the robustness of the human vision.

In this paper, the neuromorphic convolutional recurrent neural network (CRNN) is introduced for the robust abstraction of visual object recognition under limited visibility. It is inspired by the fundamental function of primary visual cortex, and the deep neural networks of multiple processing layers with multiple levels of abstraction [1], [2], [3]. The proposed neuromorphic CRNN was investigated for both the blind spot cyclist detection for road safety and the detection of attacking ramming vehicle for safety near the road.

The neuromorphic CRNN has its basis on how we humans process visual information. Hubel and Wiesel's research on cat's visual cortex have established the concept of orientation selective neuron within the visual cortex. The neuromorphic vision of our earlier convolutional neural network demonstrated robust performance in pedestrian and cyclist detection under all-weather conditions and low light conditions [3]. The *neuromorphic CRNN is enhanced with Down Up* network for evaluating the saliency, which is one of widely adopted ideas in vision [4].

Main body of this paper explains how the neuromorphic visual processing is inspired by the visual processing of primitive feature selectivity. The NVP with the controlled rectifier neuron in Down Up network is able to accomplish accurate locating of target objects for further processing [5].

As a test of neuromorphic CRNN, the proposed algorithm is applied on detecting cyclist in vehicle's blind spot without complicated object matching templates. This process is applied further with the example of attacking ramming vehicle detection.

The feasibility of new NVP is evaluated for applications of mobile camera based road safety and CCTV based roadside security, more advanced from earlier NVP [6].

2 Neuromorphic Visual Processing Inspired by Primary Visual Cortex

The NVP has its basis on how mammals analyze visual information. Result of Hubel and Wiesel's experiment on cat's visual cortex put forward the concept of the simple neuron cell. The neuron cell in the visual cortex responds to the orientation of light stimulus as illustrated in Fig. 1.

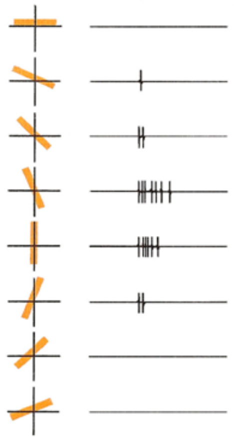

Fig.1. Response of cat's visual cortex neuron when a rectangular slit of light of different orientations is shown [1]

The NVP in Fig. 2 is inspired by a visual cortex neuron's behavior in Fig. 1, with various orientation selective features. The NVP has three key operations based on neural network, which are: 1) orientation feature extraction using convolution by neuromorphic orientation filter, 2) summation of orientation features at each neuron representing pixel, then 3) detection of human head by convolutional neural network with template filter in Fig. 3 and post processing by neural network or classifier.

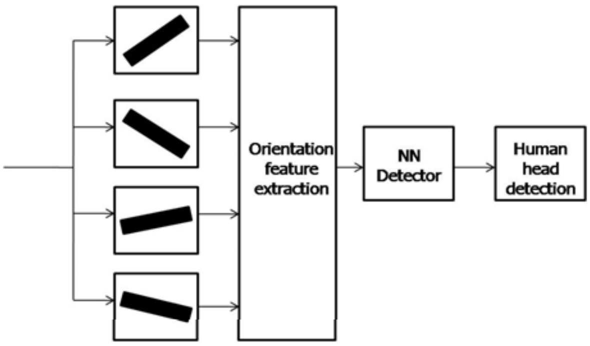

Fig. 2. Neuromorphic visual processing (NVP) for human detection, mimicking the visual cortex neuron

Template filters of NN detector can be tailered by statistcal evaluation or neural network training.

Fig. 3. Human head template configured as a kernel for convolutional neural network,

Fig. 4 shows the early NVP design of deep neural network. It is inspired by a neuromorphic neuron of simple cell in visual cortex, for various orientation selective features. The NVP in Fig. 4 has multiple layers in its process which are: 1) orientation feature extraction by NVP mimicking the brain, with the convolutional neural network (CNN) with the filters (C) and the summation neural network (S), 2) neural networks of template convolution (C) is then applied to the orientation extracted image with the summation (S) yielding the low level saliency, 3) posterior processing using denoising convolution filters (C) with the summation (s) yielding the high level saliency, and finally 4) the object recognition and detection is made by the evaluation layer (N). The image in Fig. 4 shows the detection of human object wearing a poncho in the stormy weather, with a yellow box of detected object.

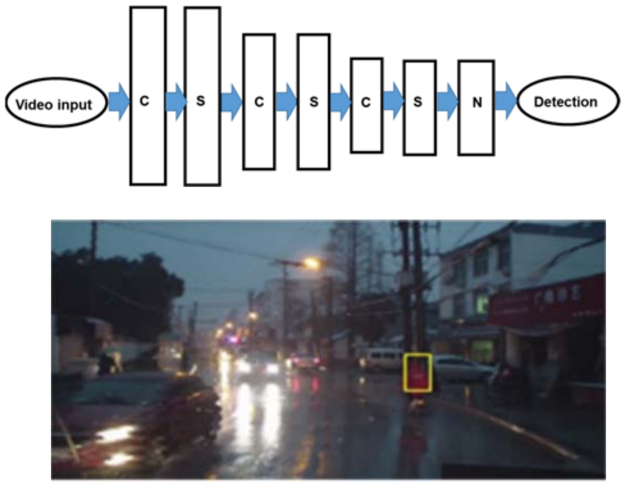

Fig. 4. NVP of Fig. 2, detection based on template kernels and deep convolutional neural network (top diagram), an application to human detection in all weather (bottom image)

3 Neuromorphic Convolutional Recurrent Neural Network for Salience-based Detection

One of the challenges encountered in earlier NVP research is being able to account for the noise from the background or surroundings. The denoising layer of posterior processing in Fig. 4 refines the saliency for improving object detection. It caused the substantial demand of deep neural network for denoising purposes, requiring complicated learning/designing processes of convolution filters as well as extra computational resources. The neuromorphic CRNN in Fig. 5 is proposed to settle those issues arising in noisy visual environment. The feature extraction process is enhanced by cascaded convolution (C) and summation (S), also by added recurrent neural network (RNN). The Down(D) Up(U) network of Fig. 5 is introduced to boost the Saliency evaluation, enhanced by the controlled linear rectifier [5]. Finally, the detection is determined by the evaluation layer (N).

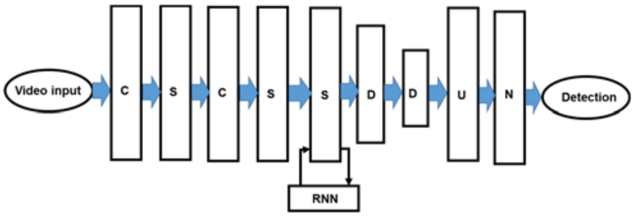

Fig. 5. Neuromorphic convolutional-recurrent neural network (CRNN) for salience-based object detection

Gabor filter is employed as an optional convolutional filter to the hand-cut one, mimicking the orientation feature extracting in visual cortex. Gabor filter in Fig. 6 was implemented by two equation introduced below, and 6 orientations was selected from 16 angles of Gabor filters.

$$g(x,y;\lambda,\theta,\psi,\sigma,\gamma) = \exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda}+\psi\right)$$

$$g(x,y;\lambda,\theta,\psi,\sigma,\gamma) = \exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)\sin\left(2\pi\frac{x'}{\lambda}+\psi\right)$$

The neuromorphic CRNN works on the real part output of complex Gabor filter.

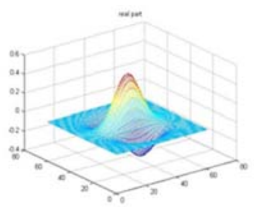

Fig. 6. Gabor complex filter illustrating the orientation feature extraction, in real-part

Autoencoder is provided as one of convolution filter banks as shown in Fig. 5. Autoencoder was trained by anonymous medical image patches. Autoencoder was prepared for the wide coverage of application, particularly for noisy visual environments. The training dataset is based on 11x11 image patches, for maintaining the flexible integrity with another convolution process. The filter parameters from Autoencoder are illustrated in Fig. 7, which appear as pro-orientation selective characteristics.

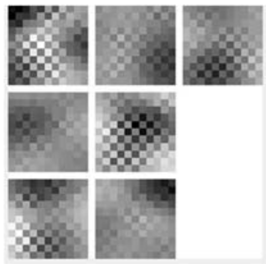

Fig. 7. Neuromorphic convolutional filters from Autoencoder, trained by 11x11 images for 7 hidden nodes

The increasing awareness about road safety with the improved road infrastructure has allowed for greater survival rate in case of road accidents. Yet despite the falling number of fatalities, the number of cyclist fatalities have gone up slightly, with some European nations having quarter of its road fatality from cyclists alone. According to the EU's CARE database, blind spot collision is one of the four most common accidents involving vulnerable road users. Blind spot detection is mostly dependent upon robust and accurate video data analysis, to make detection for approaching cyclists. There are several issues in blind spot cyclist detection, and one of the critical issues would be the large difference in size of cyclist between the nearby and the far-off. It can be influenced by the speed of cycle and would demand far more computational resources as well as more complicated templates, in case of NVP in Fig. 4. The proposed neuromorphic CRNN in Fig. 5 has the intrinsic advantage of finding fast objects even with the moving image sensor, or removing the random noise. It can be equally found as an efficient saliency evaluation compensating the issue of limited visual or computational environment.

The blind spot cyclist detection was evaluated using the video footage of real road traffic, captured via the window of passenger car by custom designed instrument with a car image sensor of 640x480. The test was conducted by both the MATLAB simulation and the Python-based real-time implementation in embedded system. A simple statistical analysis was employed for an evaluation layer N of neuromorphic CRNN, at this stage. The road traffic was recorded during the south bound drive on Tower Bridge in London, in the afternoon of August 2017.

The results in Fig. 8 represent no cyclist in the blind spot range, and no detection of cyclist. The middle figure represents the intermediate outcome after processing of RNN and feature extraction. The results in Fig. 9 show the cyclist detection for the far-off blind spot region, while those in Fig. 10 for the nearby. The final evaluation was applied to the saliency from Down Up processing of the intermediate output of middle image. In this paper, the detection is determined by the processed saliency by hard-limiting neuron with the following activation function.

Neuron output=1 (if pixel saliency >= Threshold)
output=0 (if pixel saliency < Threshold)
where Threshold=0.5* Max(saliency data)

Fig. 8. Neuromorphic CRNN of Fig. 5 for blind spot cyclist detection, (left) input video frame, (middle) intermediate result of RNN processing, (right) no cyclist detection yet

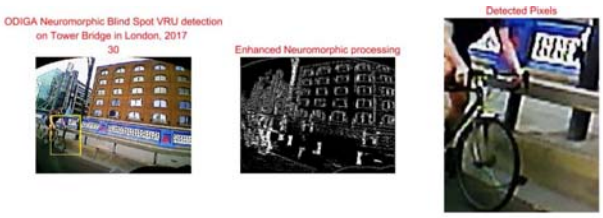

Fig. 9. Neuromorphic CRNN for blind spot cyclist detection, (left) input video frame with yellow box for detected cyclist, (middle) intermediate result of RNN processing, (right)cyclist detection with segmented image

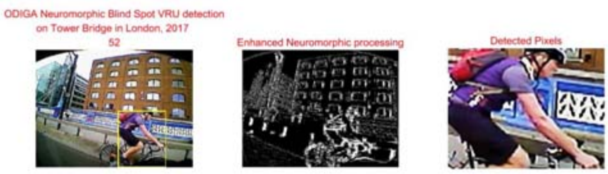

Fig. 10. Neuromorphic CRNN for blind spot cyclist detection, (left) input video frame with yellow box for detected cyclist, (middle) intermediate result of RNN processing, (right) cyclist detection with segmented image

The results in Fig. 11 show the inaccurate location of cyclist for the nearby blind spot region, while it happens rarely and is supposed due to the limitation of simple but hard limiting neuron for the strong landscape noise in the river zone.

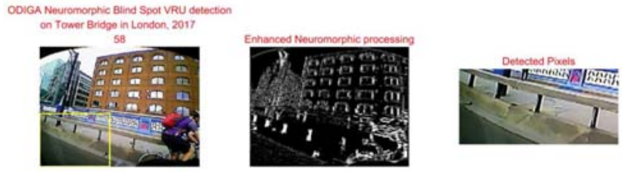

Fig. 11. Neuromorphic CRNN for blind spot cyclist detection, (left) input video frame with yellow box of erroneous detection, (middle) intermediate result of RNN processing, (right) erroneous cyclist detection with segmented image

The results in Fig. 12 show the detection of cyclist over the time during approaching in the blind spot of the vehicle, which was also moving forward. The speed of cycle was faster than the vehicle, and the cyclist or cycle in the blind spot was successfully detected throughout the period. The neuromorphic CRNN operation would be free from the outlook of cyclist or cycle, because being based on the saliency.

Fig. 12. Neuromorphic CRNN for blind spot cyclist detection for 3 moments over 1 second, (left) input video frames with yellow box for detected cyclist, (middle) intermediate result of RNN processing, (right) cyclist detection with segmented image

Recent outbreaks and incidents of vehicular ramming attacks highlight the emergence of new terrorist tactic. To respond to the urgent need in an urban security environment, a swift and simultaneous detection will be required. The neuromorphic CRNN is capable of immediate detection and warning against suspicious and dangerous vehicle manoeuvres. The identical neuromorphic CRNN was evaluated for detection of ramming vehicle terror

Fig. 14. Neuromorphic CRNN for ramming car detection for 0.3 second, based on the saliency evaluation

Fig. 15. Neuromorphic CRNN for monitoring crowd reaction for 5 seconds, after the ramming attack

attack, using the CCTV video footage (320x240) of terror attack in Jerusalem, 2017.

Fig, 14 shows detection of the ramming truck as highlighted in the saliency image – even under the poor visibility. The crowd reaction was monitored reasonably through the saliency in Fig. 15, without any further processing.

4 Conclusion

We have demonstrated the neuromorphic CRNN as more effective analysis method for noisy video data, by evaluating applications of a blind spot cyclist detection or a detection of ramming vehicle attack. Integrating recurrent neural network into NVP gave the way for target object segmentation without requiring supervised learning. The neuromorphic CRNN has shown the 96.4% accuracy for blind spot cyclist detection using the basic design. In addition to the successful detection of ramming vehicle from the low-quality video data, the real-time operation was tested successfully using a GPU (Nvidia TX) embedded system. The neuromorphic CRNN proved its feasibility of a real-time and robust intelligent vision system by the evaluation result of applications and the small scale neural computing resources.
.

Acknowledgement
This research is partly supported by the grant for 'Improving Cyclist Safety via Neuromorphic Visual Processing: Robust Cyclist Detection System' by Department for Transport of United Kingdom.

References:

[1] D. H. Hubel and T. N. Wiesel, Receptive fields of single neurones in the cat's striate cortex, *J. Physiol.*, 148, 1959, pp 574-591.

[2] Y. LeCun et al, "Deep learning," *Nature*, 521, 2015, pp. 436-444.

[3] W. Han and I. S. Han, All Weather Human Detection Using Neuromorphic Visual Processing, *Intelligent Systems for Science and Information*, Springer International Publishing, 2014, pp25-44.

[4] M. K. Lim et al, Detection of salient regions in crowd scenes, *Electronics Letters,* 50, 2014, pp. 1860-1867.

[5] W. Han and I. S. Han, "Bio-Inspired Neuromorphic Visual Processing with Neural Networks for Cyclist Detection in Vehicle's Blind Spot and Segmentation in Medical CT Images," Proc. SAI Computing Conference 2017, pp744-750.

[6] W.J. Han et al, "Bio-inspired visual information processing- the neuromorphic approach," WSEAS Transactions on Circuits and Systems, 9, 2010, pp. 441-452