Neuromorphic luminance-edge contextual preprocessing of naturally obscured targets

Alexander James White
Institute of Systems Neuroscience
National Tsing-Hua University
Hsinchu City, Taiwan
ajw@lolab-nthu.org

Chou P. Hung[†]
DEVCOM Army Research Laboratory
Adelphi, MD, USA
chou.p.hung.civ@army.mil

Andre V. Harrison
DEVCOM Army Research Laboratory
Adelphi, MD, USA
andre.v.harrison2.civ@army.mil

Chung-Chuan Lo[†]
Brain Research Center & Inst of Systems Neurosci
National Tsing-Hua University
Hsinchu City, Taiwan
cclo@mx.nthu.edu.tw

ABSTRACT

Contextual grouping mechanisms in early visual cortex are thought to aid in perception of ambiguous textures, including partially obscured targets under real-world high dynamic range (HDR) luminance. Yet, deep neural networks struggle with naturalistic obscuration and illumination while requiring millions of neurons and power-hungry GPUs for processing. We hypothesized that contextual grouping mechanisms for edge and luminance processing may aid in localization of targets under natural obscuration and illumination. To address this issue, we developed a novel small (< 10,000 neurons) spiking neural network (SNN) that uses spike time correlations to leverage the combined luminance and orientation similarity of nearby image regions for image pre-processing, to support downstream deep neural network (DNN) target localization. The network has leaky integrate-and-fire neurons with current based (CuBa) synapses and is simulated using the Nengo LOIHI API, with potential application via Intel's LOIHI neuromorphic hardware. We collected 89 HDR images of a target dummy in a heavily wooded environment under varying occlusion and illumination. We used SNN preprocessing to adjust local image contrast based on the grouping mechanism, followed by a DNN classifier (Detectron2) to localize the target. Results show that a small SNN for image preprocessing can aid image segmentation and localization of occluded targets, marking an initial step towards more efficient and accurate target recognition under natural illumination and occlusion.

†Corresponding authors

CCS CONCEPTS

• Computing methodologies~Artificial intelligence~Computer vision~Computer vision problems~Image segmentation
•Computing methodologies~Artificial intelligence~Computer vision~Computer vision problems~Object detection
•Computing methodologies~Artificial intelligence~Computer vision~Computer vision problems~Shape inference

KEYWORDS

image segmentation, occlusion, high dynamic range (HDR), grouping, object recognition, shape perception, neuromorphic

1 Introduction

When searching and navigating a visual environment, it is necessary to make sense of areas of high illumination and dark shadows simultaneously, particularly in high dynamic range (HDR) luminance visual scenes with clutter and locally varying illumination. This poses a challenge for cameras due to their

limited dynamic range, as cameras either capture limited light or become oversaturated [1,2,3]. To address this, the industry standard is to blend together three low dynamic range (LDR) images taken at different exposures [3,4,5] However, this process can require offline processing of the video [1], and results are poor even with large, power-hungry convolutional neural networks [4,5]. Moreover, it does not account for natural variations in illumination direction and cast shadows e.g. under a forest canopy. In contrast, biology solves this problem with fewer resources by using divisive normalization to dynamically compensate for HDR light levels [2], along with contextual grouping mechanisms to improve our ability to disentangle illumination to perceive lightness and shape, as demonstrated in animals [3, 4] and in humans [5]. Proper normalization in HDR environments is crucial for downstream computations, such as target recognition and localization. The problem is compounded by obscuration in a heavily wooded environment, where leaves can create artifacts such as locally varying illumination (including direction) and shadows that can make it difficult to estimate shape and that hinder DNNs from accurately localizing the target. Therefore, normalizing for illumination is a vital preprocessing step for interacting with a real environment.

It is well known that contour perception depends on integration across nearby regions [4, 6, 7] and that shape perception is confounded by the impossibility of disentangling the combined effects of illumination and surface slant on apparent lightness (the 'inverse problem'). We hypothesized that the brain performs illumination-shape disentanglement by inferring the likely illumination from nearby surfaces and grouping related elements to compute the shape feature of a target surface [8], supported by interactions between luminance detectors and edge detectors in visual cortex [9]. Previous reports showed that the luminance of surrounding elements strongly modulates perceived lightness of a central target [10], including under HDR luminance conditions [11-13]. Recently, we extended our understanding of luminance-edge interactions to the HDR domain by discovering a psychophysical phenomenon, where the contextual grouping of nearby co-oriented patches with similar HDR luminance biases the perceived orientation of a central patch [8]. With a previous version of this SNN, we were able to reproduce this human perceptual phenomenon[14]. However, it was tested only on artificial images and was not tested on natural images, and the performance of our neuromorphic algorithm was not tested as a pre-processor for a DNN.

Here, we sought to extend our SNN to enhance target localization under partial occlusion. Our approach is supported by several discoveries from prior cortical research. Grouping of similar visual elements is known to result in contextual facilitation [6], i.e. an improved sensitivity to a weak target which occurs due to specific interconnections between luminance clusters and orientation clusters, as well as feedback from higher cortical regions [9]. The micro-structure of the cortex [15] and inhibitory interneurons [16-19] contribute to the grouping of similar stimuli, including lateral, feedback inhibition, and mutual inhibition. Mutual inhibition enables cortical regions to segregate similar visual information into different distinct sets of neurons with distinct neural activity [16, 17, 19]. This is achieved through inhibitory neurons

"deciding" the group a particular visual feature belongs to [16-18, 20, 21]. Additionally, once the visual features have been grouped, they remain in those groups due to the robustness of the segmentation process to perturbations (such as shifting luminance in high contrast situations), with the visual percepts "stuck" in their basin of attraction [16, 17, 19, 20] .

With this understanding of the basic computational structure for edge-luminance interactions, our aim was to evaluate an SNN's ability to improve automatic target recognition in a heavily occluded HDR forest setting. We modified our previously developed spiking neural network [14] to preprocess images by automatically reducing local contrast (e.g. of the background forest) before passing them to a DNN for target detection and localization. To do this, we used LOIHI neuromorphic hardware emulated with Nengo [22, 23].

2 Methods

2.1 HDR images in a Wooded Environment

We used a Nikon D7100 camera to capture 424 shots featuring targeting dummies in a forested area during the summer (Figure 1). Each shot comprised a 6000 x 4000 pixel high dynamic range (HDR) lossless 14 bit NEF image and an 8 bit JPG image. For each camera and dummy location, we captured multiple shots by fixing the ISO at 200, f 2.2, and logarithmically varying the exposure duration. To pose real-world challenges for automatic target recognition, we placed the targets across a range of lighting and occlusion conditions and camera distances, and we pseudorandomly balanced the target positions across the image frame to evenly span x-y pixel positions across different levels of occlusion. We perceptually grouped the images into four categories according to the approximate level of occlusion by leaves, varying from 0-24% ('0%'), 25-49% ('25%'), 50-74% ('50%'), and 75-90% ('75%') occlusion (Table 1).

Figure 1: HDR illumination in a wooded area. Example images of a targeting dummy in a wooded area captured with (A) high exposure, (B) low exposure, and (C) an HDR image. The different exposures highlight the importance of proper exposure control in outdoor shooting scenarios. (D,E,F) A scene with more regional variation in HDR illumination, highlighting the need for localized luminance normalization. The heavily occluded target is at the horizontal midline, just right of center.

Table 1: Number of images

%Occlusion	0%	25%	50%	75%	TOTAL
Training	7	10	12	5	34
Validation	3	5	2	5	15
Test	10	10	10	10	40
TOTAL	20	25	24	20	89

For this initial study, we ignored the NEF images and selected the 89 JPG images whose exposure captured the widest range of luminances in the image. We then decimated the images to 600 x 400 pixels and saved them in JPEG format. We subdivided the 89 images into training (34), validation (15), and test (40) sets. The training and validation images were used to train our deep neural network (DNN, Detectron2, https://github.com/facebookresearch/detectron2) and to develop parameters for neuromorphic analysis. The test images were withheld from the network until final testing. The resulting dataset is a valuable resource for future studies in the field of computer vision, particularly in the development and evaluation of algorithms designed for image processing and analysis in forest environments.

2.2 Simulation Environment

We used Nengo, a Python-based scripting language for neural networks [22], to emulate the LOIHI chip spiking neural network (SNN) developed by Intel [22, 23]l, which contains LIF neurons connected by Current Based (CuBa) synapses. The LIF neuron is a differential equation that captures the essence of a biological neuron's action potential and synaptic dynamics, and can be described by a system of differential equations:

$$C_m \frac{dV_i}{dt} = g_L(V_i - E_L) + u_i \quad \text{when } V_i > V_{th} \text{ then}$$
$$V_i \rightarrow V_{rest}$$

$$u_i(t) = g_{ij}\, \alpha(t) * \sum_j \delta_j(t - t_{spike})$$

where V is the membrane voltage, C_m is the membrane capacitance, g_L is the leak conductance, E_L is the leak current, V_{th} is the voltage threshold, and V_{rest} is the reset voltage. Here $u_j(t)$ is the synaptic current. We used the standard parameters for a LOIHI LIF neuron [22]. The LIF neuron captures the essence of a biological neuron's action potential and synaptic dynamics and can be described by a system of differential equations. Whenever the membrane voltage V exceeds the voltage threshold Vth, the neuron fires a spike, which is used to calculate the synaptic current using the equation $u_i(t) = g_{ij}\, \alpha(t) * \sum_j \delta_j(t - t_{spike})$. Here the current is a convolution with exponential decay term $\alpha(t) =$

$\frac{e^{\frac{t}{\tau}}}{\tau}$ and the sum of all the presynaptic spikes $\sum_j \delta_j(t - t_{spike})$. The synapse weight is g_{ij} is excitatory when $g_{ij} > 0$, and it is inhibitory otherwise. Nengo provides a convenient environment for parameter constraining and network simulation, and we use the default parameters for LOIHI in Nengo [22, 23].

2.3 SNN Architecture

We used the LIF model implemented by LOIHI to construct the network. All synaptic weights are in Table 2. We previously reported that a recurrent spiking neural network could reproduce experimental results from our previous study [8] [14]. Here, we extended this SNN to real-world images. We divided the visual field into patch arrays of 5x5 pedestals, each pedestal comprising 5x5 pixels (Figure 2A). For each pedestal, we processed two types of sensory information. Luminance is calculated as the average luminance of the pedestal. Orientation is computed as a convolution of the pedestal with a Sobel filter oriented horizontally or vertically.

We segregated dark and light luminance pedestals using ON/OFF neurons with sigmoidal activation curves. ON neurons had a sigmoid that increased with luminance, while OFF neurons had a sigmoid that decreased with luminance. The threshold and steepness of the sigmoid were the same for all pedestals, with the steepness modulated by the variance of all pedestals (set to a minimum if variance was zero) and the threshold modulated by the average luminance across all pedestals. Each patch array was processed by its own similar subcircuit (see Figure 2B).

Each pedestal was also encoded by an orientation group comprised of two excitatory neurons encoding horizontal and vertical orientations, both of which are connected to a central inhibitory neuron that suppresses the two excitatory neurons. Each of the 25 pedestals within a patch array had a separate 3-neuron group. This created a decision network, where only one of the excitatory neurons could be active at a time.

Within each patch array, all 25 excitatory neurons of the same orientation were connected to each other, resulting in two all-to-all networks. This boosted weak inputs of similar neighboring pedestals, resolving the apparent orientation of pedestals with ambiguous orientation. We also included a global inhibitory neuron to control the firing rate of the excitatory populations. We duplicated the same circuit for the luminance pedestals. However, here the receptive fields were the ON and OFF neurons for each pedestal, resulting in the pedestals being grouped into dark versus light pedestals.

Layer 1 captured the boosting of weak stimuli and HDR via gain normalization but did not capture the effect of contextual modulation of the combination of luminance and orientation pedestals. Our previous behavioral experiment tested two hypotheses of contextual grouping by testing stimuli with different combinations of luminance and orientation similarity. To integrate these two modalities (luminance and orientation) for contextual grouping, we used a similar decision network in layer two, replacing the two excitatory neurons of the first layer with two copies of a 4-neuron circuit (Figure 2C).
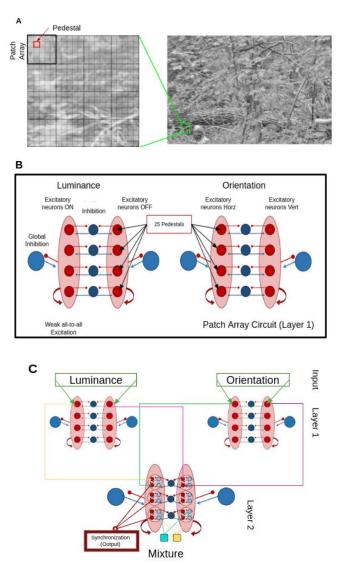
Table 2: Network parameters

Layers 1 & 2	All-to-All Excitation	Excitation to Local Inhibition	Local Inhibition to Excitation	Excitation to Global Inhibition
	0.01	0.5	-0.5	0.005
Layers 1 & 2	Global Inhibition to Excitation	Layer 1 to Layer 2 connection	Synaptic Time constant	
	0.025	6.0	0.005 sec	
Layer 2 only	Excitation to Excitation (within 4-neuron microcircuit)	Excitation to Inhibition (within 4-neuron microcircuit)	Inhibition to Excitation (within 4-neuron microcircuit)	
	1.0	0.25	-0.25	
Layer 2 only	Inhibition to Inhibition (within 4-neuron microcircuit)	Bias into inhibition 1 neuron (yellow)	Bias into inhibition neuron 2 (cyan)	
	-0.5	-0.1	0.0	

Figure 2: *Spiking neural network (SNN) circuit schematic. (A) Every image is subdivided into patch arrays composed of 5x5 pedestals. Each pedestal is 5x5 pixels. Neural circuit architecture: (B) In Layer 1, two all-to-all excitatory neuron groups (red) represent each grid location, with an inhibitory neuron (blue) enforcing exclusivity between stimulus groups. (C) Luminance and Orientation information are passed from Layer 1 to Layer 2. The network has no output in the traditional sense; the synchronization of Layer 2 neurons is the output. Additionally, a global inhibitory neuron maintains optimal firing rates. In Layer 2, 4-neuron circuits replace excitatory neurons and allow for the integration of different modalities. Excitatory control neurons (cyan and yellow) regulate the contribution of each modality from the first layer. See Table 2 for synaptic weights.*

The 4-neuron circuit is a flexible motif that is easy to control [14, 19]. It has two mutually connected excitatory neurons and two mutually connected inhibitory neurons, with the inhibitory neurons providing inhibitory feedback to the excitatory neurons. This gives it the ability to integrate information flexibly and make decisions based on the information presented. The circuit is controllable with top-down control onto the inhibitory neurons (represented by the yellow and cyan bias currents), allowing it to titrate which feature (orientation or luminance) to use for grouping. In the case of ambiguous stimuli, titration allows for the other property to contribute to the grouping. Like Layer 2, the grouping was mediated by a central inhibitory neuron that inhibited the 4-neuron circuit's excitatory neurons. In this way, at least one of the 4-neuron circuits was active, thus segregating the image into two populations. Just like Layer 1, there were also two global inhibitory neurons to help constrain firing rates. The network parameters for this circuit are listed in Table 2.

To pre-process images with our SNN, we initially converted them to grayscale. The grayscale image was then fed into the luminance detection side of the network. For the texture detection side, we first applied a 7-pixel Gaussian blur to the image to reduce noise (Figure 3A). Following the SNN processing described above, we obtained firing rate (Figure 3B) and synchrony (Figure 3C) from the activity of neurons in Layer 2. Further image pre-processing was based on the synchrony, and the firing rate information was not used.

2.4 Contrast Adjustment and DNN

To avoid the granularity associated with firing rates, we implemented Golomb's metric for synchronization in our network [24]. Synchronization $\chi^2 = \frac{\sigma_V^2}{\frac{1}{N}\Sigma\sigma_{V_i}^2}$ where σ_V^2 was the variance of the sum of all voltage traces across all neurons, while $\sigma_{V_i}^2$ was the variance of a single voltage trace. The synchronization metric involved looking for spike time correlations between excitatory neurons within Layer 2. Synchrony implied the pedestals had similar luminance and texture properties. We calculated the synchronization across all 25 pedestals within a patch array (not across all patch arrays), that is, the synchronization of 1500 neurons in layer 2. Figure 3C shows the synchronization indices across an image.

Subsequently, we employed the synchronization metric to attenuate the contrast of specific patch arrays in the image. The premise is that highly synchronized patch arrays indicate interesting objects for the network to attend to, while poorly synchronized ones can be disregarded. Therefore, we adjusted the new pixel value c downward based on the synchronization level χ^2, with $c = (f - 0.5)\frac{1+tanh(k\,\chi^2)}{2} + 0.5$ where f is the original pixel value (between 0 and 1). Note that this equation has a free parameter $k = -0.6$ which controls how much the patch array's contrast is attenuated by lack of synchrony. We tested the network with different k values to determine the optimal k. Figure 3D shows an example of an image after SNN-based contrast adjustment. This adjusted image was then passed to a deep neural network (DNN, Detectron2), a standard "GeneralizedRCNN" meta architecture, which consisted of a resnet Feature Pyramid Network (FPN) backbone trained on the COCO dataset, running on Python 3.7.9, CUDA 11.0, and Tesla P100-16GB on an HPE SGI 8600 system. The FPN had a top-down architecture with lateral connections to extract features at different scales (Ren et al., 2015). By using desynchronization to attenuate local image contrast, we aimed to draw the DNN's attention away from background noise and emphasize important visual features that are more likely to be correlated and grouped across nearby regions of the image, facilitating effective object recognition by the DNN.

3 Results
3.1 Color Images

We used the Detectron2 DNN to detect a targeting dummy using 34 training images, 15 validation images, and 40 test images. We compared the output of the DNN without versus with SNN preprocessing. In the case of SNN preprocessing, the DNN was both trained and tested on pre-processed images. Figure 4 shows sample output images, including true positives (Figs. 4A and 4C) and false positives (Figs. 4B and 4D). Figs. 4A and 4B show results of training and testing without pre-processing, and Figs. 4C and 4D show results of training and testing on pre-processed images. We tested 10 images in each of the four occlusion brackets (0%, 25%, 50%, and 75%), with versus without pre-processing. The results showed that the
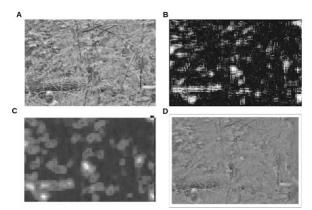

Figure 3: Image preprocessing example. Analysis of a target dummy in a wooded environment. (A) Raw grayscale image of the dummy (lower left) captured in its natural environment. (B) Firing rate output of the SSN, with white indicating higher firing rates. (C) Synchronization index for the patch arrays, with white indicating higher synchronization. (D) Contrast-adjusted image for enhanced visual clarity. These results provide insight into the neural responses to visual stimuli in a naturalistic setting.

true positive rate decreased as the level of occlusion increased. SNN preprocessing of the image weakly improved the true positive rate for all occlusions except 0% (Fig. 4E), and this improvement was significant at 50% occlusion (31.25% TP without preprocessing, 38.75% TP with preprocessing, i.e. a 24% relative improvement with SNN, p < 0.02, n = 8 permutations of 34 Training and 15 Validation images) without significantly affecting false positives except in the 0% occlusion condition (Fig. 4F). This suggests that our preprocessing circuit can be effective in increasing true positive rate in more challenging conditions.

3.2 Grayscale Images

Next, we examined the effect of SNN pre-processing on grayscale versions of the same training, validation, and test images. Again, we compared the DNN output with versus without SNN preprocessing across the four occlusion brackets. Figure 5 shows sample images of true positives (Figs. 5A and 5C) and false positives (Figs. 5B and 5D) resulting from the DNN, with (Fig. 5C, 5D) and without SNN preprocessing (Fig. 5A, 5B). Overall the DNN had better performance (more true positives, fewer false positives) for color versus grayscale images. SNN preprocessing weakly but non-significantly improved true positives across all four occlusion brackets (Fig. 5E). SNN preprocessing tended to reduce false positives in the 75% and 50% occlusion bracket and increase false positives in the 0% and 25% occlusion brackets, but these differences were not significant (Fig. 5F). These results with grayscale images
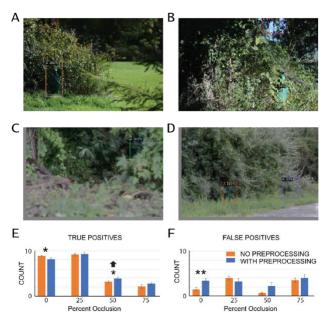
Figure 4: Color SNN preprocessing improved detection under occlusion. Analysis of deep neural network (DNN) performance on target detection in wooded environments. (A) Example of a DNN correctly localizing the targeting dummy. (B) Example of DNN result with a true positive (yellow box) and a false positive (red box). (C) Example of an SNN-preprocessed DNN output correctly identifying the targeting dummy. (D) Example of an SNN-preprocessing image with a true positive (red box) and false positive (blue box). (E) Comparison of DNN true positive results without (orange) versus with (blue) SNN preprocessing, across levels of occlusion. (F) Comparison of DNN false positives grouped by occlusion. * p < 0.02, ** p < 0.01, n = 8, uncorrected for multiple comparisons

are generally consistent with those with color images, although here the differences were not significant.

4 Conclusions

The results of our study demonstrate the effectiveness of using spiking neural networks for image preprocessing, to improve machine vision performance in detecting targets under partial occlusion and natural illumination in a wooded environment. We showed that for color images, SNN preprocessing significantly improved the true positive detection rate by 24% at moderate (50%) occlusion without significantly affecting the false positive rate. For grayscale images, SNN preprocessing had weak but non-significant improvements to both true positives and false positives, and the improvement was better for the occluded conditions (25% and greater).

It is important to note that these improvements in DNN performance were achieved with our initial attempt at an SNN image preprocessing circuit, building upon prior developments in SNN-based decision circuitry. Due to the limitations of software emulation, we were only able to test a simplified
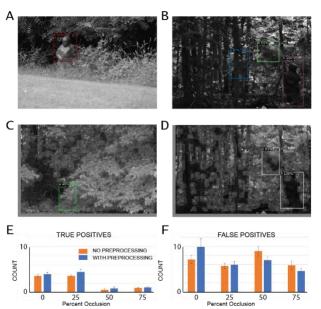

Figure 5: Grayscale preprocessing had no significant effect on true positives and false positives. (A) An example of an unprocessed image resulting in a true positive, correctly identifying the targeting dummy. (B) A DNN result with a true positive (red box) and two false positives (green and blue boxes). (C) Same as above, but with SNN preprocessing showing a true positive, correctly identifying the targeting dummy. (D) A preprocessing result with a true positive and a false positive. (E) DNN true positives grouped by occlusion level, comparing results without (orange) versus with (blue) SNN preprocessing. (F) DNN false positives grouped by occlusion level. p = n.s., n = 8

circuit with two orientations, two luminances, one spatial frequency, and highly spatially-restricted contextual processing. Also, although our circuit is designed with the goal of implementing it on HDR images, these initial tests were with SDR images, so they do not leverage the full capabilities of SNN preprocessing. Nonetheless, our findings suggest that our SNN preprocessing circuit approach can improve the real-world robustness of DNN-based object detection systems, especially in challenging conditions with high occlusion and non-uniform high-dynamic range illumination.

5 Discussion and Future Work

Our results suggest that it may be useful to reconsider what is a pre-attentive feature for visual search. Wolfe and Utochkin [25] argue that whereas color and orientation are features, the conjunction of color and orientation is not a 'feature' because it can be decomposed to color versus orientation. However, this definition of feature is driven by the notional ideals of feature separability for design of experimental stimuli, whereas ample neurophysiological evidence supports a continuum of feature processing and tight integration within early visual cortex of cats and monkeys [8, 26] in support of pre-attentive vision

(likely also in support of top-down attention). Our SNN image pre-processing results show that grouping effects from combining luminance and edge processing can improve target recognition by a DNN, and this is also apparent from the pre-processed images themselves (Fig. 3D).

Our results for color images differed from those for grayscale images, in that neuromorphic pre-processing of color images improved true positives but did not improve false positives, whereas for grayscale images it weakly but non-significantly improved both true and false positives. We believe this is due to two factors. First, the DNN was pre-trained on and optimized for color images, which resulted in better performance on the color images. However, the neuromorphic pre-processing was limited to the grayscale version of the color image due to the limited number of neurons we could include in the emulation. Second, the lower accuracy of the DNN for grayscale images resulted in fewer true positives and false positives, which rendered the results non-significant even though they were in the right direction. We believe these factors can be addressed in future work in which our network is hardware accelerated.

The limited but promising improvements provided by our current SNN model suggest that this first-order SNN approximation of cortical pre-attentive visual processing may be a fruitful path to pursue. To improve object detection beyond just vertical and horizontal orientations in our SNN model, we need to increase the number of edge detection groups to three or more. This can be accomplished by expanding the number of excitatory groups in Layer 1, with one decision inhibitory neuron servicing multiple excitatory neurons. To synthesize this information in Layer 2, a simplified version of a dendritic arbor can be introduced [27]. Each branch of the dendritic tree represents one possible grouping of pedestal luminances, allowing for multimodal integration. An inhibitory control neuron can use shunting inhibition to gate which dendritic branch is silenced, based on the local connectivity of VIP and SST neurons that is key for incorporating contextual information in the cortex [20]. This neural architecture could allow the network to handle arbitrary synthesis and perception of groupings while still reproducing the particular results of the experiment. Additionally, because of the dendritic tree, this model could easily scale to other modalities such as color, shape, or texture.

This approach has the potential to significantly improve object detection by allowing for more complex and nuanced grouping of visual information. However, implementing this model will require a larger number of neurons and layers, particularly if it is to be extended to natural scenes. For example, additional improvements may be gained by increasing the number of spatial frequencies and phases in our orientation neurons, increasing the sampling density, and increasing the effective distance and variety of contextual interactions (e.g. to include co-circular interactions; Zucker, 1985 [28]), and extending the model for color processing [3]. Furthermore, the effectiveness of this approach will need to be compared to other models such as convolutional neural networks (CNN) and connectionist models, and validated against real psychometric experiments using maximal likelihood frameworks [29].

In addition to increasing the number of luminance groups and introducing a dendritic arbor, there are also other ways to improve contrast adjustment. One potential issue that may be affecting the results is the presence of darker regions, which can negatively impact image segmentation and object detection. However, we note that the primary visual cortex preferentially responds to dark versus light [30], which we speculate may support processing of HDR scenes. Additionally, the choice of the optimal k value for adjusting the contrast could be further optimized. Testing a wide range of k values could help us zero in on the optimal k.

Furthermore, it would be interesting to benchmark this algorithm against traditional methods such as histogram equalization[31] or adaptive thresholding [32] can be used to improve contrast and mitigate the impact of dark regions.

In summary, our study demonstrates the importance of SNN preprocessing techniques in improving the performance of DNN-based object detection systems and highlights the potential benefits of our preprocessing circuit, especially in challenging conditions. Our method can be applied to other object detection systems, and we believe that it has the potential to enhance the performance on tasks such as optic flow and obstacle avoidance, particularly in the real world under non-uniform illumination and high clutter. Our findings thus provide valuable insights for future research across the fields of neuromorphic computing, computer vision, and deep learning and advance research into their combined development to address real-world challenges.

ACKNOWLEDGMENTS

REFERENCES

[1] Hafner, D., Demetz, O. and Weickert, J. *Simultaneous HDR and optic flow computation*. IEEE, City, 2014.

[2] Lazar, A. A., Ukani, N. H. and Zhou, Y. Sparse identification of contrast gain control in the fruit fly photoreceptor and amacrine cell layer. *The Journal of Mathematical Neuroscience*, 10, 1 (2020), 1-35.

[3] Albright, T. D. and Stoner, G. R. Contextual influences on visual processing. *Annual review of neuroscience*, 25, 1 (2002), 339-379.

[4] Gilbert, C., Ito, M., Kapadia, M. and Westheimer, G. Interactions between attention, context and learning in

primary visual cortex. *Vision research*, 40, 10-12 (2000), 1217-1226.

[5] Adelson, E. H. *Lightness Perception and Lightness Illusions*. MIT Press, City, 2000.

[6] Chen, C.-C. and Tyler, C. W. Lateral modulation of contrast discrimination: Flanker orientation effects. *Journal of Vision*, 2, 6 (2002), 8-8.

[7] Cass, J. R. and Spehar, B. Dynamics of collinear contrast facilitation are consistent with long-range horizontal striate transmission. *Vision research*, 45, 21 (2005), 2728-2739.

[8] Hung, C. P., Callahan-Flintoft, C., Fedele, P. D., Fluitt, K. F., Odoemene, O., Walker, A. J., Harrison, A. V., Vaughan, B. D., Jaswa, M. S. and Wei, M. Abrupt darkening under high dynamic range (HDR) luminance invokes facilitation for high-contrast targets and grouping by luminance similarity. *Journal of Vision*, 20, 7 (2020), 9-9.

[9] Hung, C. P., Ramsden, B. M. and Roe, A. W. A functional circuitry for edge-induced brightness perception. *Nature neuroscience*, 10, 9 (2007), 1185-1190.

[10] Zemach, I. K. and Rudd, M. E. Effects of surround articulation on lightness depend on the spatial arrangement of the articulated region. *JOSA A*, 24, 7 (2007), 1830-1841.

[11] Allred, S. R. and Brainard, D. H. A Bayesian model of lightness perception that incorporates spatial variation in the illumination. *Journal of Vision*, 13, 7 (2013), 18-18.

[12] Allred, S. R., Radonjić, A., Gilchrist, A. L. and Brainard, D. H. Lightness perception in high dynamic range images: Local and remote luminance effects. *Journal of vision*, 12, 2 (2012), 7-7.

[13] Radonjić, A., Allred, S. R., Gilchrist, A. L. and Brainard, D. H. The dynamic range of human lightness perception. *Current Biology*, 21, 22 (2011), 1931-1936.

[14] White, A. J., Hung, C. P. and Lo, C.-C. *HDR luminance normalization via contextual facilitation in highly recurrent spiking neural networks*. SPIE, City, 2022.

[15] Wang, X.-J. Decision making in recurrent neuronal circuits. *Neuron*, 60, 2 (2008), 215-234.

[16] Kogo, N., Kern, F. B., Nowotny, T., van Ee, R., van Wezel, R. and Aihara, T. Dynamics of a mutual inhibition circuit between pyramidal neurons compared to human perceptual competition. *Journal of Neuroscience*, 41, 6 (2021), 1251-1264.

[17] Koyama, M. and Pujala, A. Mutual inhibition of lateral inhibition: a network motif for an elementary computation in the brain. *Current opinion in neurobiology*, 49 (2018), 69-74.

[18] Lin, Y.-S., Chen, C.-C. and Greenlee, M. W. The role of lateral modulation in orientation-specific adaptation effect. *Journal of Vision*, 22, 2 (2022), 13-13.

[19] Liu, B., White, A. J. and Lo, C.-C. Augmenting Flexibility: Mutual Inhibition Between Inhibitory Neurons Expands Functional Diversity. *bioRxiv* (2020), 2020.2011.2008.371179.

[20] Wang, X.-J. and Yang, G. R. A disinhibitory circuit motif and flexible information routing in the brain. *Current opinion in neurobiology*, 49 (2018), 75-83.

[21] Cavanaugh, J. R., Bair, W. and Movshon, J. A. Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *Journal of neurophysiology*, 88, 5 (2002), 2530-2546.

[22] Bekolay, T., Bergstra, J., Hunsberger, E., DeWolf, T., Stewart, T. C., Rasmussen, D., Choo, X., Voelker, A. R. and Eliasmith, C. Nengo: a Python tool for building large-scale functional brain models. *Frontiers in neuroinformatics*, 7 (2014), 48.

[23] Davies, M., Srinivasa, N., Lin, T.-H., Chinya, G., Cao, Y., Choday, S. H., Dimou, G., Joshi, P., Imam, N. and Jain, S. Loihi: A neuromorphic manycore processor with on-chip learning. *Ieee Micro*, 38, 1 (2018), 82-99.

[24] Golomb, D. and Rinzel, J. Dynamics of globally coupled inhibitory neurons with heterogeneity. *Physical review E*, 48, 6 (1993), 4810.

[25] Wolfe, J. M. and Utochkin, I. S. What is a preattentive feature? *Current opinion in psychology*, 29 (2019), 19-26.

[26] Chu, C. C., Chien, P. F. and Hung, C. P. Tuning dissimilarity explains short distance decline of spontaneous spike correlation in macaque V1. *Vision Research*, 96 (2014), 113-132.

[27] Li, S., Liu, N., Zhang, X., McLaughlin, D. W., Zhou, D. and Cai, D. Dendritic computations captured by an effective point neuron model. *Proceedings of the National Academy of Sciences*, 116, 30 (2019), 15244-15252.

[28] Zucker, S. W. Early orientation selection: Tangent fields and the dimensionality of their support. *Computer Vision, Graphics, and Image Processing*, 32, 1 (1985), 74-103.

[29] Prins, N. and Kingdom, F. A. Applying the model-comparison approach to test specific research hypotheses in psychophysical research using the Palamedes toolbox. *Frontiers in psychology*, 9 (2018), 1250.

[30] Yeh, C.-I., Xing, D. and Shapley, R. M. "Black" responses dominate macaque primary visual cortex v1. *Journal of Neuroscience*, 29, 38 (2009), 11753-11760.

[31] Chin, A. L., Lin, C. Y., Fu, T. F., Dickson, B. J. and Chiang, A. S. Diversity and wiring variability of visual local neurons in the Drosophila medulla M6 stratum. *Journal of Comparative Neurology*, 522, 17 (2014), 3795-3816.

[32] Gonzalez, R. C. and Woods, R. E. Thresholding. *Digital image processing* (2002), 595-611.