2025 IEEE International Conference on Image Processing Workshops (ICIPW) | 979-8-3315-7799-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICIPW68931.2025.11386210

BLACK-MIRROR LIGHT-PROBE COW MODEL FOR SPECULAR HIGHLIGHT REMOVAL TO IMPROVE HOLSTEIN CATTLE IDENTIFICATION

Manu Ramesh and Amy R. Reibman

School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN USA

ABSTRACT

Computer-vision methods for identifying Holstein cattle based on their appearance depend on the correct distinction of the black and white regions on their coats. Bright specular highlights on black regions of cows could be registered as white regions, confusing the identification algorithm and leading to misidentifications. As a solution, we develop a per-pixel specular highlight removal method designed to handle non-uniform lighting from multiple illuminants on a per-scene basis. It is fast as it does not compute image statistics or use neural-network inference. It first segments and tracks a black, mirror-like object – a fully black cow – in a video of the scene, and then uses it as a light probe to map the scene illumination at the pixel level. This illumination map is then used to remove specular highlights on black regions of cows appearing at other times in the scene. Following this, we binarize the cow-pixels using a low threshold to separate the white regions from the black ones. When coupled with an available cattle recognition system, our method outperforms both the traditional and the recent deep-learning color-constancy methods, in improving recognition performance. Our method improves the performance of this cattle recognition system by almost 7%.

***Index Terms*—** specular highlight removal, illumination maps, color-constancy, light-probes, cattle identification

1. INTRODUCTION

Scene illumination can adversely impact the performance of computer-vision tasks. Specular reflections often become a nuisance factor for downstream vision tasks because they hide the true colors of the underlying objects. This is especially true for the task of segmenting between black and white regions on an image. For instance, when scanning a barcode, specular highlights can corrupt the code by rendering black regions as white. Further, since these colors are achromatic, segmentation based on hue or saturation cannot be used.

This problem also affects identification of Holstein cows as they too have black and white coat patterns. Errors due to specular-highlights thus propagate to cattle analytics applications built on top of identification systems such as individual tracking, feed and behavior monitoring, and affect their quality and reliability.

The problem becomes more complex when the scene is illuminated by multiple light sources with different hues and intensities. This spatially non-uniform scene lighting would cause parts of a Holstein cow to appear with different colors (Hue, Saturation, and Lightness), as it moves around in the scene, hindering its identification.

To overcome this problem, in this paper we devise a method to use completely black cows in the scene as mirror-like light probes to sample the per-pixel scene illumination. Considering the illumination values as the new black-points at each pixel location, we apply black-point correction to every video-frame of the scene. The black and white regions on the cows of interest are then separated by binarizing the color-corrected pixels using a low threshold value. This color-correction is performed before any downstream vision task. In this way, our color-constancy method is designed to handle non-uniform illumination from multiple-illuminants.

Our highlight removal method works fast because it neither computes image statistics, nor does it use neural-network inference. All the heavy lifting is done only once – during the computation of the illumination map. This inference speed is crucial for downstream tasks such as real-time cattle tracking.

The key contributions we make in this paper are:

- We develop a fast, per-pixel color-constancy/specular-highlight-removal method that maps illumination using black-mirror objects in the scene. This method focuses on improving the performance of downstream computer vision tasks instead of making the image perceptually pleasing.
- We show that this method improves performance of a cattle recognition system, and outperforms both traditional and recent color-constancy methods for this task.

We first discuss the related works in Sec. 2. Next, we describe our method of using black-mirror light-probes for per-pixel specular highlight removal in Sec. 3. We then provide the details of applying our method to the task of cattle recognition, along with the datasets and evaluation metrics used for the task in Sec. 4. Later, in Sec. 5 we present both

Thanks to students and workers at the Purdue Dairy for collecting data for our dataset.

quantitative and qualitative comparison of results from applying our method to the task of cattle recognition. We conclude our findings in Sec. 6.

2. RELATED WORKS

2.1. Color-constancy methods

Scene illumination has been a long standing challenge for computer vision. Traditional methods for reducing illumination effects were designed to work only for uniformly illuminated scenes. Of these, the White Patch Retinex algorithm color-corrects or white-balances an image by re-scaling it using the set of brightest pixels in the scene. The Grayworld assumption algorithm color-corrects an image by scaling each of its color channels with the inverse of the corresponding channel mean. Later, methods such as Homomorphic Filtering were designed to handle non-uniform scene illuminations. This method color-corrects the image by suppressing the low frequency components of the logarithmic intensity of each color channel. Detailed descriptions of these traditional methods are in [1].

All the above methods model the observed surfaces to be Lambertian, implying that they cannot handle specular reflection. Unlike these traditional methods, ours does not compute image statistics while color-correcting each image of the same scene, and thus is considerably faster.

The more recent color-constancy methods tackle scene illumination effects using deep neural-networks [2, 3, 4]. The method in [2] uses a two-stage encoder-decoder neural architecture (M2-Net) for coarse and fine removal of specular highlights in natural scenes. An end-to-end highlight removal system that uses adaptive hybrid attention mechanisms with two transformers (DHAN-SHR) is described in [3]. The Intrinsic Image Decomposition (IID) based method in [4] uses a multi-stage procedure to compute the diffuse reflectance, diffuse shading, and the specularity components of color images with the help of encoder-decoder neural-networks. We can then obtain the specularity-free image by multiplying the diffuse reflectance with the de-saturated diffuse shading.

The deep-learning methods tend to work well on data that is within their training distribution, and may not adapt well to new and out-of-distribution data. Our method on the other hand, samples ground truth illumination of the required scene directly. This makes it more accurate. Also, our method works faster, and needs significantly lower compute than these learning based end-to-end color-correction methods because unlike them, ours does not apply neural-network inference.

2.2. Cattle recognition

We apply our method to a computer vision implementation of cattle recognition that uses their coat patterns in the top view for recognition. This is because the top view is usually free from occlusions; the body of a cow also covers significant scene-area making it a useful light-probe. Some of these methods include [5, 6, 7, 8, 9, 10]. Of these, [5] uses a dataset containing only still images of cows, and [6] has no black cows in its dataset. The methods in [7, 8] use a dataset that comes with video segments of cows walking across a non-uniformly illuminated scene. This dataset also contains four black cows that can be used to test our method. However, we choose to use the recognition system in [9, 10] because of its simplicity, modularity, and its dataset that has video segments of isolated individual cattle moving in the scene which helps us to easily track the required black cow for light-probing.

3. METHOD

In this section, we first explain how we model a black-mirror light-probe in Sec. 3.1. Later, the procedures for using a black cow light-probe to sample illumination at different locations in the scene and for using this illumination information for color-correction are explained.

3.1. Modeling cows as black-mirror light-probes

To sample scene illumination, we use black cows in the scene modeled as perfect black-mirrors. For this, we use the Dichromatic Reflection Model [1], which models the camera sensor response as a linear combination of diffuse/matte and specular reflection components as per Eq. 1. For a camera with narrow-band sensors i for primary color wavelengths λ_i, $i \in \{r, g, b\}$, we have,

$$I_i = s_M R_{M,i} L_i + s_S R_{S,i} L_i \tag{1}$$

where I_i is the response of sensor i, $R_{M,i}$ and $R_{S,i}$ are the reflectances with respect to matte and specular reflections at wavelength λ_i, L_i is the illuminant, and s_M and s_S are two scale factors that depend on object (cow) geometry.

We simplify the above model by imposing the following conditions.

- The cow used to probe the illumination is completely black and mirror like, with $R_{M,i} = 0$ and $R_{S,i} = 1 \ \forall i$.
- The camera implements a scaled orthographic projection, and the directions of object (cow) normal, the illumination, and the camera are all parallel to each other. So, $s_M = 1$ and $s_S = 1 \ \forall i$.

The simplified model for each pixel location (x, y) becomes,

$$I_{i,(x,y)} = L_{i,(x,y)}, \tag{2}$$

indicating that the light sampled by the camera at each pixel location is the light from the illuminant irradiating on the black cow at that location. In other words, the sampled reflection map is the scene illumination map.

3.2. Illumination mapping and color-correction

Using an instance segmentation neural-network model, we track the black-mirror light-probe cow described above to sample illumination at different locations as it moves in a video of the scene – like a brush painting on a canvas. For pixel locations where the illumination is sampled in multiple video frames due to overlapping instance masks, the maximum value of illumination among all the frames is retained as the final value. Considering areas not covered by the probe-cow as the 'uncertain regions', we estimate illumination at those uncertain pixel locations using normalized convolution [11] with a Gaussian kernel as our Point Spread Function (PSF), in linearRGB color space. This creates an illumination map of the scene.

To color-correct a given image, we use the mapped illumination as the black-point at each pixel location and black-point correct the image in linearRGB color-space. Eq. 3 gives the formula to compute values of each sub-pixel P_i, ($i \in R, G, B$), in the color-corrected image.

$$P_{i,(x,y)} = clip_0^1 \left(\frac{p_{i,(x,y)} - \mathcal{L}_{i,(x,y)}}{1 - \mathcal{L}_{i,(x,y)} + \epsilon} \right), \tag{3}$$

where, p_i is the sub-pixel value before color-correction, $\mathcal{L}_i$ is the sampled illumination from Eq. 2 in linearRGB color-space which is used as the black-point, (x, y) indicate the pixel location in the image, and ϵ is a small number included to avoid division by 0. The color space of the color-corrected frame is then converted back to sRGB, and binarized with a low threshold to separate the black and white regions. Note that the black-point corrected white regions of the other Holstein cows will have colors complementary to the illuminant color (black-point) at each pixel. However, they will be converted to white after the thresholding operation.

4. EXPERIMENTS

We evaluate our specularity removal method by applying it to the chosen Holstein cattle recognition system to separate the black and white regions of cows. We use the cattle-identification performance metrics to measure the effectiveness of our highlight-removal method. Further, we also compare our method with other color-correction/specularity-removal methods in the literature on the same task.

In this section, we first present a brief overview of the selected cattle recognition system [9, 10], the dataset, and the evaluation metric we use. Later, we explain how we couple our specularity removal method with it.

4.1. The cattle recognition system

We present a simplified diagram of the selected cattle recognition system [9, 10] with our color-correction module embedded in Fig. 1. To identify a cow, the system localizes and extracts the cow instance using keypoints and instance masks, aligns it into a canonical pose, thresholds it to create a matrix barcode, and finds the nearest matching barcode in its dataset of training cows. The feature space of all the matrix barcodes generated from the training cow data is called the 'cattlog' [9].

In this paper, we replace their keypointRCNN [12] keypoint detector with an HRNet [13] one to improve cow localization. Further, while the authors of [9] use only one training image to learn the identity of an individual cow into the cattlog, we use multiple images of the cow in different locations to reduce the influence of illumination.

4.2. Datasets

For all experiments in this paper, we use the same datasets as [9] to train and evaluate the cattle recognition system. The 'cow videos dataset' contains video-segments of cows walking one at a time along a short path on two different days. These segments are called 'cut-videos'. These videos have a pixel resolution of 1920×1080, and frame rate of $30FPS$. The 153 cut-videos from the first day are used to build the cattlog and the 148 cut-videos from the second day are used for evaluating the identification system.

4.3. Evaluation metric

We use the cow identification performance on the testing dataset of cut-videos to measure the performance of our color-correction method. We use the metric *Number of correct identifications* [10], which represents the total number of cow instances that are correctly identified in all frames of all cut-videos that are used for evaluation. An increase in this metric indicates the success of our color-correction method.

In cases where color-constancy methods result in multiple cows being mapped to the same matrix barcode in the cattlog, we consider all cows sharing the same matrix barcodes to be equivalent. So, the identification of a cow as any other equivalent cow is also considered a correct identification.

4.4. Illumination mapping and color-correction

The color-correction block in Fig. 1 applies our method described in Sec. 3 to reduce the effect of scene illumination on cattle recognition. To implement this method, we use a completely black cow as our black-mirror light probe. To satisfy the conditions listed in Sec. 3.1, we model the back of the probe-cow to be completely black, flat and mirror like, the camera to apply scaled orthographic projection, and also each pixel location on the cow to be lit only by a point light source directly above it.

We sample the scene illumination from the cut-video of the probe-cow from the training day using a Mask R-CNN [12] model, and generate the scene illumination map shown in Fig. 2a. The red and blue regions in this illumination map

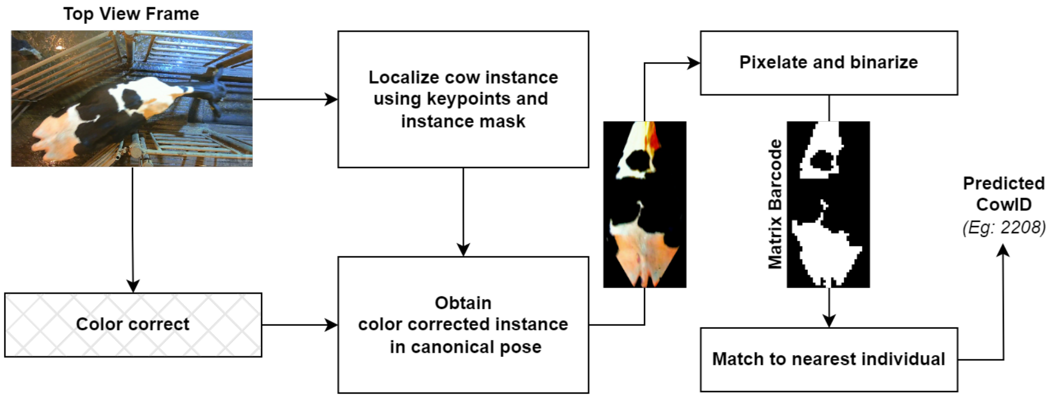
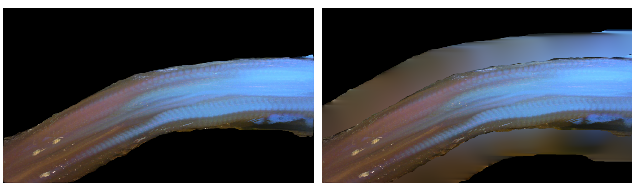


Fig. 1: Simplified flow diagram of the cattle recognition system [9, 10], with our color-correction block embedded and marked with cross-hatching. The color-corrected frame is used to obtain the template aligned image and estimate the cow identity.



(a) Scene illumination map (b) Extended illumination map

Fig. 2: The illumination maps generated from the black probe-cow. Observe the change in hue, intensity and saturation from left to right in the images. The white spots in the bottom left are from the white spots on the legs of the cow that seep into the instance mask.

are due to the probe-cow reflecting light from the red and blue colored lamps illuminating the respective regions in the scene.

Despite the cows being constrained to the path by fences as shown in the Top View Frame in Fig. 1, they do have some room for lateral movement. So, parts of other cows that walk through this scene can appear outside the illumination-mapped region. This unmapped region corresponds to the black region in Fig. 2a where we are uncertain of the scene illumination. Therefore, to cover the entire region of interest, we extend the illumination map slightly into this uncertain region using normalized convolution with a Gaussian kernel of size 301 to obtain the extended illumination map in Fig. 2b. This extended illumination map is used during evaluation to eliminate specular-reflection on all cow instances in the testing set using black-point correction as per Eq. 3. This is followed by a binarization operation with a low threshold, and is carried out by the 'Pixelate and binarize' block in Fig. 1.

5. RESULTS

In this section, we compare our highlight removal method with a few other methods from the literature discussed in Sec. 2. We use multiple instances per cow from the cut-videos of the training set to create the cattlog for all the methods listed below. We present both a quantitative and a qualitative comparison.

5.1. Quantitative comparison

Table 1 shows the values for the 'Number of correct identifications' metric described in Sec. 4.3. The baseline results use the cattle recognition system as described in Sec. 4.1, but without the color-correction block.

As discussed in Sec. 3, with the diminished specular highlights on black regions, we use a low binarization threshold of 75 for our method. For all other methods, and for the baseline, we use a binarization threshold of 127 as in [9]. However, since the IID method makes the black regions of cows very dark, we apply a much lower threshold of 50 to obtain better results.

Both our method and IID found the same three cows to be completely black, and they were all considered equivalent. Similarly, M2-Net found a different set of three cows to be fully black. Further, Homomorphic Filtering, and DHAN-SHR found the same set of two equivalent black cows. The baseline, Gray-world assumption, and the Retinex White Patch methods could not find any equivalent black cows.

From Table 1, among the traditional color-constancy methods, the Retinex White Patch method performs the worst. The results from Gray-world assumption and the Homomorphic Filtering methods are close to the baseline. Among the deep-learning based methods, M2Net [2] performs worse than the baseline, while DHAN-SHR [3] performs better than the baseline. The IID method [4] performs much better. Our

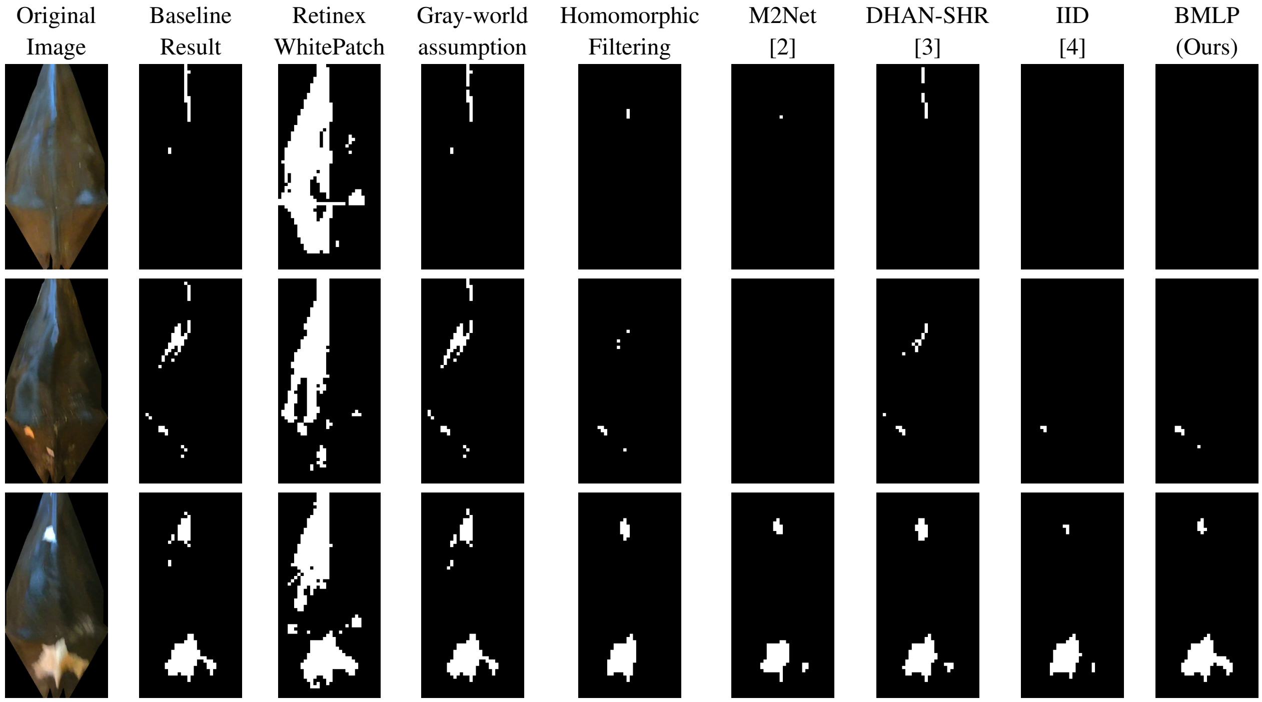


Fig. 3: Qualitative comparison of cattlog matrix barcodes produced by different color-constancy methods. Note that the original image in the leftmost column is just one of the many instances of the cow used to generate the averaged matrix barcodes. In the first row, only ours and the IID method generates the correct barcode. In the second row, only our method identifies the smaller white patch on the bottom of the cow. In the third row too, only our method correctly identifies the top white spot while also identifying bottom white patch as a single connected component.

Experiment Name	# correct IDs
Retinex White Patch	3398
M2Net [2]	3814
Homomorphic Filtering	3914
Baseline	3925
Gray-world	3930
DHAN-SHR [3]	3962
IID Highlight Removal [4]	4068
BMLP (ours)	4198

Table 1: Number of correct identifications for each color-constancy method considered, and the baseline result.

method Black-Mirror Light-Probe (BMLP), produces the best result and outperforms the baseline by almost 7%.

5.2. Qualitative comparison

To observe how the different color-constancy methods perform in reducing lighting effects we present the pixelated binary images in the form of matrix barcodes for a few cows in Fig. 3. Observe how the baseline method falsely classifies the highlights on a completely black cow to be white regions. Our method correctly binarizes the cow instance when some other highlight removal methods fail.

6. CONCLUSION

In this paper, we present a fast highlight removal method for improving the separation of black and white regions on Holstein cows to assist computer-vision systems that identify them. Our method can handle scenes with non-uniform illumination hue and brightness values at the pixel level. The method tracks black mirror-like cows in the scene based on shape cues, and produces an illumination map that is then used to color-correct other cows in the same scene.

We apply our method to an available cattle recognition system and show that our method outperforms both the traditional color-constancy methods and the recent deep-learning based highlight removal methods for this task. Our method could be applicable beyond cattle identification to domains involving identifying actual barcodes on much simpler flat surfaces.

Because our algorithm samples the illumination only in areas swept by the probe cow, it has limited ability to color-correct cows of interest in the unswept areas. Future work will develop non-flat cow models to handle environments with strong directional lighting.

7. REFERENCES

[1] Marc Ebner, *Color Constancy*, Wiley, 2007.

[2] Zhaoyangfan Huang, Kun Hu, and Xingjun Wang, "M2-Net: Multi-stages Specular Highlight Detection and Removal in Multi-scenes," *ArXiv*, 2022.

[3] Xiaojiao Guo, Xuhang Chen, Shenghong Luo, Shuqiang Wang, and Chi Man Pun, "Dual-Hybrid Attention Network for Specular Highlight Removal," *MM 2024 - Proceedings of the 32nd ACM International Conference on Multimedia*, pp. 10173–10181, 2024.

[4] Chris Careaga and Yağiz Aksoy, "Colorful Diffuse Intrinsic Image Decomposition in the Wild," *ACM Transactions on Graphics*, vol. 43, no. 6, pp. 1–12, 2024.

[5] William Andrew, Jing Gao, Siobhan Mullan, Neill Campbell, Andrew W. Dowsey, and Tilo Burghardt, "Visual identification of individual Holstein-Friesian cattle via deep metric learning," *Computers and Electronics in Agriculture*, vol. 185, pp. 106133, 2021.

[6] Thi Thi Zin, Cho Nilar Phyo, Pyke Tin, Hiromitsu Hama, and Iku Kobayashi, "Image Technology based Cow Identification System Using Deep Learning," in *Proceedings of the International Multiconference of Engineers and Computer Scientists*, 2018, pp. 236–247.

[7] Jing Gao, Tilo Burghardt, William Andrew, Andrew W. Dowsey, and Neill W. Campbell, "Towards Self-Supervision for Video Identification of Individual Holstein-Friesian Cattle: The Cows2021 Dataset," *ArXiv ID: 2105.01938*, May 2021.

[8] Jing Gao, Tilo Burghardt, and Neill W. Campbell, "Label a Herd in Minutes: Individual Holstein-Friesian Cattle Identification," in *Image Analysis and Processing. ICIAP 2022 Workshops*, Pier Luigi Mazzeo, , Emanuele Frontoni, , Stan Sclaroff, , and Cosimo Distante, Eds., vol. 13374 LNCS, pp. 384–396. Springer International Publishing, 2022.

[9] Manu Ramesh, Amy R. Reibman, and Jacquelyn P. Boerman, "Eidetic recognition of cattle using keypoint alignment," *Electronic Imaging*, vol. 35, no. 7, pp. 279–1 – 279–6, January 2023.

[10] Manu Ramesh and Amy R Reibman, "SURABHI: Self-Training Using Rectified Annotations-Based Hard Instances for Eidetic Cattle Recognition," *Sensors*, vol. 24, no. 23, pp. 7680, November 2024.

[11] H. Knutsson and C.-F. Westin, "Normalized and differential convolution," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1993, pp. 515–523.

[12] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick, "Mask R-CNN," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 386–397, 2020.

[13] Ke Sun, Bin Xiao, Dong Liu, and Jingdong Wang, "Deep high-resolution representation learning for human pose estimation," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2019-June, pp. 5686–5696, 2019.