The Enrichment of Texture Information to Improve Optical Flow for Silhouette Image

Bedy Purnama[1]*, Mera Kartika Delimayanti[2]
Kunti Robiatul Mahmudah[3], Fatma Indriani[4], Mamoru Kubo[5], Kenji Satou[6]

Graduate School of Natural Science and Technology, Kanazawa University, Kanazawa 9201192, Japan[1, 2, 3, 4]
Telkom of Computing, Telkom University, Bandung 40257, Indonesia[1]
Department of Computer and Informatics Engineering, Politeknik Negeri Jakarta, Depok 16425, Indonesia[2]
Institute of Science and Engineering, Kanazawa University, Kanazawa 9201192, Japan[5, 6]

Abstract—Recent advances in computer vision with machine learning enabled detection, tracking, and behavior analysis of moving objects in video data. Optical flow is fundamental information for such computations. Therefore, accurate algorithm to correctly calculate it has been desired long time. In this study, it was focused on the problem that silhouette data has edge information but does not have texture information. Since popular algorithms for optical flow calculation do not work well on the problem, a method was proposed in this study. It artificially enriches the texture information of silhouette images by drawing shrunk edge on the inside of it with a different color. By the additional texture information, it was expected to give a clue of calculating better optical flows to popular optical flow calculation algorithms. Through the experiments using 10 videos of animals from the DAVIS 2016 dataset and TV-L1 algorithm for dense optical flow calculation, two values of errors (MEPE and AAE) were evaluated and it was revealed that the proposed method improved the performance of optical flow calculation for various videos. In addition, some relationships among the size of shrunk edge and the type and the speed of movement were suggested from the experimental results.

Keywords—Optical flow; silhouette image; artificial increase of texture information

I. INTRODUCTION

Today, a number of cameras are installed in various devices such as smart phones, PCs, security devices, etc. They generate a large amount of still images and videos every day resulting the demand to analyze images and videos by computers has been growing rapidly. In the research field of computer vision, machine learning methods including deep learning have been actively studied for the purpose of achieving better performance. In addition, some studies which have been done by employing the combination of these methods significantly improved the performance [1].

One of the rapidly developing studies of computer vision and deep learning are the detection of moving objects [2] [3], where the most fundamental information is optical flow [4] [5] [6] that indicates corresponding points in two images. In the case of video data, typically the two images are i^{th} and $(i+1)^{th}$ video frames. It is useful for detecting and tracking moving objects in a video.

Among many studies in computer vision, this study focused on the analysis of animal behavior. Similar to the analysis of human behavior, analysis of animal behavior brings many benefits to bionomics, ethology, medical science, fishery, pet industry, etc. It has been studied a long time, mainly in the level of animal's location tracking (coarse-grain level). There are some studies on applying optical flow to detect animal's presence such as outdoor animal detection using Neural Network [7], elephant detection by using color model [8], pig and deer detection by using CNN [9] [10]. On the other hand, analysis of animal's action or motion (fine-grain level) has also attracted interests of researchers in the interdisciplinary field between life science and computer vision.

For the analysis of animal's action or motion, dense optical flow (i.e., optical flow for every pixel of object) should be calculated accurately. The basis of optical flow was developed by Horn and Schunck [11]. In general, optical flow is calculated based on two kinds of information, that is, edges and texture of objects. However, sufficiently rich texture information of animal is not always available. For example, a unicolored animal like a black cat may provide clear edge information, but does not provide any texture information on its body. Moreover, even if an animal has clear texture information on its body surface, it is considered as a silhouette in a backlit video. The previous research focused on how the accuracy of optical flow for videos containing silhouette images of animals can be improved [12]. Since no texture information is available for a silhouette image, popular methods of optical flow calculation provided returned wrong or missing optical flow, especially for the center area of silhouette images (in contrast, optical flows near to the edge are relatively accurate since edge information is available also for silhouette images). To solve the problem, some areas to be improved and additionally calculated by an inpainting algorithm using perspective transformation have been detected in this method in order to solve the problem. However, this method has not been applied to more complex silhouettes.

A new method for improving the accuracy of optical flow for silhouette images is proposed in this paper. In short, the texture information of a silhouette image has been enriched by drawing shrunk edge on the inside of it with a different color (it is called object-in-object). In this research, it is assumed that the object (animal) has already been separated from the background.

*Corresponding Author

To estimate the effect of the proposed method, some videos of animals from a publicly available video dataset were used to estimate the effect of the method. In the selected videos, only one animal is seen. After segmenting an animal from the background of each video, a video which consists of silhouette images was generated. Through the comparison of optical flows for original and silhouette video, it was shown the accuracy of optical flow for silhouette images can be improved in this method. The main contributions of this research are as follows.

- Focusing on silhouette images, seeking novelty to improve the accuracy of existing dense optical flow methods.

- The addition of texture (object-in-object) can increase the accuracy of optical flow for silhouette images.

- The improvement was estimated by using a publicly available video dataset.

II. BACKGROUND

The research regarding the use of feature texture in order to improve the optical flow estimation performance has been conducted by the researchers. It was initiated by Arredondo [13] using differential method to estimate optical flow from feature texture. He used mathematical approach for this research.

On the other hand, the empirical approach was used by Andalibi [14]. Andalibi added the static texture in images with poor texture. The addition of texture mask, where the image area had its optical flow component approaching to zero, did not significantly improve the optical flow estimation. Nevertheless, this was an initial approach to solve the problems on poor texture in optical flow.

In the researches [13] and [14], there are some chances to improve the optical flow estimation for silhouette image sequences cases. On silhouette image sequences, there was almost nothing to find out the feature texture. It was expected that with the addition of texture information to silhouette image sequences, it improved the optical flow performance.

III. MATERIALS AND METHODS

A. Silhouette Image and Video

The previous research [12] focused on animal silhouette images animated by a program (i.e. rotation). The uniqueness of the silhouette animal was that it only had edge information with no texture information. In real conditions, it is believed that animals can be found if they are walking at dusk or they have unicolor, such as black horses. To evaluate the new method proposed in this paper, the DAVIS 2016 dataset [15] was used in this study. Since important objects in each video in the dataset have already been segmented, it assumes that the animal in a video has already been separated from the background. Practically, the result of segmentation is given as a mask image to the original image of each frame of a video. The region of an animal indicated by the mask image is also referred as a region of Interest (RoI). By applying the mask image to the original image, the silhouette image is generated, where the RoI is black and the background is white.

The use of image silhouettes from animal videos revealed a variety of natural animal movements, where each body part of an animal is freely moving or not moving. In case of a simple movement such as walking, only some parts of body are moving. In contrast, almost every body parts are moving in a more complex movement like jumping or flying. From the dataset DAVIS 2016, which contains 50 videos with 480p resolution, 10 videos shown in Table I were chosen in accordance with the criteria that a unicolored animal is taking natural motion in the video. Walk movement shows that animals walk on their feet. Move forward movement indicates that the animal moved without walking. Deformation movement indicates a change in the direction of the animal's orientation. Slow indicates camera / animal movement tends to be slow while fast refers to fast camera / animal movement.

B. Optical Flow

The basic concept of optical flow was initiated by Horn and Schunck [11]. The formulation in (1) is a general equation for the optical flow. There are two solutions that must be resolved, namely, data term and smoothness term. There are traditional developments for optical flow algorithms such as Brox [16], Black [17], Lucas-Kanade [18], TV-L1[19]. There are also the studies that use deep learning such as Flownet [20], DeepFlow [21], EpicFlow [22].

$$E(u,v) = \int_{\Omega} \underbrace{(I_x u + I_y v + I_t)^2}_{data\ term} + \underbrace{\lambda(|\nabla u|^2 + |\nabla v|^2)}_{smoothness\ term} dxdy \quad (1)$$

To form the ground truth of this study, DeepFlow was used [21]. RGB images from the dataset were used to form. This DeepFlow was divided into two processes, namely, feature search using DeepMatching [23] to handle large displacements. The next process was DeepFlow itself, which used deep learning. The output of the DeepFlow is .flo file format that was adjusted to the Middlebury dataset standard [24] [25].

To generate optical flow from silhouette images, the TV-L1[26] [27] was used since this method is the most commonly used as a baseline for optical flow. TV-L1 is based on the coarse method to fine framework numerically, while DeepFlow is based on coarse to fine in the deep learning framework.

TABLE I. DATASET USED IN THIS STUDY

Data Name	Number of Frames	Type of Movements
Bear	82	Walk, slow
Blackswan	50	Move forward, slow
Camel	90	Walk, deformation, slow
Cows	104	Walk, deformation, slow
Dog	60	Walk, deformation, fast
Elephant	80	Walk, deformation, slow
Flamingo	80	Walk, deformation, slow
Goat	90	Walk, deformation, fast
Mallard-fly	70	Walk, fly, high deformation
Mallard-water	80	Move forward, deformation, slow

C. Framework

Fig. 1 illustrates the framework of the experiment in this study. Fig. 2 is the initial stage of the framework. It starts with the formation of Ground Truth from two consecutive RGB JPEG images with a resolution of 480p for each data. DeepFlow [21] was used and the output of this process (i.e. the Ground Truth) is named optical flow t.

The next process described in Fig. 3 is the novelty of this study. From the DAVIS 2016 dataset, annotations (i.e., mask images) with 480p resolution have been provided. In this process, for the enrichment of texture information, a shrunk edge line with white color was drawn at the inside of RoI with black color. From here, a shrunk edge line is called as an artificial texture. This artificial texture was located inside the RoI, where the RoI centroid coincided with the centroid of the artificial texture. In the experiment, the magnification sizes from 10, 20, 30, 40, 50, 60, 70, 80, and 90 percentages were tried.

The silhouette image that has been enriched with an artificial texture was named silhouette image t', which was the output from the silhouette image input t. As shown in Fig. 4, silhouette image t' and silhouette image t' +1 were the input for TV-L1 to form a new optical flow t'.

The final stage was performance evaluation. From Fig. 5, it can be observed that mean endpoint error (MEPE) was used and average angular error (AAE) [24] for evaluating the accuracy of optical flow calculated from silhouette images. MEPE was used to calculate the average magnitude deviation of the vector flow from the estimated optical flow to the ground truth. MEPE was calculated by (2). Meanwhile, AAE was to calculate the average angular deviation of vector flow from the estimated optical flow towards the ground truth. AAE was calculated using (3).

$$EPE = \sqrt{(u - u_G)^2 + (v - v_G)^2} \qquad (2)$$

$$AAE = cos^{-1}\left(\frac{1.0 + u \times u_G + v \times v_G}{\sqrt{1.0 + u^2 + v^2}\sqrt{1.0 + u_G^2 + v_G^2}}\right) \qquad (3)$$
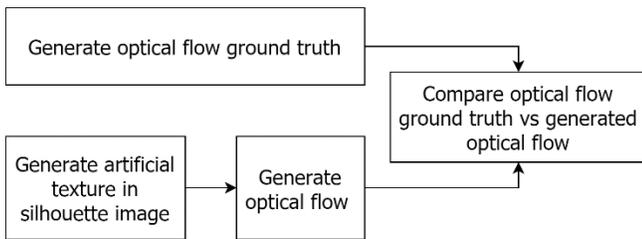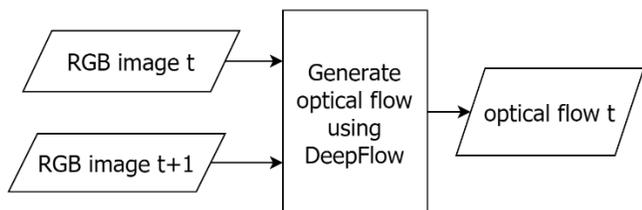

Fig. 1. The Framework of the Experiment.

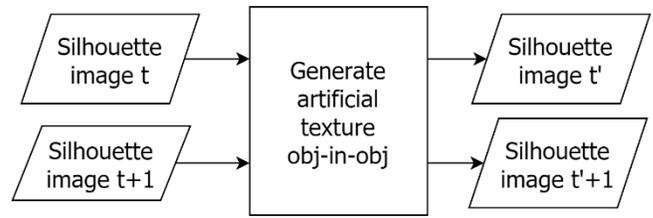Fig. 2. Generated Ground Truth of Optical Flow.
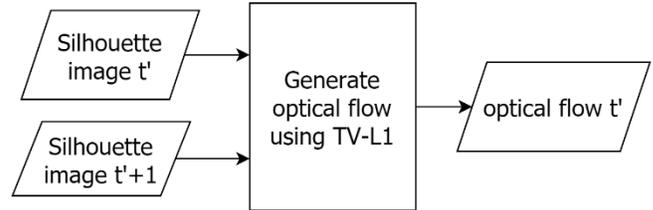

Fig. 3. Generate Artificial Texture.

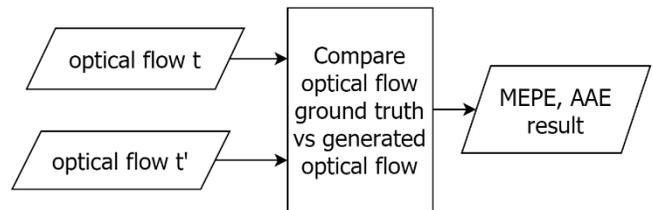Fig. 4. Generated Optical Flow from Sequence of Silhouette Artificial Texture.

Fig. 5. Compare Optical Flow Performance.

IV. RESULT

In this study, the supporting applications used were:

- OpenCV version 4.11 with contrib which provided TV-L1.

- DeepMatching version 1.2.

- DeepFlow version 2.

The experimental results can be seen in Fig. 6 and Fig. 7. Fig. 6 shows an example of ground truth, where two RGB image sequences were processed by DeepFlow algorithm to generate ground truth optical flow. Using the same example, Fig. 7 shows the silhouette images with artificially enriched texture. The first column of it shows the percentage of artificial texture which is gradually added by 10% until it reaches the maximum of 90%. The second and third columns show the silhouette images corresponding to t^{th} and $(t+1)^{th}$ frames respectively. The last column shows the optical flow calculated by TV-L1 algorithm with two silhouette images in the same row as input. In the row named "without" in Fig. 7, it is clearly shown that an almost white area is observed at the center of RoI. It indicates that TV-L1 does not work well for silhouette images without enrichment. Moreover, it is demonstrated that the enrichment of texture might solve this problem.
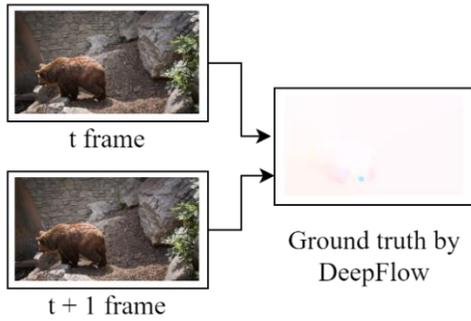
Fig. 6.   Two Consecutive Original Images and its Optical Flow (Ground Truth) Calculated by DeepFlow.

Table II shows the MEPE calculated from optical flow of silhouette images against optical flow ground truth. Starting from no increments, increments of 10 percent to 90 percent of the original size. Meanwhile, Table III shows the AEE for the same case.

In this study, the results of optical flow were evaluated with two measurements, namely, MEPE and AAE were evaluated. Both were achieved by comparing the optical flow of the experimental results with ground truth. To determine the quality of MEPE and AAE, it can be observed from the value, the smaller the value the better the performance.

Table II revealed that there were six animal data (Camel, Cows, Dog, Elephant, Flamingo˒ and Mallard-fly) that achieved the lowest MEPE value with the formation of 90% of magnification artificial texture. In contrast, in the Bear and Blackswan data, the lowest MEPE value was in the formation of 80% of magnification artificial texture. What was different from the others were that Goat and Mallard-water each achieved the lowest MEPE value at the formation of 50% and 70% of magnification artificial texture.

Table III revealed that there were five animal data (Camel, Cows, Dog, Elephant, and Flamingo) which achieved the lowest AAE value in the formation of 90% of magnification artificial texture. In contrast, in the Bear and Blackswan data, the lowest AAE value was in the formation of 80% of magnification artificial texture. What was different from the others were that Goat, Mallard-fly, and Mallard-water, where each achieved the lowest AAE values at the formation of 50%, 60%, and 70% of magnification artificial texture.
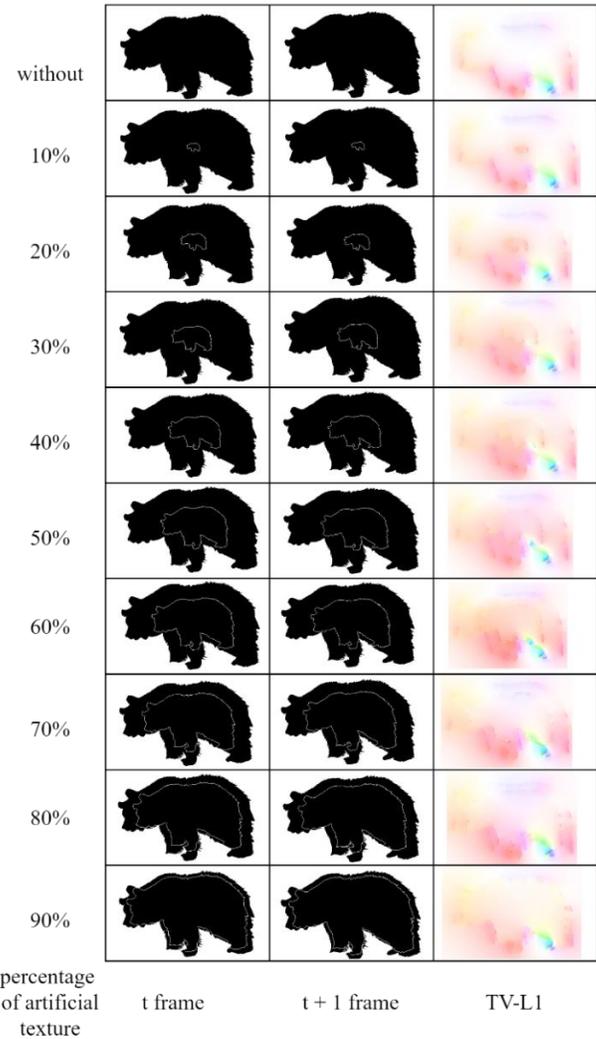

Fig. 7.   RoI with Artificially Enriched Texture and its Optical Flow by TV-L1.

TABLE II.   MEPE OF GROUND TRUTH BY TV-L1 ARTIFICIAL TEXTURE

	Percentage of magnification artificial texture									
	w/o	*10*	*20*	*30*	*40*	*50*	*60*	*70*	*80*	*90*
Bear	1,136	1,098	1,069	1,046	1,028	1,018	1,006	0,999	**0,997**	1,000
Blackswan	1,844	1,793	1,762	1,737	1,713	1,693	1,687	1,664	**1,641**	1,650
Camel	1,432	1,392	1,361	1,333	1,320	1,308	1,299	1,288	1,270	**1,260**
Cows	1,489	1,450	1,415	1,381	1,364	1,351	1,343	1,339	1,338	**1,333**
Dog	6,133	6,110	6,093	6,070	6,067	6,065	6,092	6,125	6,125	**5,940**
Elephant	1,425	1,417	1,409	1,400	1,386	1,373	1,368	1,359	1,351	**1,344**
Falmingo	1,717	1,708	1,697	1,686	1,675	1,668	1,661	1,655	1,647	**1,639**
Goat	3,897	3,865	3,834	3,800	3,780	**3,775**	3,787	3,809	3,836	3,841
Mallard- fly	5,447	5,436	5,427	5,418	5,412	5,412	5,416	5,421	5,423	**5,411**
Mallard -water	2,109	2,082	2,060	2,042	2,029	2,020	2,017	**2,015**	2,024	2,029

TABLE III. AAE OF GROUND TRUTH BY TV-L1 ARTIFICIAL TEXTURE

	Percentage of magnification artificial texture									
	w/o	*10*	*20*	*30*	*40*	*50*	*60*	*70*	*80*	*90*
Bear	36,813	35,541	34,519	33,630	32,863	32,419	31,882	31,583	**31,504**	31,553
Blackswan	46,164	44,334	43,235	42,316	41,425	40,635	40,352	39,422	**38,470**	38,688
Camel	36,734	35,496	34,435	33,438	32,964	32,544	32,233	31,821	31,261	**30,760**
Cows	42,889	41,395	40,014	38,635	37,937	37,397	37,081	36,842	36,722	**36,383**
Dog	51,337	50,488	49,627	48,741	48,263	47,744	47,897	48,167	48,044	**44,988**
Elephant	37,370	37,000	36,604	36,190	35,591	35,051	34,877	34,497	34,065	**33,786**
Falmingo	43,213	42,869	42,445	42,059	41,625	41,330	41,060	40,797	40,395	**39,866**
Goat	56,694	55,754	54,742	53,680	52,859	**52,388**	52,446	52,823	53,271	53,352
Mallard- fly	56,054	55,621	55,159	54,669	54,270	54,070	**53,946**	54,030	54,261	54,406
Mallard -water	56,054	55,153	54,458	53,873	53,478	53,160	52,979	**52,794**	52,930	53,018

V. DISCUSSION AND CONCLUSIONS

All the experiments revealed that the addition of an artificial texture to the silhouette image can improve performance. Thus, the proposed method worked well for realistic silhouette images generated from DAVIS 2016 dataset. Moreover, it can be observed that there was no MEPE or the smallest AAE in the w/o column (without adding an artificial texture).

In Table II and Table III, from five animal data (Camel, Cows, Dog, Elephant, and Flamingo) the lowest MEPE and AAE were in subcolumn 90. If it is connected with Table I which revealed the motion type, it can be noticed that five of them had the same type of motion type on walk, deformation and slow. Only Dog was the fast type. This experiment revealed that for the case of relatively slow object movement also little object deformation and adding an artificial texture that is very close to the original texture can improve performance.

For the case of Bear and Blackswan data, the smallest MEPE and AAE values were in the 80 subcolumn, where both motion types were slow. For Bear was the walk movement while Blackswan was the moves forward movement. It was revealed that in the case of slow moving objects without deformation, the addition of artificial texture was still required to improve performance. However, it was close enough to the original texture at about 80% of the original size.

The rest of the three animal data were slightly different from the others. Mallard-fly achieved the lowest MEPE in subcolumn 90, but it achieved the lowest AAE in sub-column 60 where the types of movement were walk, high deformation, and fly. There was an inconsistency between MEPE and AAE, which were usually in the same subcolumn. This was due to a large deformation change, from walking to flying. As for Mallard-water, the lowest MEPE and AAE values were in the 70 subcolumn. In terms of type of movement, it was almost similar to Blackswan's, namely, move forward and slow movements. However, there was additional deformation of the object's motion orientation. From the Mallard-water experiment, it can be observed that where the movement was slow and where the deformation was progressive, it was still necessary to add an artificial texture, but it was a bit far from the original. The last one that was the most different from the other data were the goat data. The lowest MEPE and AAE values fell in sub column 50, with the type of walk movement, deformation and fast. What distinguishes it from the others is primarily the deformation of the object. This was due to the movement of the fur. Afterward, the camera movement was great for following objects. From this experiment, it was revealed that with large object deformation and camera movement, only 50% of the artificial texture was needed to improve performance.

Improved silhouette image performance by adding artificial texture (object-in-object method) can be categorized as follows:

- The addition of an artificial texture as much as 90% of the original texture, for slowly moving objects and simple deformations.

- The addition of 70% or 80% artificial texture of the original texture, for fast moving objects and simple deformations.

- The addition of artificial texture 60% 50% of the original texture, for fast moving objects and large deformations.

This study attempted to compare the ground truth from real video in the form of RGB images and optical flow which was formed from modified image silhouette. The contribution of the proposed method to modify the silhouette by adding an artificial texture can actually improve its performance. The performance measurements used were MEPE and AAE. The limitation of this study was it was only for a single segmented object. For the case of multiple objects, the object of occlusion was not included in this study and it can be studied for further research.

REFERENCES

[1] N. O'Mahony et al., "Deep Learning vs. Traditional Computer Vision," in Advances in Computer Vision, vol. 943, K. Arai and S. Kapoor, Eds. Cham: Springer International Publishing, 2020, pp. 128–144.

[2] M.-N. Chapel and T. Bouwmans, "Moving objects detection with a moving camera: A comprehensive review," Comput. Sci. Rev., vol. 38, p. 100310, Nov. 2020, doi: 10.1016/j.cosrev.2020.100310.

[3] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object Detection With Deep Learning: A Review," IEEE Trans. Neural Netw. Learn. Syst., vol. 30, no. 11, pp. 3212–3232, Nov. 2019, doi: 10.1109/TNNLS.2018.2876865.

[4] Y. Zhang, J. Zheng, C. Zhang, and B. Li, "An effective motion object detection method using optical flow estimation under a moving camera," J. Vis. Commun. Image Represent., vol. 55, pp. 215–228, Aug. 2018, doi: 10.1016/j.jvcir.2018.06.006.

[5] D. H. Ye, J. Li, Q. Chen, J. Wachs, and C. Bouman, "Deep Learning for Moving Object Detection and Tracking from a Single Camera in Unmanned Aerial Vehicles (UAVs)," Electron. Imaging, vol. 2018, no. 10, pp. 466-1-466–6, Jan. 2018, doi: 10.2352/ISSN.2470-1173.2018.10.IMAWM-466.

[6] J. Huang, W. Zou, J. Zhu, and Z. Zhu, "Optical Flow Based Real-time Moving Object Detection in Unconstrained Scenes," ArXiv180704890 Cs, Jul. 2018, Accessed: Jan. 03, 2021. [Online]. Available: http://arxiv.org/abs/1807.04890.

[7] M. Bonneau, J.-A. Vayssade, W. Troupe, and R. Arquet, "Outdoor animal tracking combining neural network and time-lapse cameras," Comput. Electron. Agric., vol. 168, p. 105150, Jan. 2020, doi: 10.1016/j.compag.2019.105150.

[8] M. Zeppelzauer, "Automated detection of elephants in wildlife video," EURASIP J. Image Video Process., vol. 2013, no. 1, p. 46, Dec. 2013, doi: 10.1186/1687-5281-2013-46.

[9] L. Zhang, H. Gray, X. Ye, L. Collins, and N. Allinson, "Automatic Individual Pig Detection and Tracking in Pig Farms," Sensors, vol. 19, no. 5, p. 1188, Mar. 2019, doi: 10.3390/s19051188.

[10] W. J. Hans, V. Sherlin, and N. Venkateswaran, "On-Road Deer Detection for Advanced Driver Assistance using Convolutional Neural Network," Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 4, 2020, doi: 10.14569/IJACSA.2020.0110499.

[11] B. K. P. Horn and B. G. Schunck, "Determining optical flow," Artif. Intell., vol. 17, no. 1–3, pp. 185–203, Aug. 1981, doi: 10.1016/0004-3702(81)90024-2.

[12] B. Purnama et al., "On the Problem about Optical Flow Prediction for Silhouette Image:," 2019, pp. 289–293, doi: 10.5220/00075723 02890293.

[13] M. A. Arredondo, K. Lebart, and D. Lane, "Optical flow using textures," Pattern Recognit. Lett., vol. 25, no. 4, pp. 449–457, Mar. 2004, doi: 10.1016/j.patrec.2003.11.007.

[14] M. Andalibi, Lawrence. L. Hoberock, and H. Mohamadipanah, "Effects of texture addition on optical flow performance in images with poor texture," Image Vis. Comput., vol. 40, pp. 1–15, Aug. 2015, doi: 10.1016/j.imavis.2015.04.008.

[15] F. Perazzi, J. Pont-Tuset, B. McWilliams, L. Van Gool, M. Gross, and A. Sorkine-Hornung, "A Benchmark Dataset and Evaluation Methodology for Video Object Segmentation," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, Jun. 2016, pp. 724–732, doi: 10.1109/CVPR.2016.85.

[16] T. Brox, "Optical Flow: Traditional Approaches," in Computer Vision, Cham: Springer International Publishing, 2020, pp. 1–5.

[17] M. J. Black and P. Anandan, "The Robust Estimation of Multiple Motions: Parametric and Piecewise-Smooth Flow Fields," Comput. Vis. Image Underst., vol. 63, no. 1, pp. 75–104, Jan. 1996, doi: 10.1006/cviu.1996.0006.

[18] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in In IJCAI81, 1981, pp. 674–679.

[19] C. Zach, T. Pock, and H. Bischof, "A Duality Based Approach for Realtime TV-L 1 Optical Flow," in Pattern Recognition, vol. 4713, F. A. Hamprecht, C. Schnörr, and B. Jähne, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 214–223.

[20] P. Fischer et al., "FlowNet: Learning Optical Flow with Convolutional Networks," ArXiv150406852 Cs, May 2015, Accessed: Jan. 03, 2021. [Online]. Available: http://arxiv.org/abs/1504.06852.

[21] P. Weinzaepfel, J. Revaud, Z. Harchaoui, and C. Schmid, "DeepFlow: Large displacement optical flow with deep matching," Sydney, Australia, Dec. 2013, [Online]. Available: http://hal.inria.fr/hal-00873592.

[22] J. Revaud, P. Weinzaepfel, Z. Harchaoui, and C. Schmid, "EpicFlow: Edge-Preserving Interpolation of Correspondences for Optical Flow," 2015.

[23] J. Revaud, P. Weinzaepfel, Z. Harchaoui, and C. Schmid, "DeepMatching: Hierarchical Deformable Dense Matching," ArXiv150607656 Cs, Oct. 2015, Accessed: Jan. 01, 2021. [Online]. Available: http://arxiv.org/abs/1506.07656.

[24] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, and R. Szeliski, "A Database and Evaluation Methodology for Optical Flow," Int. J. Comput. Vis., vol. 92, no. 1, pp. 1–31, Mar. 2011, doi: 10.1007/s11263-010-0390-2.

[25] D. J. Butler, J. Wulff, G. B. Stanley, and M. J. Black, "A Naturalistic Open Source Movie for Optical Flow Evaluation," in Computer Vision – ECCV 2012, vol. 7577, A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 611–625.

[26] T. Pock, M. Urschler, C. Zach, R. Beichel, and H. Bischof, "A Duality Based Algorithm for TV-L 1-Optical-Flow Image Registration," in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007, vol. 4792, N. Ayache, S. Ourselin, and A. Maeder, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 511–518.

[27] J. Sánchez Pérez, E. Meinhardt-Llopis, and G. Facciolo, "TV-L1 Optical Flow Estimation," Image Process. Line, vol. 3, pp. 137–150, Jul. 2013, doi: 10.5201/ipol.2013.26.