International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

To Design and Analyze an Method for Unsupervised Recurrent All-Pairs Optical Flow Field Transform Algorithm

Ms. Rajashree Revaji Shinde
Research Scholarship
Electronics & Communication Engineering
Dr. A. P. J. Abdul Kalam University
Indore, India
rajashreeshinde177@gmail.com

Dr. Santosh Pawar
Research Supervisor
Electronics & Communication Engineering
Dr. A. P. J. Abdul Kalam University
Indore, India
spawarrkdf@gmail.com

Abstract—The SMURF, a powerful technique for optical flow unsupervised learning that closes the gap with supervised methods, exhibits good cross-dataset generalization, and even allows for "zero-shot" depth estimation. SMURF introduces several important improvements: full-image warping for learning to predict out-of-frame motion, multi-frame self-supervision for improved flow estimates in occluded regions, and most importantly, modifications to the unsupervised losses and data augmentation that allow the RAFT architecture to operate in an unsupervised setting. These developments, in our opinion, take unsupervised optical flow one step closer to becoming truly practical, enabling optical flow models trained on unlabeled videos to deliver accurate pixel-matching in areas where labeled data is lacking.

Keywords- Self-supervised learning, Multi-frame, Full-Image Warping

I. INTRODUCTION

Multi-frame unsupervised self-learning RAFT use the entire image Warping is a technique for determining the solo optical stream that exceeds administered approaches like as PWC-Net and FlowNet2, and works on the cutting edge of all benchmarks from 36% to 40% (compared with the previous best UFlow technique). Our method for composing improvements from the administered optical stream, called the RAFT model, integrates novel concepts for solo learning, including mindful self-management for loss of arrangement, a strategy for handling off-outline movement, and a learning approach. For inductions utilizing credible multi-outline video information, only two casings are needed. [1].

Optical stream shows a close-up pixel-by-pixel correspondence between two images, identifying each pixel in the original image with its corresponding location in the second image. The vector field demonstrates the general places of the pixels of the obvious development or "stream" between the two pictures. Assessing this stream field is a key issue in PC vision, and any progression in stream assessment is helpful for some downstream exercises, for example, visual odometry, Multi view profundity assessment and video object observing.

Traditional techniques make optical transition assessment as an improvement issue. They produce, for a given pair of pictures, a stream field that boosts the perfection and closeness of the coordinated with pixels. Maybe, late directed learning approaches train profound neural organizations to assess optical transition from instances of picture sets commented on to the truth of the earth. Since it is incredibly hard to detect the progression of truth with genuine pictures, administered learning is normally restricted to engineered information [3]. Albeit these strategies have given brilliant outcomes in the preparation field, it is hard to sum up when the hole between the objective space and the manufactured preparing information is excessively huge.

Since unsupervised learning enables the training of optical flow patterns from unlabeled video on any domain, it is a promising approach to solving this challenge. The unsupervised technique trains the same neural networks as a supervised approach but optimizes them with objectives like uniformity and photometric similarity of classical methods. It does this by fusing the concepts of supervised learning and classical approaches. [2]. Unsupervised approaches, in contrast to these traditional techniques, optimize for the full

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

training set rather than just a pair of photos. Given that unsupervised optical flow draws influence from both classical and supervised learning approaches, a perfect combination of novel concepts and insights from these two domains might lead to substantial advancements. We exactly that in this text, and we add the following three things:

1. To adequately fine-tune this model for unsupervised learning, we merge the state-of-the-art supervised model, RAFT, with unsupervised learning and make significant modifications to the data gain and loss functions. [1].
2. To compute unsupervised losses, we utilize the entire image after learning the unsupervised image bar. This method, which we refer to as full-image heating, enhances the flow quality in the vicinity of the image's edges.
3. For Unsupervised Recurrent All-pairs Field Transform of Optical Flow, we design and analyze a new algorithm SMRUF-Combining Augmentation with geometric transformation. Moreover, we leverage a traditional multi-frame stream tuning technique to produce better labels for self-control from multi-frame input. By joining no more than two frames for inference, this method enhances performance, particularly in closed regions. [3].

The photometric misfortune in the Full-Image warping is crucial for unattended Optical motion assessment is typically limited to motion vectors that remain inside the image's edge because, when the vectors point outside the frame, there are no pixels to consider for determining their photometric appearance. To overcome this limitation, we calculate the stream range from a cropped version of the I1 images and, in addition, I2 for the complete non-cryptographic image, misshaping it with the evaluated stream V1, before calculating the photometric error. [4]. Since we didn't score any longer these transition vectors that move outside the casing give a picture as they happen as a learning signal for the model at this point. We utilize full picture twisting for all datasets with the exception of Flying Seats, where we tracked down that the presentation of the little pictures is as of now hurt[7].

II. EXISTING SYSTEM

Existing System

A. Stone et al, SMURF: Self-Teaching Multi-Frame Unsupervised RAFT with Full-Image Warping present SMURF, an unsupervised optical flow learning method that outperforms many supervised methods, such as PWC-Net and FlowNet2, and improves the state of the art by 36% to 40% across all benchmarks. Our method combines novel unsupervised learning concepts, like a sequence-aware self-supervision loss, a way to deal with out-of-frame motion, and a strategy to effectively learn from multi-frame video data while requiring only two frames for inference, with architecture enhancements from supervised optical flow, i.e., the RAFT model. Traditional and supervised learning approaches serve as models for unsupervised optical flow; by appropriately combining new ideas with what we know from these two fields, we can achieve substantial advancements. In this paper, we do precisely that, and we make the following three contributions:

1. We combine the best supervised model currently available, RAFT, with unsupervised learning and make significant modifications to the loss functions and data augmentation in order to correctly regularize this model for unsupervised learning..
2. We perform unsupervised learning on picture crops and compute unsupervised losses using the entire image. This method, which we call full-image warping, enhances flow quality close to the limits of the image.
3. To get more accurate labels for self-supervision from multi-frame input, we utilize a classical technique for multi-frame flow refinement [19]. With this method, performance is enhanced, particularly in obscured areas, and inference takes no more than two frames. [1].

F. Aleotti et al, Reversing the cycle: Self-supervised deep stereo through enhanced monocular distillation in The author of this work notes the noteworthy generalization abilities that address domain shift problems. Self-supervised learning methods are quickly bridging the gap with supervised approaches in a number of sectors. Both monocular and stereo depth estimates exhibit this property, with the latter often acting as a dependable source of self-supervision for the former. Alternatively, to reduce traditional stereo distortions, we propose a novel self-supervised paradigm that flips the relationship between the two. In order to train deep stereo networks, we purposefully condense knowledge using a monocular completion network. This architecture uses single-image cues and a small number of sparse points generated by traditional stereo algorithms to estimate dense yet accurate disparity maps through a consensus mechanism across repeated estimations. We carefully evaluate the impact [2].

B. Lucas, A stereo vision use of an iterative image registration approach Applications for image registration in computer vision include motion analysis, pattern recognition, and picture matching for stereo vision. Unfortunately, current picture registration methods are often expensive. Furthermore, they frequently fall short in handling picture aberrations like rotation. We introduce a new image registration method in this research that finds the best match by using spatial intensity gradient information to guide the search. This method finds the greatest match between two photos with considerably fewer image comparisons than methods that look at potential registration places in a predetermined order because it considers additional information about the images.. Our method capitalizes on the fact that the two images are often already roughly registered in many applications. This method can be expanded to handle any kind of linear image distortion, including rotation. After that, we go over a stereo vision system that makes use of this registration process and offer some more research directions for effectively applying this technique to the interpretation of stereo images. [3].

D. Maurer, Proflow: Acquiring the ability to forecast optical flow The author of this work addresses both issues. Rather than presuming a moving camera with associated rigidity limitations, we introduce a unique optical flow technique that leverages a convolutional neural network (CNN) to develop appropriate motion models. Our contributions in this context are fourfold: (i) Our method learns the models live, i.e., during the estimation, in contrast to existing approaches that train a network beforehand. (ii) Moreover, our models are trained utilizing initial flow estimates of the actual sequence, rather than depending on possibly inappropriate ground truth data sets, e.g., data sets that solely contain motion patterns that differ from the happening motion. The benefit of such an unsupervised training1 is that suitable models can be learned for every sequence. (iii) Lastly, our method trains a single model for every frame in every sequence, in addition to one model per sequence. Apparently, this leads to a significant degree of flexibility in terms of changing the substance of the scenario. (iv) Lastly, the acquired models exhibit spatial variation, meaning they rely on their location. Consequently, the issue of autonomously moving objects is addressed. It is finally possible for us to forecast the forward flow from the backward flow after learning such specialized motion models. This makes it feasible to enhance the estimation in areas—such as obstructed regions—where the forward flow is unavailable. Experiments clearly show the advantages of our new approach. Not only do they consistently outperform an unpredicted baseline method, but they also get excellent outcomes across the board for all significant benchmarks.

When it comes to optical flow estimate, temporal coherence is a useful information source. Finding an appropriate motion model to make use of this data is a difficult task, though. In this paper, we propose a convolutional neural network (CNN) based unsupervised online learning strategy that estimates a motion model for each frame separately. These trained models, which establish a relationship between forward and backward motion, not only enable the inference of important motion information from the backward flow, but they also aid in enhancing performance at occlusions, where a trustworthy forecast is very helpful. Furthermore, our acquired models exhibit spatial variation, hence enabling the estimation of non-rigid motion according to construction. Consequently, this makes it possible to get around a significant drawback of more recent rigidity-based methods that aim to enhance the estimation by adding more stereo/SfM constraints. Tests show that our novel strategy is beneficial. In comparison to a baseline without prediction, they not only consistently demonstrate improvements of up to 27% for all major benchmarks (KITTI 2012, KITTI 2015, and MPI Sintel), but they also achieve the best results for the MPI Sintel benchmark, which is one of the three benchmarks with the greatest amount of non-rigid motion. [4].

N. Mayer et al, Three such datasets are presented in this huge dataset to train convolutional networks for disparity, optical flow, and scene flow estimates. The datasets were created with Blender3, an open source 3D creative suite, customized. Our work is conceptually comparable to the Sintel benchmark. Unlike Sintel, our dataset offers ground truth for scene flow and is big enough to make convolutional network training easier. Specifically, it contains ground truth for bidirectional disparity, motion boundaries, bidirectional optical flow and disparity change, and stereo color pictures and object segmentation. Moreover, our collection includes RGBD data, meaning that the complete camera calibration and 3D point coordinates are accessible. We are unable to fully utilize this dataset in a single publication, but we have already shown how it may be used in a number of scenarios when combined with convolutional network training. We train a network for disparity estimation that produces competitive results on prior benchmarks, particularly when compared to real-time approaches. In conclusion, we also provide a network for estimating scene flow and offer the first quantitative data on the entire scene flow on a sizable enough test set. [5].

S. Meister et al, The era of end-to-end deep learning is characterized by Unflow: Unsupervised learning of optical flow with a bidirectional census loss. A lot of the advancements in computer vision are driven by vast amounts of labeled data. However, it is challenging to get dense perpixel ground truth for real sceneries in the optical flow environment, which makes such data uncommon. Consequently, the supervision of modern end-to-end convolutional networks for optical flow relies on synthetic datasets, but the problem of domain mismatch between training and test situations still exists. Motivated by traditional energy-based optical flow techniques, we create an unsupervised loss that avoids the requirement for ground truth flow by utilizing occlusion-aware bidirectional flow estimates in conjunction with the robust census transform. Our unsupervised method significantly beats earlier unsupervised deep networks on the KITTI benchmarks, and is even more accurate than comparable supervised approaches trained solely on synthetic datasets. Our method provides competitive optical flow accuracy on the KITTI 2012 and 2015 benchmarks by optionally fine-tuning on the KITTI training data. This also makes generic pre-training of supervised networks possible for datasets with limited ground truth. [6].

M. Menze et al, Cooperative 3D vehicle and scene flow estimation The authors of this work expand on the methodology of Menze and Geiger (2015) while going one step further: Rather than breaking down the scene into a collection of separately moving regions with a shared rigid motion, we break down the scene into three-dimensional objects and also model their shape and position in three dimensions in addition to their rigid motion. In order to do this, we apply a deformable 3D vehicle model to the scene flow estimation procedure. More precisely, we take advantage of the Eigenspace-based representation of (Zia et al., 2011), which has been applied to pose estimation from a single image in the past. Our model simultaneously infers a rich 3D scene flow field, the number of vehicles, and their form and posture parameters given two stereo pairs as input. Energy minimization on a conditional random field that encourages projected object hypotheses to agree with the estimated velocity and depth is how the problem is formalized. Fig. 1 presents a representative result that includes the outcome of model-based reconstruction along with scene flow estimates projected to disparity and optical flow. [7].

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

A. Ranjan, An other method that incorporates the best features of both approaches is optical flow estimation using a spatial pyramid network author. Many years of flow research have resulted in well-designed systems and useful concepts. However, these techniques have several assumptions that restrict how well they work. In order to achieve the following objectives: 1) improve performance over current neural networks and the classical methods upon which our work is based; 2) achieve real-time flow estimates with accuracy better than the much slower classical methods; and 3) reduce memory requirements to make flow more practical for embedded, robotic, and mobile applications. As a result, here we apply machine learning to address the weak points while maintaining the engineered architecture. Two challenges must be resolved in order to compute flow. The first involves finding long-range correlations, and the second involves finding exact motion bounds and intricate sub-pixel optical flow. FlowNet, the prior neural network technique, aims to simultaneously learn both of these. On the other hand, we solve the former using current techniques while addressing the latter utilizing deep learning. [8].

A. Ranjan et al, Competitive cooperation: Unsupervised joint learning of motion segmentation, camera motion, optical flow, and depth The author discusses the unsupervised learning of multiple related low-level vision problems, including optical flow, camera motion estimation, single view depth prediction, and segmenting a video into moving and static areas. Our main finding is that geometric limitations establish a connection between these four basic vision issues. Because the solutions can support one another, learning to tackle them together simplifies the issue. By specifically utilizing geometry and dividing the scene into stationary and mobile areas, we surpass earlier research. In order to achieve this, we provide Competitive Collaboration, a framework that makes it easier to coordinate the training of several specialized neural networks to tackle challenging issues. Similar to expectation-maximization, competitive collaboration makes use of neural networks that collaborate with a moderator to determine whether pixels are stationary or moving independently, as well as compete to explain pixels that relate to static or moving regions. Our unique approach unifies these issues into a single framework while reasoning about the camera motion, depth of the static scene structure, moving object optical flow, and scene segmentation into moving objects and the static background. Our model attains state-of-the-art performance among combined unsupervised approaches on all sub-problems, having been trained without any supervision. [9].

Z. Ren et al, Deep learning without supervision for optical flow estimates According to this research, a deep network trained with our unsupervised scheme performs on par with fully supervised training. We think this is mostly because of our end-to-end training, which enables the network to use context data over a wide area to infer local motion. In light of this, the following is a summary of our primary contributions. 1) To the best of our knowledge, this is among the first studies on deep neural networks for autonomously learning optical flow. Our work differs significantly from the most advanced learning-free techniques. DeepFlow (Weinzaepfel et al. 2013) or EpicFlow (Revaud et al. 2015), and the supervised deep learning approach FlowNet (Fischer et al. 2015) and DispNet (Mayer et al. 2016). 2) For end-to-end unsupervised learning for optical flow estimation, we present a novel optical flow network that is similar to the Spatial Transformer Network pipeline (Jaderberg et al. 2015). It leverages the loss function used in variational approaches (Brox et al. 2004) without supervision. Even though the improvements are not great, we think this is a worthwhile area to investigate further. 3) Lastly, we will make a public Caffe (Jia et al. 2014) implementation available to facilitate comparison and more innovation. [10].

F. Steinbrucker et al, Computation of large displacement optical flow without warping In order to estimate large displacement optical flow without the requirement for warping techniques, the author suggested a unique algorithm. We deconstruct the original non-convex functional into a functional that can be reduced by alternating two globally optimum steps using a quadratic relaxation approach. The approach only switches between a convex optimization that considers the smoothness requirement and a thorough search with respect to the non-convex (but point-wise) data term. As a result, the flow estimating procedure may be broken down into two phases: discontinuity preserving smoothing and finding suitable correspondents. The suggested method, in contrast to warping approaches, can naturally use arbitrary data terms, such as norms on color values or local patches, and non-convex, non-differentiable terms. The suggested quadratic decoupling approach, in contrast to cutting-edge warping schemes, enables the computation of flow fields that accurately fit small-scale structures over huge displacements, as demonstrated by the author's numerous experiments. [11].

III. Proposed Work

Model RAFT Our initial ablation explores how situational awareness, as opposed to the UFlow art technique and its PWC model RAFT, can lead to dimensions improvement. It is no accident that the approach improves but rather lowers performance when the model is replaced with unsupervised learning, as demonstrated by the results in Table 4[13]. We have identified and developed the methods described here to improve unsupervised learning with RAFT through considerable experimentation [14]. Perhaps as a result of the more constrained architecture, earnings from these strategies are substantially lower with the PWC model. The following are key elements in the design, development, and analysis of "An algorithm SMRUF-Combining Augmentation with geometric transformation for Unsupervised Recurrent All-pairs Field Transform of Optical flow." [16],

- RAFT Model
- Multi-Frame Self-Supervision
- Image Warping
- Unsupervised RAFT

3.1 Architecture Of Proposed System

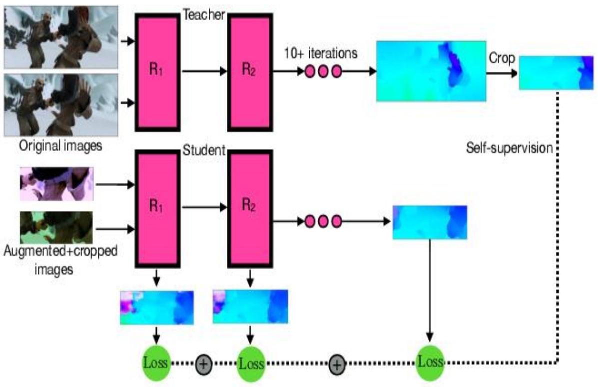

Fig. 1. Self-supervision with sequence loss and augmentation.

We employ a single model for the roles of "teacher" and "student." We use the model on complete, non-augmented photos in our role as the teacher. The model only sees a trimmed and enhanced version of the same photographs while acting as a student. After cropping, the teacher's final output is utilized to monitor the students' predictions throughout all of their iterations (including applying smoothness and photometric losses). The main advantages of this self-supervision method are as follows:

(1) The model becomes capable of ignoring photometric augmentations.
(2) The model gains the ability to forecast more accurately at the borders and in obscured portions of the image.
(3) Early iterations of the recurrent model learn from the output at the final iteration images.

Next, this cost volume is periodically fed into a recurrent network, which builds and refines a flow field forecast repeatedly. To enable training with very tiny batch sizes, we substitute batch normalization with instance normalization [15], which is the only architectural change we make to RAFT. In order to accommodate the model and the more complex unsupervised training steps into memory, the batch size had to be lowered. But more importantly, we found that in order to effectively leverage RAFT's learning capacity, significant modifications need to be made to the unsupervised learning approach.
[16].

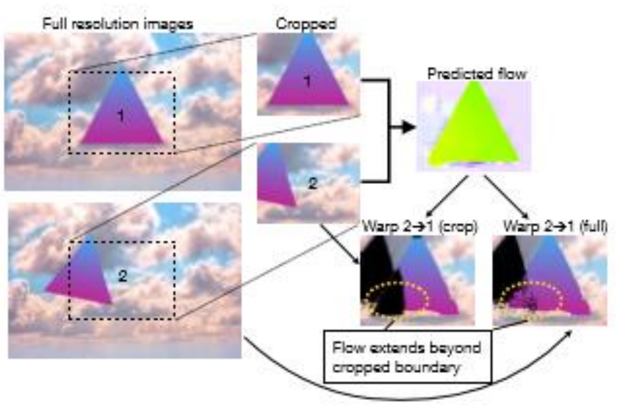

Fig. 2. Full image Warping.

Only flow vectors that stay inside the image frame usually have the photometric loss, which is essential for unsupervised optical flow estimation because they don't have any pixels to compare their photometric appearance with. This limitation is overcome by computing the flow field using a cropped version of the images I1 and I2, warping it with the estimated flow V1, and referencing the full, uncropped image I2 (see to Figure 2). As a result of their removal from the occluded classification, these flow vectors outside the image frame serve as a learning signal for the model [14]. We use full-image warping for all datasets, with the exception of Flying Chairs, where we found that cropping the already small photographs was detrimental to performance.

IV. PROPOSED SYSTEM

The study of hyper parameters is based on UFlow with slight modifications based on further hyper parameter search. KITTI dataset is required for the study.

4.1 Dataset Details

In compliance with the guidelines published in the literature, our model may be trained and tested using the following optical flow datasets: Flying Chairs, Sintel, and KITTI 2015. We pretrain on Flying Chairs before we fine-tune on Sintel or KITTI. Similar to UFlow, we did not observe any benefits from pretraining on larger amounts of non-domain data, such Flying Objects. In our approach, not a single training method uses any ground truth labels. We use the "training" section of Flying Chairs for training, and we divide the Sintel dataset into its standard train and test portions. We practice using the split with KITTI that we used with the multi-view extension in previous work [12]: Correctly analyze these models based on the test set. One model is trained on the multiview extension of the training set, and the second model is trained on the test extension. After training on the label-free "test" portion of the dataset and evaluating on the training set, we report metrics for ablations. After training only on the training component, we report results for final benchmark values[17]. For our benchmark result on Sintel, we train on a 50% blend of the KITTI and Sintel multi-frame self-supervision labels. We also report error rates ("ER") for all datasets using KITTI. If a prediction's EPE exceeds 3 pixels or 5% of the true flow vector's length, it is considered wrong. We typically compute these measures for all pixels, with the exception of "EPE (noc)," which only considers non-occluded pixels.

4.2 Data Augmentation

Wide-ranging Data Enhancement We employ the same augmentation as supervised RAFT, which is substantially stronger than what has been used in unsupervised optical flow up to this point (with the exception of the latest ARFlow), to regularize RAFT. Each image receives a random eraser augmentation, which randomly removes portions of the image. We also arbitrarily alter the image's color, brightness, saturation, stretching, scaling, random cropping, and random left/right and up/down flipping[18]. All augmentations are applied to the model inputs, but not to the images that are used to compute the smoothness and photometric losses[19] One advantage of using unaugmented images for the self-generated labels for self-supervision is that the model is trained to ignore these augmentations.

5. Outcome of Proposed System

The SMURF is a powerful unsupervised optical flow learning technique that closes the gap with supervised methods and exhibits good cross-dataset generalization, including "zero-shot" depth estimation. Significant enhancements are introduced by SMURF, the most notable of which are

1. Adjustments to the data augmentation and unsupervised losses that enable the RAFT architecture to function in an unsupervised environment.
2. Whole-image warping to anticipate out-of-frame movements, and
3. Multi-frame autonomous monitoring to enhance flow estimations in obscured areas.

We believe that our contributions are a step towards achieving genuine feasibility for unsupervised optical flow, allowing optical flow models trained on unlabeled videos to do high-quality pixel-matching in domains without labeled input

- To create and evaluate an algorithm for the Unsupervised Recurrent All-pairs Field Transform of Optical Flow using geometric transformation and SMRUF-Combining Augmentation.
- To research and evaluate methods for self-supervision modifications.
- To investigate and assess multi-frame self-supervision
- Researching and evaluating the Self-supervision with sequence loss and augmentation;
- Analyzing the RAFT Model

6. Conclusion

In this paper, In order to implement worldwide popular geometry restrictions for unsupervised optical flow learning, we have put forth some practical solutions. We used the low-rank constraint on a stationary scene in order to regularize a globally stiff structure. We suggested using the union-of-subspaces constraint to general dynamic sceneries (multi-body or deformable). Tests conducted on several benchmarking datasets have demonstrated the effectiveness and dominance of our techniques in relation to cutting-edge (unsupervised) deep flow approaches. We intend to investigate multi-frame extension in the future, which is the application of geometric restrictions over a number of frames.

References

[1] A. Stone, D. Maurer, A. Ayvaci, A. Angelova, R. Jonschkowski, “SMURF: Self-Teaching Multi-Frame Unsupervised RAFT with Full-Image Warping”, IEEE Conference on Computer Vision and Pattern Recognition), Page no.3887-3896, Honolulu, Hawaii, USA, July, 2021.

[2] F. Aleotti, F. Tosi, L. Zhang, M. Poggi, and S. Mattoccia, “Reversing the cycle: Self-supervised deep stereo through enhanced monocular distillation”, in ECCV, Vol.3, Issue 3, Page no. 1-18, Aug. 2020.

[3] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, “High accuracy optical flow estimation based on a theory for warping” in Eurppean Conference on Computer Vision, Page no.25-36, 2004.

[4] D. Butler, J. Wulff, G. Stanley, and M. Black, “A naturalistic open source movie for optical flow evaluation” in Proceedings of the 12th European conference on Computer Vision - Volume Part VI, Page no. 611-625, October 2012.

[5] Q. Chen and V. Koltun, “Full flow: Optical flow estimation by global optimization over regular grids”, in CVPR, Page no. 4706-4714, 2016.

[6] A. Dosovitskiy, P. Fischer, E. Ilg, P. Hausser, C. Hazırbas¸, V. Golkov, P. Smagt, D. Cremers, and T. Brox, “Flownet: Learning optical flow with convolutional networks”, in ICCV, Page no.1-13, May 2015.

[7] H. Hirschmuller, “Stereo processing by semi-global matching and mutual information”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Page no. 328–341, February 2008.

[8] K. Berthold and B. Schunck, “Determining optical flow”, AI, 1981. Artificial Intelligence, Volume 17, Issues 1–3, Pages 185-203, August 1981.

[9] E. Ilg, N. Mayer, T. Saikia, M. Keuper, A. Dosovitskiy, and T. Brox, “Flownet 2.0: Evolution of optical flow estimation with deep networks”, in CVPR, Page no. 2462-2470, 2017.

[10] W. Im, T. Kim, and S.-E. Yoon, “Unsupervised learning of optical flow with deep feature similarity” in ECCV, Page no. 2463-2473, 2020.

[11] J. Janai, F. Guney, A. Ranjan, M. J. Black, and A. Geiger, “Unsupervised learning of multi-frame optical flow with occlusions” in ECCV, Page no. 1235-1243, 2018.

[12] R. Jonschkowski, A. Stone, J. Barron, A. Gordon, K. Konolige, and A. Angelova, “What matters in unsupervised optical flow”, in ECCV, Page no. 4850-4861, 2020.

[13] D. Kingma and J. Adam, “A method for stochastic optimization”, in ICLR, Page no. 2243-2252, 2015.

[14] L. Liu, J. Zhang, R. He, Y. Liu, Y. Wang, Y. Tai, D. Luo, C. Wang, J. Li, and F. Huang, “Learning by analogy: Reliable supervision from transformations for unsupervised optical flow estimation”, in 4th Conference on Robot Learning, Page no. 1-13, 2020.

[15] P. Liu, I. King, M. Lyu, and J. Xu, “Flow2Stereo: Effective self-supervised learning of optical flow and stereo matching”, in CVPR, Page no. 456-467, 2020.

[16] D. S. Thosar and M. Singh, "A Review on Advanced Graphical Authentication to Resist Shoulder Surfing Attack," 2018 International Conference on Advanced Computation and Telecommunication (ICACAT), Bhopal, India, 2018, pp. 1-3, doi: 10.1109/ICACAT.2018.8933699.

[17] P. Liu, I. King, M. Lyu, and J. Xu, "DDFlow: Learning optical flow with unlabeled data distillation", AAAI, Page No.1-13, Hawaii USA, Feb 2019.

[18] Devidas S. Thosar*, Dr. Nisarg Gandhewar. (2022). An advanced image authentication using passimage algorithm to resist shoulder surfing attack. Computer Integrated Manufacturing Systems, 28(10), 52–59.

[19] P. Liu, M. R. Lyu, I. King, and J. Xu, "Selflow: Self-supervised learning of optical flow", Conference on Computer Vision and Pattern Recognition, Page no. 2423-2436, 2019