IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 2, 2020, accepted January 20, 2020, date of publication January 22, 2020, date of current version January 30, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2968763

Research on Robust Visual Tracker Based on Multi-Cue Correlation Particle Filters

YUQI XIAO AND DIFU PAN
School of Traffic and Transportation Engineering, Central South University, Changsha 410075, China
Corresponding author: Yuqi Xiao (xiaoyuqi@csu.edu.cn)

ABSTRACT For the problem of robust visual tracking in various complex tracking scenarios, a multi-cue correlation particle filter (CPF for short) visual tracker supervised by population convergence is proposed. By combining the advantages of particle filter and correlation filter, the CPF tracker gains better robustness, computational efficiency and stability for visual tracking. Meanwhile, to solve the problem of sample diversity in traditional CPF tracker, a genetic operating algorithm supervised by population convergence is proposed and introduced to the resampling process of CPF. Then considering that a single kind of feature weakens the tracking efficiency and robustness of our tracker, we propose to combine different types of features including Harris feature, HOG feature and SIFT feature based on fuzzy control theory to form a multi-cue CPF tracker (SPC-MCCPF for short). Multiple experiments on the OTB2015 and VOT2018 datasets prove that our tracker is quite effective in dealing with various challenging tracking problems.

INDEX TERMS Computer vision, target tracking, feature extraction, correlation particle filter.

I. INTRODUCTION

Visual target tracking has a wide range of applications in hotspots such as traffic monitoring, automatic driving, human-machine interaction, violence identification, suspect inspection and other fields [1]–[3]. Although great progress has been made in recent years, there still remain various challenging problems such as low resolution, deformation and occlusion [4]. This study therefore focuses on dealing with these challenges and develops a robust and efficient tracking method.

Recently, particle filter has been applied in the field of visual tracking because it is suitable for solving nonlinear problems like nonlinear target motion and is flexible in combination with various kinds of target representations. Ref. Concha *et al.* [5] presented an adequate performance analysis of the particle filtering (PF for short) algorithm for a computationally intensive 3D multi-view visual tracking problem. Zhang *et al.* [6] proposed a hybrid positioning method based on particle swarm optimization algorithm and particle filter. This method is mainly used for positioning tasks with a priori environment map. This research obtains a robust positioning method based on PF, which can work in symmetrical environments. Wang *et al.* [7] developed a new tracking method, in which target tracking is achieved using Huber loss regularization and subspace learning under the framework of particle filtering. And Huber loss function is employed to model the error between candidates and templates to achieve robust tracking. Mondal [8] advanced a discriminative appearance model based on online probabilistic neural networks and an effective observation model based on confidence scores. This updating mechanism can adapt to changes in appearance during target tracking and enables the tracker to efficiently track fully occluded object. Wang *et al.* [9] presented a low-rank sparse tracking algorithm based on PF. This algorithm uses low rank constraints and the underlying sparse to jointly learn the features of local dense scale-invariant feature transformation corresponding to the candidate samples. And it performs great in terms of accuracy. However, traditional particle filters still haven't addressed the issue of particle degeneracy and sample dilution, which limit the tracking efficiency and accuracy.

Meanwhile, correlation filters (CF for short) have also been applied in numerous visual tracking applications [10], [11]. As stated in the convolution theorem, the correlation in time domain corresponds to the multiplication of the elements in the Fourier domain. Therefore, the basic idea of CF is to calculate the correlation in the Fourier domain, avoiding the time-consuming convolution operation. Due to the high computational efficiency, CF has attracted considerable attention

The associate editor coordinating the review of this manuscript and approving it for publication was Juan A. Lara.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

in visual tracking [12]. Yan *et al.* [13] advanced a tracking approach based on multi-channel HOG features, which combines correlated filter tracking response and improved color histogram tracking response to achieve high precision tracking of small targets. Danelljan *et al.* [14] presented the multi-scale correlation filter (DSST) algorithm with HOG features to solve the problem of scale variation in the tracking process. Sui *et al.* [15] presented a new visual tracking algorithm for correlation filter learning based on the intensity of the peak of the correlation response, which effectively enhances the ability of identifying the peak of the correlation response and achieve better recognition performance. Tang *et al.* [16] introduced the multi-kernel learning (MKL for short) into KCF and uses its upper bound to reconstruct the MKL version of the CF objective function, significantly reducing the negative mutual interference between different kernels. Sun *et al.* [17] raised a robust region-of-interest merged CF tracking algorithm, which used a novel joint training formula and derived the Fourier solvers for optimizing model training. Huang *et al.* [18] advanced a tracking approach based on discriminative correlation filtering (DCF for short) and fused colour features. This algorithm consists of compressed colour name features and histogram oriented gradient features based on opponent colour space. Based on the fused colour features, two different DCF are used to estimate the scale and translation of the target, respectively. However, traditional CF relies heavily on the maximum response value of the response map and becomes unreliable when the map gets fuzzy.

In order to solve the above problems, this paper introduces the resampling progress of PF to provide enough effective target candidates for the CF tracker. In addition, to deal with the problem of sample diversity of particle filter, we propose the genetic operation supervised by population convergence (SPC for short), and introduce it into the PF resampling process to provide CF tracker with enough effective target candidates. In this way, the correlation particle filter tracker based on SPC genetic operation method (SPC-CPF for short) is proposed.

Furthermore, in recent years, various visual target features extraction methods have been developed to help optimize the performance of the trackers, e.g., HOG descriptors [19], Haar-like features [20], histograms [21], covariance descriptors [22] and scale-invariant feature transform (SIFT for short) features [23]. However, most traditional trackers just employ a single feature, which is only suitable for a specific tracking scenario and lacks robustness to various tracking conditions. To address this issue, the multi-cue feature extraction strategy is presented in this paper. The multi-cue feature extraction strategy consists of different features such as Harris, HOG and SIFT. Among them, Harris feature has rotation invariance, SIFT feature has scale invariance, and HOG feature has optical and geometric invariance. The multi-cue feature extraction strategy combines the advantages of these features and can provide more effective image features for our tracker.

Consequently, in this paper, we advance the genetic operation supervised by population convergence, and introduce it into the PF resampling process to guarantee its sample diversity. And we optimize the traditional CF tracker via the proposed improved PF resampling process to address the issue that traditional CF relies heavily on the maximum response value of the response map and becomes unreliable when the map becomes ambiguous. Furthermore, the multi-cue feature extraction strategy is proposed to fuse different features via fuzzy logic theory to combine their advantages and help improve the efficiency and robustness of our tracker. And various experiments on OTB2015 and VOT2018 datasets proved that the multi-cue correlation particle filter tracker supervised by population convergence (SPC-MCCPF for short) can show a good tracking effect in various challenging tracking scenarios.

II. CORRELATION PARTICLE FILTER TRACKER SUPERVISED BY POPULATION CONVERGENCE

Correlation filter (CF for short) is an efficient tracking method and can adaptively deal with scale variation and robustly estimate the size of the visual target. So we employ CF as our basic tracking framework. But it still faces the problem of relying heavily on the response map and getting unreliable when the map becomes fuzzy. So we propose the SPC-CPF tracker to enhance the efficiency and robustness of the traditional CF-based trackers.

A. KERNELIZED CORRELATION FILTER (KCF)

The KCF tracker [19] uses a filter w to model the appearance of the target. The filter w is based on HOG features and trained on an image block x of $M \times N$ pixels. The training samples $x^{m,n}$, $(m,n) \in \{0,1,\ldots, M-1\} \times \{0,1,\ldots, N-1\}$ are the all possible circular shifts. By minimizing the error between the regression target $y^{m,n}$ and the training sample $x^{m,n}$, the filter w can be obtained as shown below:

$$w = \arg\min \sum_{m,n} \left|\left\langle \phi(x^{m,n}), w \right\rangle - y(m,n)\right|^2 + \lambda_1 \|w\|^2 \quad (1)$$

where $\langle , \rangle$ is the inner product, ϕ represents the mapping from Hilbert space to kernel space and λ denotes a regularization parameter($\lambda \geq 0$). The correlation coefficient is calculated by the fast Fourier transformation (FFT for short), the objective function is expressed as $w = \sum_{m,n} \alpha(m,n)\phi(x^{m,n})$, and α is computed by:

$$\alpha = F^{-1}\left(\frac{F(y)}{F(k^x) + \lambda}\right) \quad (2)$$

where F and F^{-1} represent the Fourier transform and its inverse, respectively. The kernel correlation $k^x = K(x^{m,n}, x)$ is calculated in the Fourier domain from a Gaussian kernel.

A patch z which has the same size as x is cropped from the new image frame. Then the response score can be expressed as:

$$f(z) = F^{-1}(F(k^z) \odot F(\alpha)) \quad (3)$$

where $k^z = K(z^{m,n}, \hat{x})$, $\odot$ is the Hadamard product and $\hat{x}$ is the learned appearance.

B. GENETIC OPERATION SUPERVISED BY POPULATION CONVERGENCE (SPC GENETIC OPERATION)

Definition 1: Let f_t be the average fitness value, set $f_{\max}$ as the best individual fitness and denote f as the average fitness value of all individuals whose is better than f_t. Then, the population convergence value can be defined as $= f_{\max} - f$. Here, we calculate the difference between $f_{\max}$ and f, rather than the difference between $f_{\max}$ and f_t. Thus we can get rid of the individuals with fitness lower than the average, and the extent of premature convergence can therefore be estimated more accurately. The crossover and mutation probability can be calculated by:

$$\begin{cases} P_g = -1/(1 + exp(-k_1 \cdot \Delta)) + 1.5 \\ P_m = -1/(1 + exp(-k_2 \cdot \Delta)) + 1 \end{cases} \tag{4}$$

where k_1 and k_2 are positive constants, and is non-negative. Then, the ranges of the crossover probability Pg and the mutation probability Pm are [0.5,1] and [0,0.5], respectively. Traditionally, the control parameters of genetic algorithm are fixed in the whole evolution process of the population, which can easily lead to premature convergence. In our advanced SPC genetic operation method, Pg and Pm can automatically adjust themselves according to the convergence coefficient Δ. When the degree of convergence of the population is large, Pg and Pm are increased automatically to overcome premature convergence. When the population tends to diverge, Pg and Pm are reduced automatically, so that the individuals tend to converge.

The crossover and mutation of the SPC genetic operation method proposed in this paper are shown in Equations (5) and (6):

$$\begin{cases} x_{t+1}^i = \alpha \cdot x_t^i + (1-\alpha) \cdot x_t^j \\ x_{t+1}^j = \alpha \cdot x_t^j + (1-\alpha) \cdot x_t^i \end{cases} \tag{5}$$

$$\begin{cases} x_{t+1}^k = \begin{cases} x_t^k + f\left(t, m_t - x_t^k\right), & p < 0.5 \\ x_t^k - f\left(t, m_t - l_t\right), & p \geq 0.5 \end{cases}, \; x_t^k \in [l_t, m_t] \\ f(t, y) = y \cdot \left(1 - p^{(1-t/T)^b}\right) \in (0, y), \; p \in U(0, 1) \end{cases} \tag{6}$$

where α and p are random numbers with the value range (0,1); x_t^i, x_t^j and x_t^k are particles at time t; x_{t+1}^i and x_{t+1}^j are updated samples obtained by crossover operator at time (t+1); x_{t+1}^k means updated particles obtained by non-uniform mutation operator at time (t+1); T means the last iteration; b determines the heterogeneity of the mutation operator; $f(t, y)$ is an adaptive mutation operator that can adaptively adjust the size of its step. Later in the iteration, only a small neighbouring area of the current solution is searched to guarantee locking of the optimal solution and the precision of positioning.

C. CORRELATION PARTICLE FILTER TRACKER BASED ON SPC GENETIC OPERATION (SPC-CPF TRACKER)

Firstly, the traditional correlation filter tracker is strongly dependent on the response map and gets unreliable when the map turns ambiguous. So we introduce the SPC genetic operation to help resample more candidates. Specifically, if the maximum response value is greater than the threshold θ, the coordinate of the maximum response value is taken as the new position of the object. Otherwise, we should employ the SPC resampling to resample enough target candidates for redetecting the target position. It can be expressed as follows:

$$\begin{cases} \max R \geq \theta & tracking\ by\ CF \\ \max R < \theta & tracking\ by\ SPC - CF \end{cases} \tag{7}$$

Secondly, the sample renewal equation of the quantum behaved particle swarm optimization based on wave function is applied to the SPC-CF tracker to improve its computing efficiency.

The flow chart of the SPC-CPF tracker is as follows:

The flow of SPC-CPF algorithm is as follows:

1) SAMPLING INITIALIZATION

Particle set x_0^i $(i = 1, 2, \ldots, S)$ is randomly generated based on the prior probability distribution function of particle state $p(x_0) \in U(0, 1)$. Set the initial value of the weights as $w_0^i = 1/S$. And set $p(x_0|y_0) = p(x_0)$ as the initial probability density function.

2) IMPORTANT DENSITY SAMPLING [24]

a) Calculating the significant density function

$$x_t \sim q(x_t^i|y_t) = p(x_t^i|x_{t-1}^i)p(x_{t-1}^i|y_{t-1}) \tag{8}$$

b) Updating the weights

$$w_t^i = w_{t-1}^i \frac{p(y_t\,|x_t^i)p(x_t^i\,|X_{t-1}^i)}{q(x_t^i\,|X_{t-1}^i\,, Y_t)} \tag{9}$$

c) Calculating the probability density function

$$p(x_t\,|y_{1:t}) = \sum_{i=1}^{S} w_t^i \delta(x_t - x_t^i) \tag{10}$$

where δ refers to the Dirac function. Firstly, according to Equation (8), S particles are randomly generated, and the particle weight and probability density are then updated according to Equations (9) and (10).

3) SPC RESAMPLING

d) Determination of the degree of particle degeneracy

$$N_{eff} = \frac{1}{\sum_{i=1}^{S} (w_t^i)^2} \tag{11}$$

where N_{eff} indicates the degree of particle degradation. If N_{eff} is greater than the threshold N_{th}, the degradation of the

particles is not obvious and we can skip the weight resetting step e). Otherwise, serious sample degradation may occur and weight resetting step must be carried out before the updating of the particles' positions.

e) Weights resetting

Reset all of the weights as $w_t^i = 1/S$.

f) SPC genetic manipulation

Firstly, we eliminate individuals with lower fitness and replace them with the same number of individuals with higher fitness. Then we carry out SPC genetic operations based on population convergence to enrich sample diversity.

4) OBTAIN THE SPATIAL RESPONSE MAP

We can obtain the spatial response map with IFFT. The mathematical model is expressed as:

$$R_m = \sum_k F^{-1}\left(F\left(\left\langle z_m, \hat{x}\right\rangle\right) \odot F\left(a\right)\right) \tag{12}$$

where $\hat{x}$ is the discrete Fourier transform (DFT for short) $\hat{x} = F(x)$; z_m indicates the m-th particle corresponding to the image patch; $\odot$ denotes the element-wise product and R_m means the corresponding response map.

5) LOCATE THE CENTER OF THE TARGET

We choose the image patch with the largest response value as the target centre:

$$\max R = \max(\max R_1, \max R_2, \ldots, \max R_S) \tag{13}$$

where $max\ R$ represents the largest response of the whole sample, and $max\ R_s$ means the maximum response of the s-th particle.

6) DETERMINE WHETHER TO END THE ITERATION

Firstly, we should determine whether to stop the iteration based on the iterative termination conditions. If not, we can return to step 2). Otherwise, we can end the iteration and output the result.

III. MULTI-CUE SPC-CPF TRACKING FRAMEWORK

Traditional visual trackers extract single feature to achieve precise tracking in specific tracking scenarios, but they lack robustness to various challenging tracking scenarios where there exist multiple tracking interferences such as illumination variation, motion blur and deformation. To solve this problem, we develop the multi-cue feature extraction strategy. As illustrated by Fig. 2, the multi-cue feature extraction strategy consists of different features such as Harris, HOG and SIFT. Among them, Harris feature has rotation invariance, SIFT feature has scale invariance, and HOG feature has optical and geometric invariance. This multi-cue feature extraction strategy combines the advantages of these features and can provide more effective image features for our tracker.

Meanwhile, this multi-cue feature extraction strategy employs fuzzy logic rule to fuse different target features to enhance our tracker's feature recognition ability. Taking the fifth feature as an example, the calculation formula of the

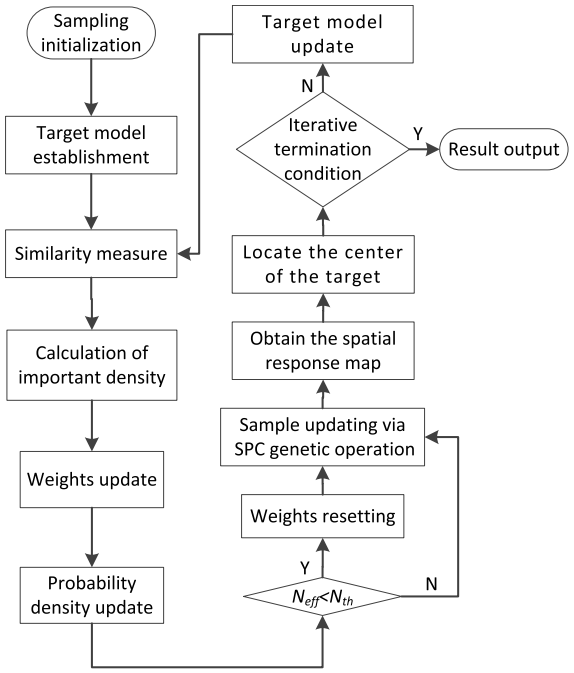

FIGURE 1. Flow chart for the SPC-CPF target tracking algorithm.

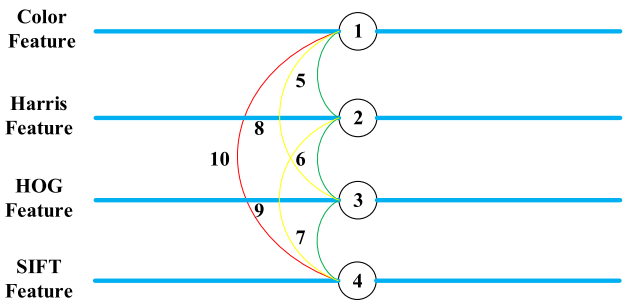

FIGURE 2. Graph illustration of the multi-cue feature extraction strategy.

similarity measure after the fusion of the Harris feature and the color feature can be obtained by:

$$\rho'(y) = \alpha\rho_C(y) + (1-\alpha)\rho_H(y) \tag{14}$$

where $\rho_C(y)$ and $\rho_H(y)$ respectively represent the similarity measurement of the color feature and the Harris feature. The weight coefficient α has been set to a fixed value in many studies. However, the actual weights of the visual features in various tracking scenes are quite different. Therefore, we propose to assign the weight parameter α according to the similarity measurement value of each feature in the current frame. For features with higher values of similarity measurement, we assign higher weights to them through fuzzy logic theory; otherwise, we assign lower weights, as shown in Table 1:

The fuzzy weight selection strategy: The fuzzy sets of $\rho_C(y)$ and $\rho_H(y)$ are set to $\{0.2, 0.4, 0.6, 0.8, 1\}$. The range of α is set to $\{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$. When the similarity measurement values of the two features are

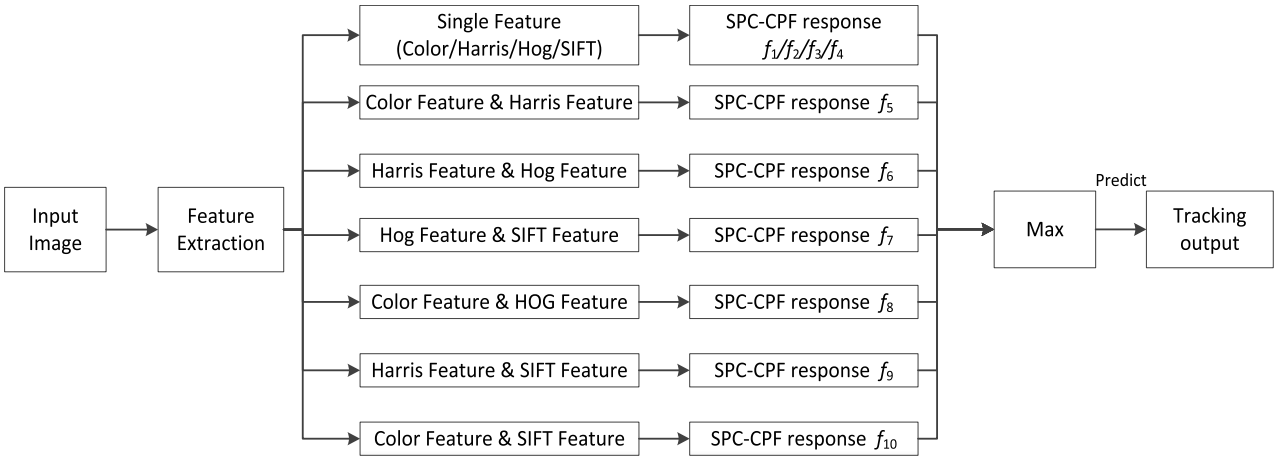

FIGURE 3. Overview of the proposed SPC-MCCPF.

TABLE 1. Fuzzy control of α.

ρ_H \ ρ_C	0.2	0.4	0.6	0.8	1.0
0.2	0.5	0.6	0.7	0.8	0.9
0.4	0.4	0.5	0.6	0.7	0.8
0.6	0.3	0.4	0.5	0.6	0.7
0.8	0.2	0.3	0.4	0.5	0.6
1.0	0.1	0.2	0.3	0.4	0.5

equal, α takes an intermediate value of 0.5. Otherwise, the position of α in the fuzzy control table is determined by the value of $\rho_C(y)$ and $\rho_H(y)$. For coefficient α, we set the upper limit of the interval of $\rho_C(y)$ as its column position; similarly, we set the upper limit of the interval of $\rho_H(y)$ as its row position. For instance, set $\rho_C(y)$ as 0.41 and $\rho_H(y)$ as 0.74. Obviously, 0.41 belongs to (0.4,0.6) while 0.74 belongs to (0.7,0.8). Therefore it can be determined that the position of α is row 0.6 and column 0.8, and the corresponding weight coefficient value of this position is 0.6.

Based on the multi-cue feature extraction strategy, fuzzy logic theory and the SPC genetic operation, the multi-cue correlation particle filter tracker supervised by population convergence (SPC-MCCPF for short) is developed. In particular, we extract multiple target features including color, Harris, SIFT and HOG features. And we fuse these different features with each other by fuzzy logic theory and obtain six new fusion features. For better robustness and precision, the proposed SPC-CPF method is then employed to handle with these four single features and six new fusion features. And we choose the one with the max response map as the best result of the current frame. The tracking framework of the proposed SPC-MCCPF algorithm is illustrated as below:

IV. EXPERIMENTAL ANALYSIS

We used MATLAB to implement the proposed tracker on a machine equipped with an Intel Core i-7-4790 CPU @ 3.60GHz with 16 GB RAM and runs at 23 frames per second. For experimental verification, we employ the OTB2015 [32] and VOT2018 [33] datasets to make qualitative comparisons, quantitative comparisons and statistical comparisons between our tracker with the state-of-the-art tracking methods.

A. QUALITATIVE COMPARISONS

In qualitative comparisons, we use the tracking benchmark dataset [32]. Our tracker is compared with 9 typical tracking methods including the C-COT [25], SRDCFdecon [26], MUSTer [27], Staple [28], SAMF [29], DSST [14], CNT [30], KCF [21], and Struck [31]. To better evaluate and analyze the advantages and disadvantages of theses trackers, we evaluate the trackers based on different complex interference factors including deformation, in-plane rotation, low resolution, occlusion, out-of-plane rotation, motion blur, scale variation, illumination variation, background clutter, out-of-view and fast motion. As illustrated by Fig. 4, these trackers are indicated by ten different colours. Qualitative analysis is carried out from the following aspects:

1. Illumination variation: In ''Shaking'', most of the trackers succeed in tracking despite dramatic changes in illumination. But only our tracker remains consistent with the target.

2. Fast motion: In ''Biker'', the target moves very fast. Ours, SRDCFdecon, C-COT, Staple and MUSTer can track the target successfully, but the others fail.

3. Motion blur: For the object in ''BlurOw'', motion blur occurs frequently. The MUSTer, DSST, Staple, SAMF, CNT, KCF and Struck lose the target easily when serious motion blur happens, while ours, SRDCFdecon and C-COT can track the target in the whole tracking process.

4. Deformation: In ''Dancer2'', target deformation occurs. Ours can always track the whole target, while others lose part of the target easily. Among these trackers, the C-COT, Staple, and SRDCFdecon can lose less part of the target than the others.

5. Background clutters: In ''MountainBike'', the tracking drift arises in SAMF, Struck, DSST, CNT and KCF when background clutter appears in the tracking region, such as #159 and #225. In addition, the KCF and Struck lose the target totally.

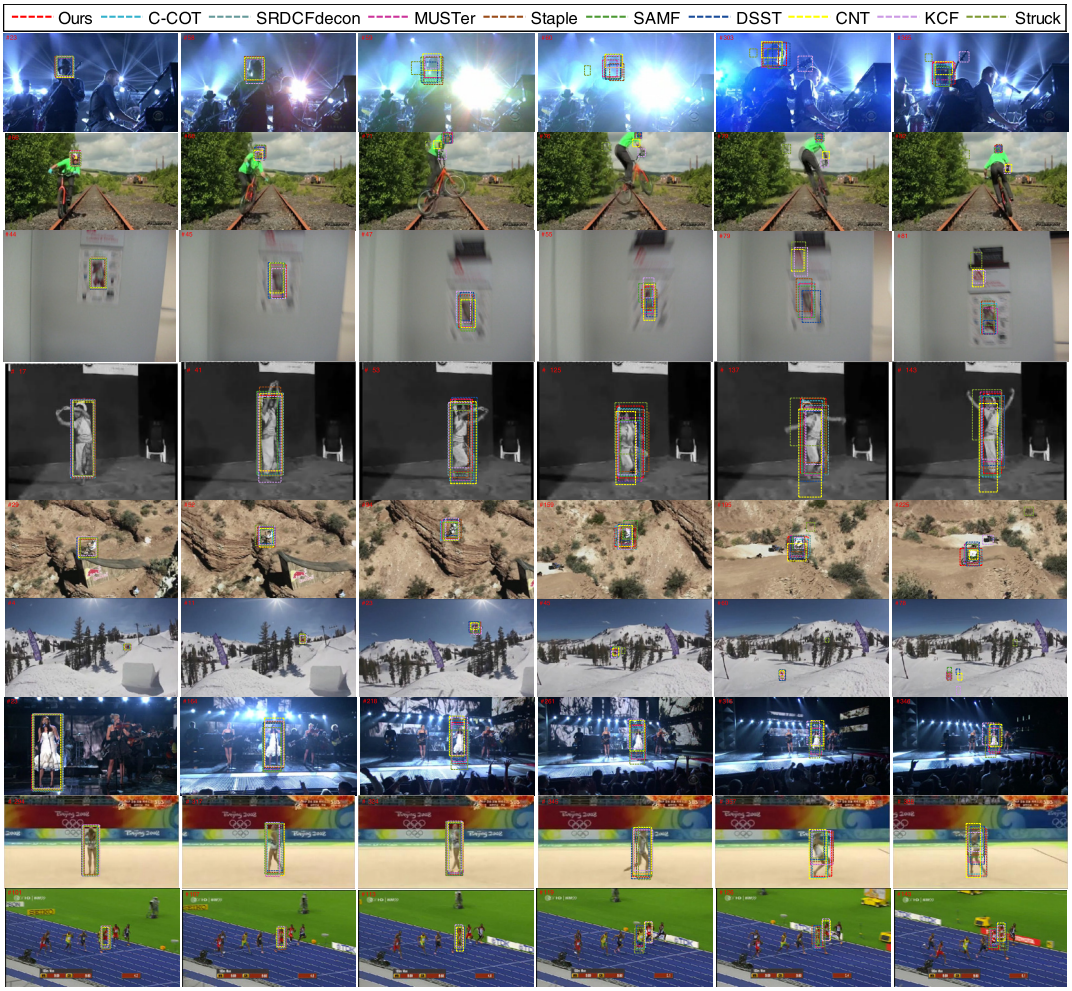

FIGURE 4. Qualitative comparisons of 10 trackers on nine challenging sequences (from top to bottom are Shaking, Biker, BlurOw, Dancer2, MountainBike, Skiing, Singer1, Gym and Bolt).

6. Low resolution: In ''Skiing'', low resolution occurs because the target is very small and it is difficult to extract enough efficient feature. It becomes very obvious after frame #78: SAMF, DSST, CNT, KCF and Struck totally lose the target.

7. Scale variation: In ''Singer1'', ours can better adapt to the scale change, but tracking drift arises in MUSTer, CNT, Staple, SAMF, KCF and Struck; serious tracking drifting even occurred in Struck and CNT.

8. Rotation: It is divided into out-of-plane and in-plane rotation. Both of them occur in ''Gym''. Only our bounding box matches the target well and tracks the target accurately.

9. Occlusion: In ''Bolt'', the target is completely or partially occluded by another target. Only our tracker can always track the target quickly and precisely.

B. QUANTITATIVE COMPARISONS

In order to comprehensively and reliably evaluate our tracking algorithm, we choose two indicators of precision rate and success rate for quantitative analysis.

1. Success rate: Given a threshold, the tracker is considered to succeed if and only if the overlap rate α is larger than the threshold. The success rate is defined as the percentage of successful frames. A larger value indicates better tracking performance.

2. Precision: Precision indicates the ratio of frames in which the centre location error (CLE for short) is within a given threshold. The larger the CLE value, the better the tracking performance.

In the experiments, we quantitatively compare these tracking methods in terms of overall performance and attribute-based performance for 50 sequences [32].

For overall performance evaluation of 50 sequences in the benchmark [32], we plotted the success plots and the precision plots of the above 10 trackers. The success plots show the success rates at different overlapping thresholds in the interval [0,1], while the precision plots show the accuracy at different CLE thresholds from 0 to 50 pixels. A comparison of the overall performance of the trackers is illustrated by Fig.5. This data illustrates that the proposed tracking method

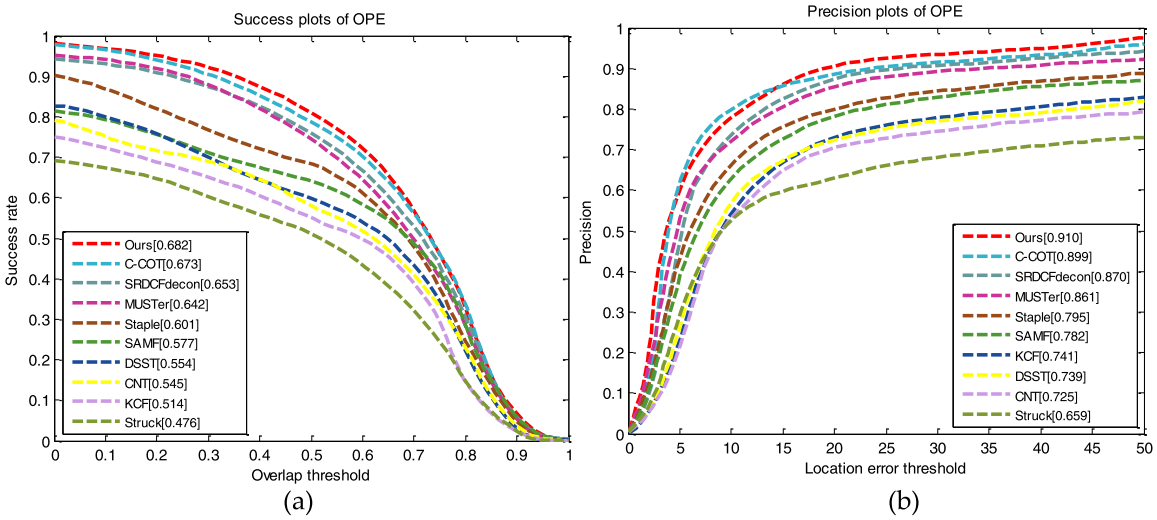

FIGURE 5. The success plots and precision plots of OPE for the trackers.

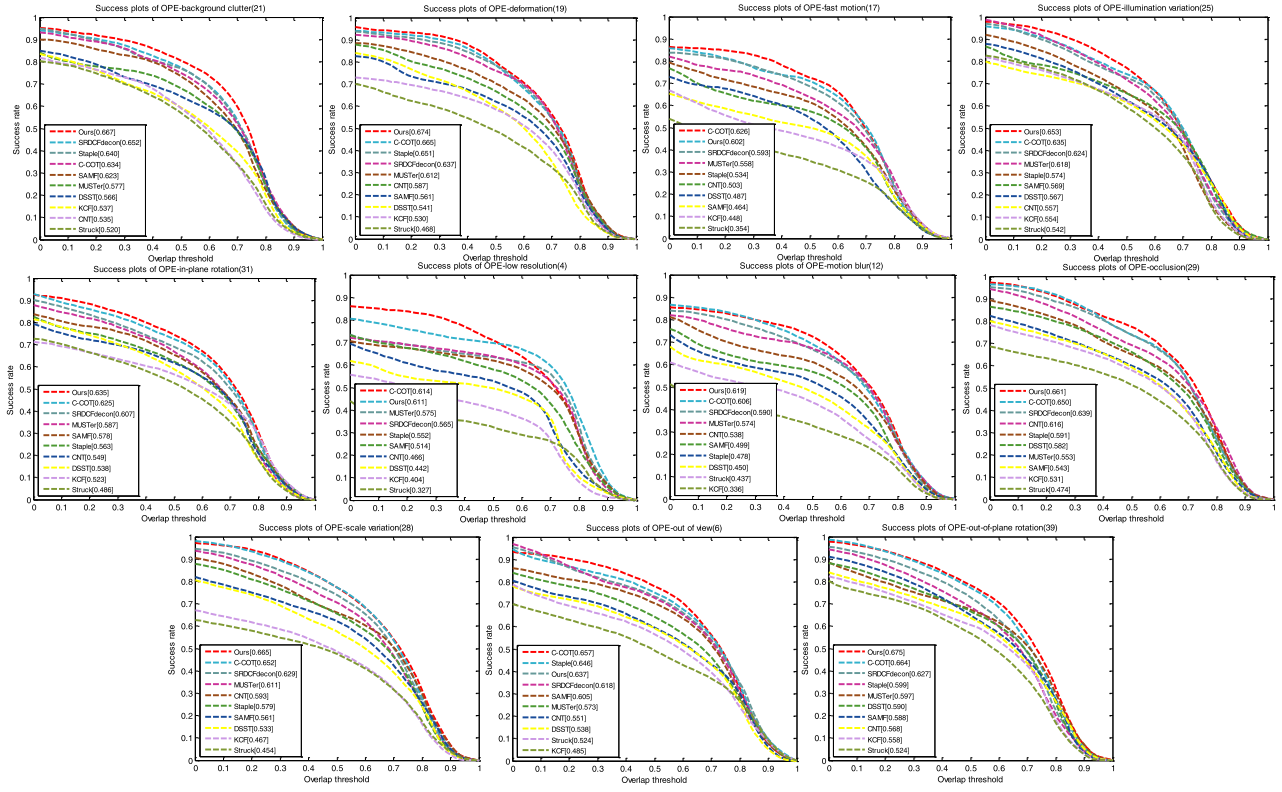

FIGURE 6. The success plots of OPE for the trackers on different attributes.

achieves satisfactory tracking results on the OTB2015 dataset and is superior to other typical trackers.

To further analyze the performance of our tracking method in different tracking scenes, we evaluate eleven attributes of the above ten trackers, which have been defined in OTB2015 [32]. Figures 6 and 7 show the success and accuracy scores of these trackers for each individual attribute on the OTB2015 dataset.

As shown in Figures 6 and 7, among these eleven attributes, whether it is a success rate score or a precision score, our tracker ranks in top one in at least eight attributes and ranks in top four in all these attributes. From these attribute-based data, we can see that our tracker has no obvious weaknesses and has achieved good tracking performance on all attributions.

C. STATISTICAL COMPARISON

To further evaluate the stability and robustness of the proposed SPC-MCCPF, a statistical comparison is performed on the VOT2018 [33] dataset, as shown in the table below:

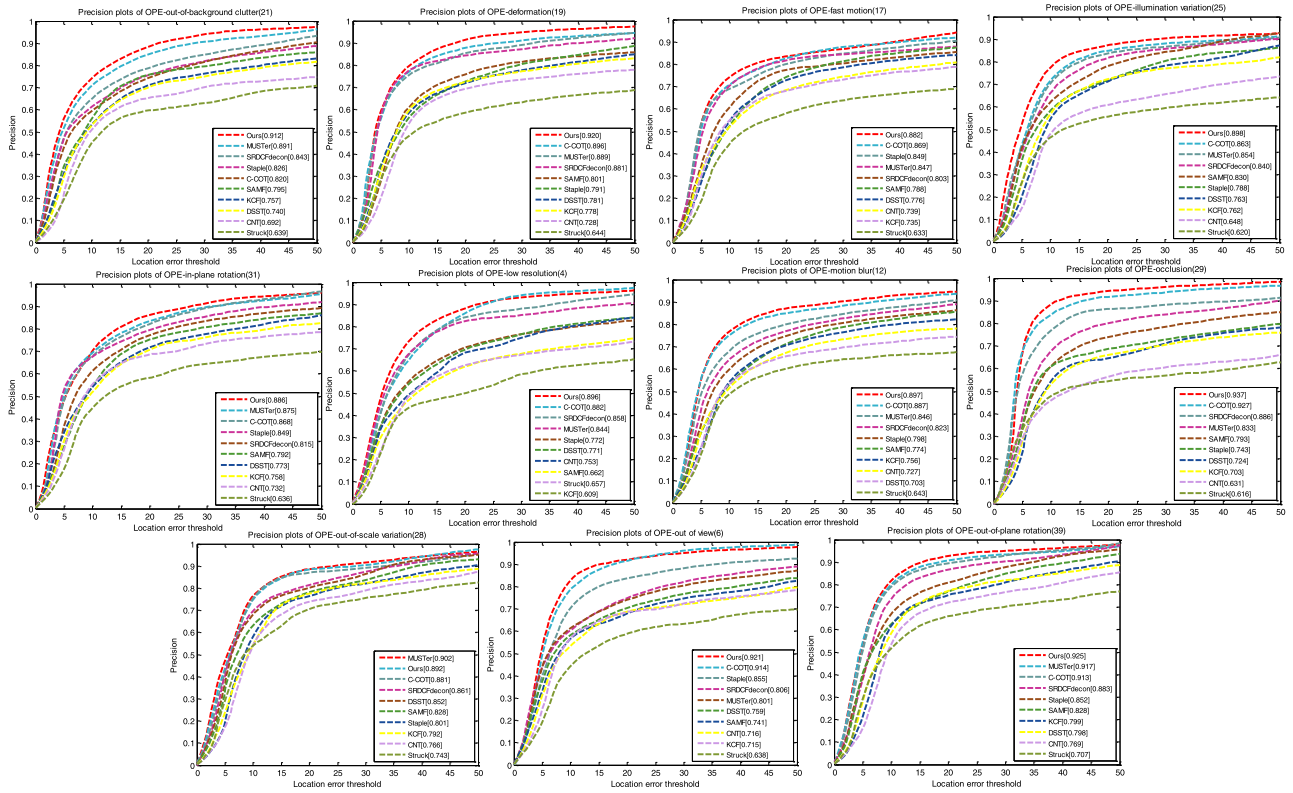

FIGURE 7. The precision plots of OPE for the trackers on different attributes.

TABLE 2. The statistical comparison on VOT2018 dataset.

Trackers	LADCF	DeepSTRCF	SPC-MCCPF	SiamVGG	ECO	C-COT	Staple	SRDCF	DSST	KCF
EAO	0.389①	0.345③	0.361②	0.286④	0.280	0.267	0.169	0.119	0.079	0.135
Acc.	0.503	0.523④	0.528③	0.531①	0.484	0.494	0.530②	0.490	0.395	0.447
R.Fail.	0.159①	0.215②	0.252③	0.318	0.276④	0.318	0.688	0.974	1.452	0.773

As shown in Table 2, we have carried out a statistical comparison on VOT2018 [33]. This several introduces several new trackers which are outstanding in VOT2018, including LADCF [34], SiamVGG [35] and DeepSTRCF [36], for more comprehensive statistical comparison. As shown above, the proposed SPC-MCCPF tracker ranks second in EAO, third in Acc. and R.Fail.. The EAO value of SPC-MCCPF is 8.3% smaller than the best one. The Acc. value of SPC-MCCPF is 0.4% smaller than the second best value. And the R.Fail. value of SPC-MCCPF is 17.2% higher than the second least one. Consequently, the proposed SPC-MCCPF tracker has obvious advantage over the state-of-the-art trackers, and has achieved good performance in all three aspects of statistical comparisons of VOT2018.

V. CONCLUSION

This paper presents a multi-cue correlation particle filter tracker supervised by population convergence. Specifically, we propose the genetic operation supervised by population convergence (SPC genetic operation for short) to supervise the resampling process of particle filter by the degree of population convergence. This resampling process can help guarantee the sample diversity and global optimization ability of the correlation filter and provide more candidates for our tracker to detect the object accordingly when the response map gets fuzzy and difficult to identify. Meanwhile we present the multi-cue feature extraction strategy through the fuzzy logic theory to help provide more efficiency and robust target features for our tracker. Extensive experimental results on the OTB2015 and VOT2018 datasets have proved the robustness and effectiveness of our tracker against the state-of-the-art trackers.

REFERENCES

[1] M. Danelljan, G. Bhat, M. Felsberg, and F. S. Khan, "ECO: Efficient convolution operators for tracking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 6931–6939.
[2] H. Fujita and D. Cimr, "Computer Aided detection for fibrillations and flutters using deep convolutional neural network," *Inf. Sci.*, vol. 486, pp. 231–239, Jun. 2019.

[3] C. Zhang, C. Liu, X. Zhang, and G. Almpanidis, "An up-to-date comparison of state-of-the-art classification algorithms," *Expert Syst. Appl.*, vol. 82, pp. 128–150, Oct. 2017.
[4] A. Gonczarek and J. M. Tomczak, "Articulated tracking with manifold regularized particle filter," *Mach. Vis. Appl.*, vol. 27, no. 2, pp. 275–286, Feb. 2016.
[5] D. Concha, R. Cabido, J. J. Pantrigo, and A. S. Montemayor, "Performance evaluation of a 3D multi-view-based particle filter for visual object tracking using GPUs and multicore CPUs," *J. Real-Time Image Process.*, vol. 15, no. 2, pp. 309–327, Aug. 2018.
[6] Q.-B. Zhang, P. Wang, and Z.-H. Chen, "An improved particle filter for mobile robot localization based on particle swarm optimization," *Expert Syst. Appl.*, vol. 135, pp. 181–193, Nov. 2019.
[7] Y. Wang, S. Hu, and S. Wu, "Object tracking based on Huber loss function," *Vis Comput*, vol. 35, no. 11, pp. 1641–1654, Nov. 2019.
[8] A. Mondal, "Neuro-probabilistic model for object tracking," *Pattern Anal. Appl.*, vol. 22, no. 4, pp. 1609–1628, Nov. 2019.
[9] Y. Wang, X. Luo, L. Ding, and J. Wu, "Object tracking via dense SIFT features and low-rank representation," *Soft Comput.*, vol. 23, no. 20, pp. 10173–10186, Oct. 2019.
[10] V. Naresh Boddeti, T. Kanade, and B. Vijaya Kumar, "Correlation filters for object alignment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 2291–2298.
[11] H. K. Galoogahi, T. Sim, and S. Lucey, "Multi-channel correlation filters," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 3072–3079.
[12] T. Zhang, S. Liu, C. Xu, B. Liu, and M.-H. Yang, "Correlation particle filter for visual tracking," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2676–2687, Jun. 2018.
[13] J. Yan, J. Du, R. Young, C. Chatwin, and P. Birch, "Real time UAV tracking of fast moving small target on ground," *J. Electron. Imag.*, vol. 27., p. 053010 1-12, 2018.
[14] M. Danelljan, G. Häger, F. Khan, and M. Felsberg, "Accurate scale estimation for robust visual tracking," in *Proc. Brit. Mach. Vis. Conf.*, Nottingham, U.K., Sep. 2014.
[15] Y. Sui, G. Wang, and L. Zhang, "Correlation filter learning toward peak strength for visual tracking," *IEEE Trans. Cybern.*, vol. 48, no. 4, pp. 1290–1303, Apr. 2018.
[16] M. Tang, B. Yu, F. Zhang, and J. Wang, "High-speed tracking with multi-kernel correlation filters," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4874–4883.
[17] Y. Sun, C. Sun, D. Wang, Y. He, and H. Lu, "ROI pooled correlation filters for visual tracking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 5783–5791.
[18] Y. Huang, Z. Zhao, B. Wu, Z. Mei, Z. Cui, and G. Gao, "Visual object tracking with discriminative correlation filtering and hybrid color feature," *Multimedia Tools Appl.*, vol. 78, no. 24, pp. 34725–34744, Dec. 2019.
[19] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "High-speed tracking with kernelized correlation filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 583–596, Mar. 2015.
[20] H. Song, "Robust visual tracking via online informative feature selection," *Electron. Lett.*, vol. 50, no. 25, pp. 1931–1933, Dec. 2014.
[21] W. Zhong, H. Lu, and M. H. Yang, "Robust object tracking via sparsity-based collaborative model," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 1838–1845.
[22] J. Gao, H. Ling, W. Hu, and J. Xing, "Transfer learning based visual tracking with Gaussian processes regression," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 188–203.
[23] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. ICCV*, 1999, vol. 99, no. 2, pp. 1150–1157.
[24] R. Al Mallah, A. Quintero, and B. Farooq, "Distributed classification of urban congestion using VANET," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 9, pp. 2435–2442, Sep. 2017.
[25] M. Danelljan, A. Robinson, F. S. Khan, and M. Felsberg, "Beyond correlation filters: Learning continuous convolution operators for visual tracking," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 472–488.
[26] M. Danelljan, G. Hager, F. Shahbaz Khan, and M. Felsberg, "Adaptive decontamination of the training set: A unified formulation for discriminative visual tracking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1430–1438.
[27] Z. Hong, Z. Chen, C. Wang, X. Mei, D. Prokhorov, and D. Tao, "Multi-store tracker (MUSTer): A cognitive psychology inspired approach to object tracking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 749–758.
[28] L. Bertinetto, J. Valmadre, S. Golodetz, O. Miksik, and P. Torr, "Staple: Complementary learners for real-time tracking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, vol. 38, no. 2, pp. 1401–1409.
[29] Y. Li and J. Zhu, "A scale adaptive kernel correlation filter tracker with feature integration," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 254–265.
[30] K. Zhang, Q. Liu, Y. Wu, and M. H. Yang, "Robust visual tracking via convolutional networks without training," *IEEE Trans. Image Process.*, vol. 25, no. 4, pp. 1779–1792, Feb. 2016.
[31] S. Hare, S. Golodetz, A. Saffari, V. Vineet, M.-M. Cheng, S. L. Hicks, and P. H. Torr, "Struck: Structured output tracking with kernels," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 10, pp. 2096–2109, Oct. 2016.
[32] Y. Wu, J. Lim, and M. H. Yang, "Online object tracking: A benchmark," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 2411–2418.
[33] M. Kristan *et al.*, "The sixth visual object tracking VOT2018 challenge results," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2018.
[34] T. Xu, Z.-H. Feng, X.-J. Wu, and J. Kittler, "Learning adaptive discriminative correlation filters via temporal consistency preserving spatial feature selection for robust visual object tracking," *IEEE Trans. Image Process.*, vol. 28, no. 11, pp. 5596–5609, Nov. 2019.
[35] Y. Li and X. Zhang, "SiamVGG: Visual tracking using deeper siamese networks," 2019, *arXiv:1902.02804*. [Online]. Available: https://arxiv.org/abs/1902.02804
[36] F. Li, C. Tian, W. Zuo, L. Zhang, and M.-H. Yang, "Learning spatial-temporal regularized correlation filters for visual tracking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4904–4913.