

Classification of Infrared Objects in Manifold Space Using Kullback-Leibler Divergence of Gaussian Distributions of Image Points

Huilin Ge ¹, Zhiyu Zhu ^{1,*}, Lou Kang ¹, Wei Wei ², Runbang Liu ¹, Robertas Damaševičius ^{3,4,*} and Marcin Woźniak ³

¹ School of Electronics and Information, Jiangsu University of Science and Technology, Zhenjiang 212003, Jiangsu, China; ghl1989@just.edu.cn (G.H.); 172030045@stu.just.edu.cn (L.K.); liurunbang212@163.com (R.L.)

² College of Computer Science and Engineering, Xi'an University of Technology, Xi'an 710048, China; weiwei@xaut.edu.cn

³ Faculty of Applied Mathematics, Silesian University of Technology, 44-100 Gliwice, Poland; marcin.wozniak@polsl.pl

⁴ Department of Applied Informatics, Vytautas Magnus University, Kaunas 44404, Lithuania

* Correspondence: zzydzz@163.com (Z.Z.); robertas.damasevicius@polsl.pl (R.D.)

Received: 3 February 2020; Accepted: 5 March 2020; Published: 8 March 2020

Abstract: Infrared image recognition technology can work day and night and has a long detection distance. However, the infrared objects have less prior information and external factors in the real-world environment easily interfere with them. Therefore, infrared object classification is a very challenging research area. Manifold learning can be used to improve the classification accuracy of infrared images in the manifold space. In this article, we propose a novel manifold learning algorithm for infrared object detection and classification. First, a manifold space is constructed with each pixel of the infrared object image as a dimension. Infrared images are represented as data points in this constructed manifold space. Next, we simulate the probability distribution information of infrared data points with the Gaussian distribution in the manifold space. Then, based on the Gaussian distribution information in the manifold space, the distribution characteristics of the data points of the infrared image in the low-dimensional space are derived. The proposed algorithm uses the Kullback-Leibler (KL) divergence to minimize the loss function between two symmetrical distributions, and finally completes the classification in the low-dimensional manifold space. The efficiency of the algorithm is validated on two public infrared image data sets. The experiments show that the proposed method has a 97.46% classification accuracy and competitive speed in regards to the analyzed data sets.

Keywords: manifold learning; feature mapping; infrared image recognition; object classification; Kullback-Leibler Divergence

1. Introduction

Feature detection and matching are the basis of many image processing applications in the computer vision domain [1–5] and elsewhere [6]. Infrared small object detection is a focus of ongoing research in numerous areas, such as aircraft tracking [7], ship detection [8], 3D scene reconstruction [9] and video surveillance [10]. Infrared small object recognition in difficult environments such as those with a complex background and object clutter or those with low illumination is highly important, and is a difficult task in infrared search and tracking systems [11]. Contrary to visible light images, the infrared images do not have color information, while luminance is influenced by the thermal radiation of the object and background. Moreover, small infrared objects miss texture

features due to the long sensing distance [12]. As a result, common object tracking methods based on visible features cannot recognize the difference between the object and its background in the infrared image [13].

Many different approaches have been proposed for infrared object tracking such as saliency extraction [9], multiscale patch-based contrast measure and a temporal variance filter [14], feature learning and fusion, reliability weight estimation based on nonnegative matrix factorization [15], Poisson reconstruction and the Dempster-Shafer theory [16], three-dimensional scalar field [17], a double-layer region proposal network (RPN) [18], Siamese convolution network [19], a mixture of Gaussians with modified flux density [20], spatial-temporal total variation regularization and weighted tensor [21], two-stage U-skip context aggregation network [22], histogram similarity map based on the Epanechnikov kernel function [23], quaternion discrete cosine transform [24], non-convex optimization [25], Mexican-hat distribution of pixels [26], and Schatten regularization with reweighted sparse enhancement [27].

Manifold learning assumes that the infrared object images to be classified are distributed as a set of points on the manifold space [28–31]. The purpose of manifold learning is to put forward a representation method for mapping manifold data point sets to a low-dimensional space [32]. The prior knowledge of the low-dimensional manifold of an image can be effectively used for the image reconstruction method, as has been demonstrated for the computer tomography images in [33]. The current infrared image classification methods in manifold space mainly map high-dimensional infrared object data point sets to low-dimensional space, and then complete classification of infrared object data points. Knowing the intrinsic structure of data, efficient manifold-based image classification methods can be constructed [34].

The infrared object classification method on the manifold takes advantage of one property of the manifold space, so that the manifold space can be regarded as a small piece of Euclidean space locally [35]. It attempts to obtain the distribution information of the infrared object data point set in the entire manifold with all the low-dimensional local maps. Most of the infrared object classification methods on the manifold are developed under the concept of describing the relationship between points in the data point set in high-dimensional manifold space [36]. In order to better describe the local relationship of infrared object data points on high-dimensional manifolds, initial research focused on the topological relationships between local region data points. The local linear embedding (LLE) method was proposed to describe the local data points in the manifold space by measuring the Euclidean distance between data points [37]. LLE considers that each point can be represented by its surrounding points. The distance between the surrounding data points and the object point within the neighborhood is used as the weight. However, LLE cannot well reconstruct a high-dimensional manifold data point set with unevenly distributed data point sets. In order to further describe the distance relationship between the data points of the infrared object image on the manifold, the isometric feature mapping (ISOMAP) method [38] introduced the concept of geodesic distance. The idea of this method is to construct a feature neighborhood map of infrared object image data points, which can represent the local information of the infrared object image data point set in a high-dimensional manifold space. However, this method is more suitable for scenarios where the manifold space of infrared image data points is relatively flat, and the calculation cost is relatively high when calculating the optimal route in the neighborhood of the infrared object image data point set [39]. The Laplace feature mapping algorithm (LE) [40] is a different idea to using the Euclidean distance between points on the manifold. LE expresses the local relationship of the infrared object image data point set in the manifold space through the graph theory. However, LE does not perform well when the data points on the manifold are far away from each other.

Image classification using non-Euclidean manifolds such as a Grassmann manifold and the Symmetric Positive Definite (SPD) manifold [41] is becoming increasingly more attractive. The weights of the iterative manifold embedding (IME) layer are learned by unsupervised strategy, which has been used to analyze the intrinsic manifolds of data sets with missing data [42]. The distribution of image data in multi-view manifold space can be captured by Multi-view Generative Adversarial Network (GAN), which can map the shape and view manifolds in a lower dimensionality latent space

[43]. The Wasserstein-driven low-dimensional manifold model (W-LDMM) can be used for noise estimation, image denoising and noisy image inpainting tasks [44]. On the other hand, Riemannian manifolds can be used to visualize geometric transformations in images [45], which has many useful applications for image augmentation aiming to improve the accuracy of classification for small data sets. Pixels corresponding to different image classes tend to be segmented better on the Riemannian manifold than in the spectral space, since image points mapped from the Gaussian probability distribution are radially distributed on the skewed surface of Riemannian manifold [46]. However, due to the specific characteristics of Riemannian manifolds, the traditional machine learning methods often fail on them [47], which motivates the exploration of new feature extraction and classification methods.

Currently, there are still many problems to be addressed in infrared object classification, such as the occlusion of objects, change of object position and change of light. [48]. Improvements in the accuracy of infrared object classification when the infrared object data set has only a few samples (the “small data” problem [49]) and lack of prior knowledge have become important research directions for manifold learning in infrared object classification.

The novelty and contribution of this paper is outlined as follows. We propose a novel manifold learning algorithm for infrared object detection and classification that uses Kullback-Leibler (KL) Divergence to minimize the loss function between two symmetrical distributions of points (the distribution in the manifold space and the distribution of the points of the infrared image), and finally completes the classification in the low-dimensional space.

2. Materials and Methods

To improve the classification accuracy of infrared objects, this manuscript proposes a manifold learning method for classification. The proposed algorithm (see Figure 1) constructs a high-dimensional manifold space by using the infrared object image pixels as the manifold space dimension. The constructed high-dimensional manifold is then mapped into a low-dimensional space. By describing the local probability distribution information of the infrared image data points, the reduced-dimensional data set can accurately retain the high-dimensional information. Finally, the difference between the probability distribution of the data point set in manifold space and the probability distribution of the low-dimensional space is minimized using the KL Divergence.

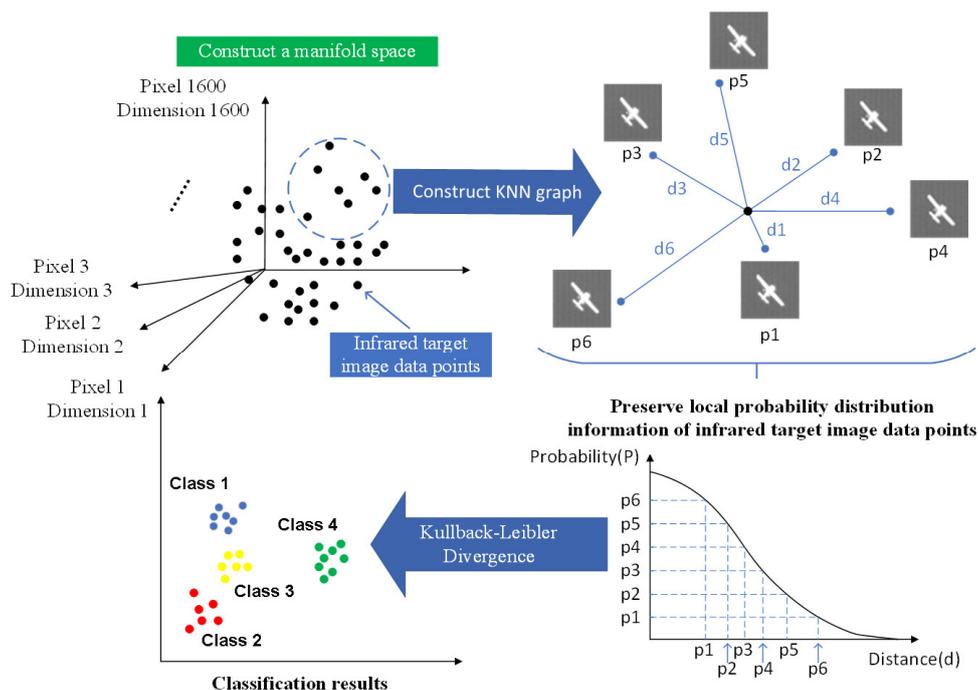


Figure 1. Schematic outline of the algorithm proposed in this article.

2.1. Construction of High-Dimensional Manifold Space

Each pixel of the infrared object image is used as each dimension in the manifold space to construct a high-dimensional manifold space. The infrared objects of various categories appear as data points in this manifold space. The infrared object image manifold constructed in this paper is a differential manifold, but its local space can be understood as Euclidean-style space. The definition of this high-dimensional infrared object manifold is as follows.

If every point p in X has an open neighborhood $Y \in X$, and in addition Y and an open subset in the Euclidean space R^n are homeomorphic, then X is an n -dimensional topological manifold.

Let X be a d -dimensional manifold and $V = \{(U_\alpha, \varphi_\alpha)\}_{\alpha \in I}$ is a set of X coordinate cards. V is a C^k differential structure of X when the following conditions are met.

The manifold M formed by $\{U_\alpha: \alpha \in I\}$ is an open cover, and the mapping of $\varphi_\alpha: U_\alpha$ to $\varphi_\alpha(U_\alpha) \subset R^d$ is homeomorphic. If $U_\alpha \cap U_\beta \neq \emptyset$, and the double mapping $\varphi_\alpha \varphi_\beta^{-1}: \varphi_\beta(U_\alpha \cap U_\beta)$ to $\varphi_\alpha(U_\alpha \cap U_\beta)$ and its inverse mapping are all k times differentiable, then $(U_\alpha, \varphi_\alpha)$ is compatible with (U_β, φ_β) .

2.2. Low-Dimensional Mapping and Classification Method of Infrared Object Manifold Space

After the high-dimensional manifold space is constructed, the dimension of the manifold space is still high, and it needs to be reduced to complete the classification of infrared objects. In order to improve the efficiency, we have adopted a dimensionality reduction method for the manifold space. The Gaussian distribution is used to describe the distribution characteristics of the infrared object image data point set in the high-dimensional manifold space, and the Student's t-distribution is used to describe the distribution characteristics of infrared object image data point set after dimensionality reduction in the low-dimensional space. Finally, the difference of two distributions is minimized to complete the infrared object classification process. The steps are mainly divided into the description of the distribution information of the infrared object image data points in two spaces and the process of obtaining the minimum value of the difference between two distributions.

2.2.1. Projection of Infrared Object Image Data Points

The infrared object image data point set is projected onto a plane by random projection. On this plane, all data point sets show a discrete distribution. The data set of infrared object image data in the n -dimensional manifold space X can be defined by Equations (1)–(3).

$$x_i, x_j \in X, \quad (1)$$

$$x_i = (x_i^{(1)}, x_i^{(2)}, x_i^{(3)})^T, \quad (2)$$

$$x_j = (x_j^{(1)}, x_j^{(2)}, x_j^{(3)})^T \quad (3)$$

Since the infrared object image data points are projected onto a two-dimensional plane, when calculating the distance between these points and the surrounding neighboring points, the distance L_p between x_i and x_j can be defined by the Euclidean distance using the following formula:

$$L_2(x_i, x_j) = \left(\sum_{l=1}^n |x_i^{(l)} - x_j^{(l)}|^2 \right)^{\frac{1}{2}} \quad (4)$$

2.2.2. Construction of the KNN Map of the Infrared Object Image Data Points

The k-Nearest Neighbor (KNN) map of the infrared object image data points can be constructed by the above steps, and the obtained KNN map initially describes the local features of the manifold data point set. The local feature information of the infrared object image data in the manifold space should also have the same local characteristics after being projected into the low-dimensional space. For infrared object image data points in the high-dimensional space, the KNN maps can be used to

represent the weighted image maps of the surrounding data points to the object data points. Here a weight is the Euclidean distance between points. We assume that a normal distribution describes the distribution probability of these neighboring points. For example, there are two infrared object image data points x_i and x_j in a high-dimensional manifold space. The infrared object image data point x_i is the center of the Gaussian distribution. x_i uses probability $P_{j|i}$ to select x_j as its nearest neighbor. $P_{j|i}$ of the neighboring points is inversely proportional to the distance of x_j and can be expressed as:

$$P_{j|i} = \frac{\exp(-\|x_i - x_j\|^2 / 2\sigma_i^2)}{\sum_{k \neq i} \exp(-\|x_i - x_k\|^2 / 2\sigma_i^2)}, \quad (5)$$

where σ_i is the variance of the Gaussian distribution with x_i as the center point.

In order to avoid the congestion problem when the Gaussian distribution is projected into the two-dimensional space, Student's t-distribution is used to describe the local relationship of the infrared object image data points in the low-dimensional space. The t-distribution in the two-dimensional space is derived from the Gaussian distribution of the infrared object image data point set in the manifold space. Assuming that the probability distribution of the neighborhood near the point set of the infrared object image data in the high-dimensional space can be represented by the normal distribution $N(\mu, \sigma^2)$, the mean value of the T distribution in the low-dimensional space can be derived by:

$$\bar{u} = \frac{1}{n} \sum_{i=1}^n u_i, \quad (6)$$

where u_i is the infrared object image data points in the KNN domain and \bar{u} is its mean.

Similarly, the variance of these points can be derived from the following formula.

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (u_i - \bar{u})^2, \quad (7)$$

The distribution of the infrared object image data points in the two-dimensional space is determined by the variance and standard deviation of the normal distribution in the manifold space and the number of data points. Therefore, the T distribution random variable can be constructed by:

$$t = \frac{\bar{u} - \mu}{\frac{s}{\sqrt{n}}} \quad (8)$$

In this way, the probability distribution of the infrared object image data points in the two-dimensional space can be constructed according to the T distribution:

$$q_{i|j} = \frac{(1 + \|y_i - y_j\|^2)^{-1}}{\sum_{k \neq i} (1 + \|y_k - y_i\|^2)^{-1}}, \quad (9)$$

where $q_{i|j}$ is the probability distribution of the infrared object image data points in the 2D space.

In order to maintain the symmetry of the two probability distributions, a uniform symmetrical distance function is introduced as shown in the following formula:

$$p_{ij} = \frac{p_{j|i} + P_{i|j}}{2n} \quad (10)$$

2.2.3. Dimensionality Reduction

Since the distribution of high-dimensional space $P_{j|i}$ and the distribution of low-dimensional space $q_{i|j}$ should be as similar as possible, the KL divergence is used as a loss function to minimize the difference of this two distributions. The process of dimension reduction and classification is converted to the process of obtaining the minimum value of the loss function:

$$C = \sum_i \sum_j p_{j|i} \log \frac{p_{j|i}}{q_{i|j}} \quad (11)$$

In order to control overfitting that occurs due to the small number of samples, a degree of confusion is also set to avoid overfitting. The degree of confusion can be defined by

$$\text{Perp}(p_i) = 2^{H(P_i)}, \quad (12)$$

where $H(P_i)$ is the Shannon entropy with p_i defined as follows:

$$H(P_i) = - \sum_j p_{j|i} \log_2 p_{j|i} \quad (13)$$

The degree of confusion changes in proportion to the entropy. With this feature, the value of entropy can be changed by adjusting the degree of confusion.

The next step is to construct the objective function. The positive samples of the infrared object image data points projected into the 2D space should be clustered, and the negative samples of the infrared object image data points should be placed far away from the positive samples. The weight between the points can be defined by Equation (14),

$$P(e_{ij} = 1) = f(\|y_i - y_j\|^2), \quad (14)$$

where y_i and y_j represent two points in the low-dimensional space. A binary edge e_{ij} between these two points has a weight value of 1, and $P(e_{ij} = 1)$ represents the probability that these two points exist. When the distance between y_i and y_j becomes closer, the value $P(e_{ij} = 1)$ becomes larger. The expression with weight w_{ij} is used in practical applications as follows:

$$P(e_{ij} = w_{ij}) = P(e_{ij} = 1)^{w_{ij}} \quad (15)$$

The positive sample set in the infrared object image data point set is defined as E , and the negative sample set is defined as \bar{E} . These sets are obtained from the KNN diagram as follows.

$$O = \prod_{(i,j) \in E} p(e_{ij} = 1)^{w_{ij}} \prod_{(i,j) \in \bar{E}} (1 - p(e_{ij} = 1))^\gamma, \quad (16)$$

where γ is the weight of the negative samples.

The optimization process can be understood as maximizing the probability of weighted edges of positive samples in the KNN graphs and minimizing the probability of weighted edges of negative samples in the KNN graphs. For the convenience of calculation, the above optimization formula can be transformed into the following formula:

$$O = \sum_{(i,j) \in E} w_{ij} P(e_{ij} = 1) \sum_{(i,j) \in \bar{E}} \gamma (1 - p(e_{ij} = 1)) \quad (17)$$

The negative sample \bar{E} increases the computational complexity, and it is not easy to directly use gradient descent for training. Therefore, a negative sampling algorithm is selected in this paper, and a negative sample is formed based on a randomly selected number of infrared object image data points that conform to the noise distribution $P_n(j)$. The objective function is expressed by Equation (18).

$$O = \sum_{(i,j) \in E} w_{ij} (p(e_{ij} = 1)) + \sum_{k=1}^M E_{jk} \sim P_n(j) \gamma \log(1 - p(e_{ij} = 1)) \quad (18)$$

2.2.4. Classification in Infrared Object Manifold Space

Finally, the low-dimensional representation of the manifold space and the classification results of different categories of objects are obtained. First, we find the distance between the input infrared object image and each point set in a two-dimensional space. Then, a set of points with the smallest distance from the input object image is considered as the category to which the object image belongs.

The problems of gradient vanishing and gradient explosion still occur during training. This problem can be solved by converting the edge between two points into a binary edge with the number of w_{ij} . When more edges with large weights are converted into binary edges, the calculation cost will become higher. Therefore, we use random sampling in these transformed binary edges to solve the

problem of high computational cost. In the optimization loss function, we use the asynchronous stochastic gradient descent (ASGD) algorithm [50] to improve the execution performance.

2.2.5. Illustration of the Method Stages

To illustrate the different stages of the method operation, we used the thermal infrared image “ambassador_morning” from the CSIR-CSIO Moving Object Thermal Infrared Imagery Dataset (MOTIID) [51]. Figure 2 shows the probability distributions calculated for different classes of image data points and the constructed KNN tree.

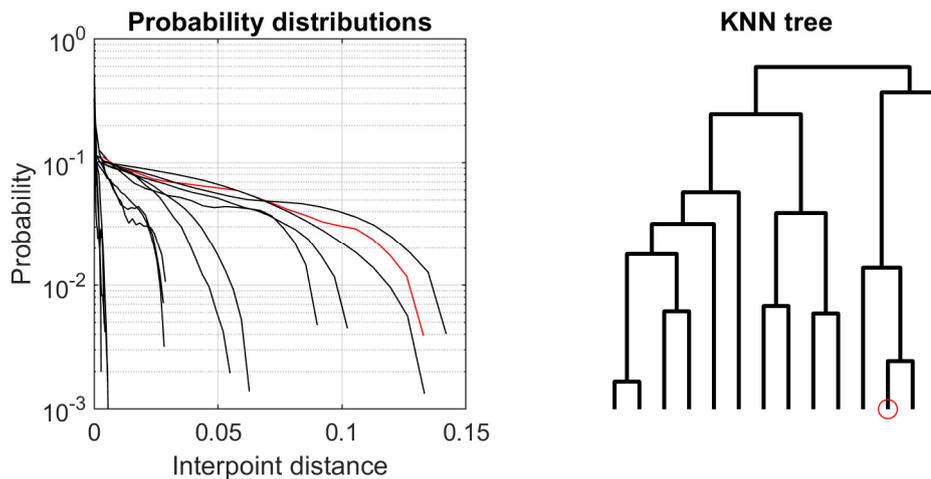


Figure 2. Illustration of the construction of the k-Nearest Neighbor (KNN) tree (right) using probability distributions of image points (left) for the “ambassador_morning” image. The red color shows a tree node corresponding to the tracked infrared object.

Figure 3 illustrates the low-dimensional embedding of the “ambassador_morning” image into a two-dimensional manifold and the corresponding tracking result. Note that the sets of closely related image points form clusters in the manifold space, which correspond to the particular areas of the target infrared image such as the tracked object.

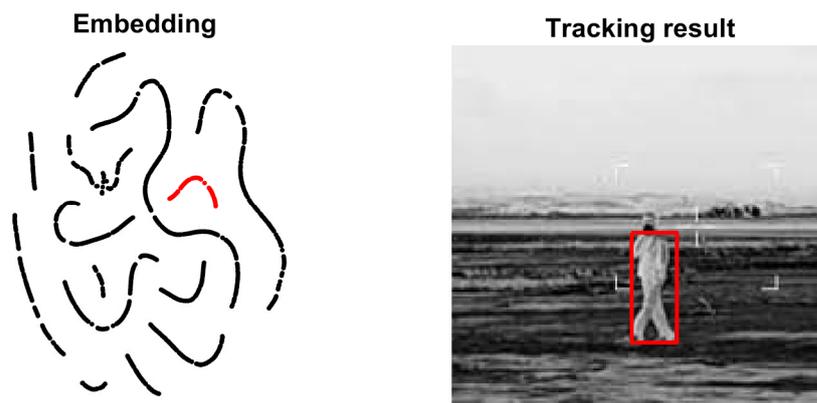


Figure 3. Illustration of the embedding of the image points into a two-dimensional manifold (left) and the corresponding tracked infrared object (right) for the “ambassador_morning” image. The red color outlines the cluster of a closely related set of image points in the two-dimensional manifold and the tracked object.

3. Experimental Verification

The hardware platform used in this experiment is a Microsoft Surface Pro laptop with Intel Core M3-7Y30 1.61GHz CPU and 4 GB RAM.

The experiments in this paper used the following infrared object image data sets.

- Two infrared video sequences named "8_quadrocopter1" and "8_horse" in the open-source LTIR data set (v1.0) of the Computer Vision Laboratory of Linköping University [52].
- The infrared video sequence named "data1" in the data set for dim-small object detection and tracking of aircraft in infrared image sequences of the ATR Key Laboratory of National University of Defense Technology [7].
- An infrared video sequence named "6a" in the OSU Color-Thermal Database data set [53].
- The four infrared image sequences of the CSIR-CSIO Moving Object Thermal Infrared Imagery Dataset (MOTIID) data set named "ambassador_morning", "auto_partially_occluded", "bike_far" and "dog_evening" [51].

The characteristics of image sequences are summarized in Table 1.

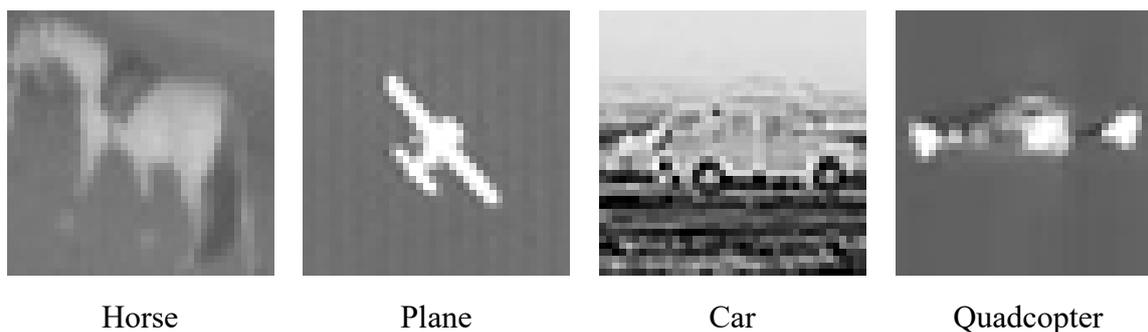
Table 1. Characteristics of image sequences.

Image Sequence	No. of Images (Frames)	Resolution, px	Bit Depth
8_quadrocopter1	178	640 × 480	8
8_horse	348	324 × 256	8/16
data1	398	256 × 256	24
6a	1652	320 × 240	8
ambassador_morning	155	640 × 480	24
auto_partially_occluded	219	640 × 480	24
bike_far	202	640 × 480	24
dog_evening	69	640 × 480	24

Eight kinds of objects are selected in the eight infrared object image data sets. There are 20 samples of each object. The size of these infrared object images is resized to 40 × 40 px, and the corresponding category label data is also added. The data set information is shown in Table 2, while the examples of images are shown in Figure 4.

Table 2. Description of different infrared object categories.

Category	Distance	Background	Label
Horse	Far	Ground	1
Plane	Far	Sky	2
Car	Close	Ground	3
Quadcopter	Far	Wall	4
Tricycle	Close	Ground	5
Gog	Close	Ground	6
Motorcycle	Far	Ground	7
Pedestrian	Close	Ground	8



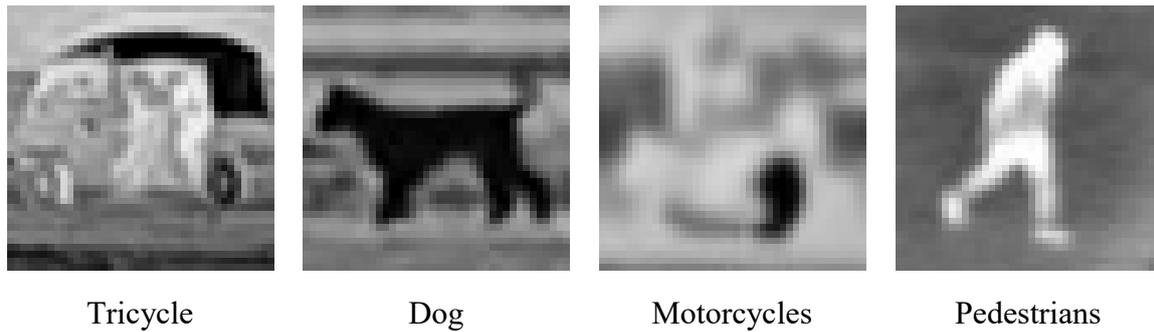


Figure 4. Sample images of different infrared object categories.

After reducing the dimensionality of the manifold space of the infrared object images, different categories of infrared object image data points are represented in the form of point sets in the low-dimensional space. The set of points with different label numbers in Figure 5 represents different categories of infrared object images. We can see that different categories of infrared objects have been effectively classified, and there are obvious gaps between different categories. There were no misclassifications observed, partly because of the small number of samples, which were separated in the manifold space well.

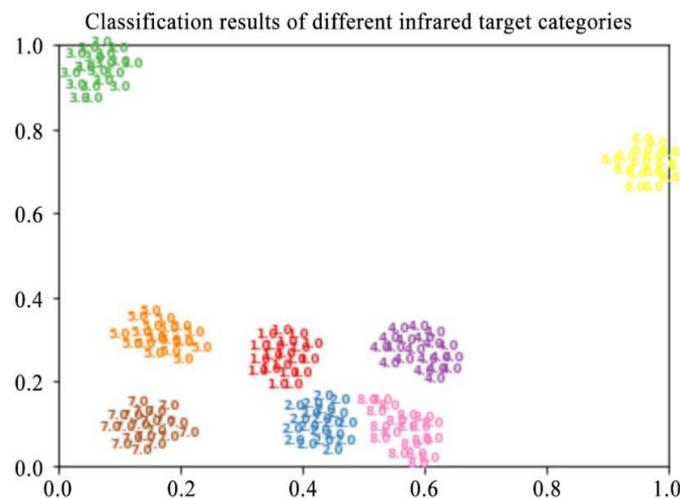


Figure 5. Classification results of different infrared object categories.

The reason for the clear classification results is that the characteristics of different infrared object types are quite different from each other. The number of infrared object categories we used is small, so the results of infrared object classification do not overlap with each other. The background changes of categories 1, 2, 4, 5, 7, 8 are not obvious, and the differences between objects and background are more obvious. These six categories of infrared object images have similar image distribution characteristics, so the infrared object point sets of these categories are closer in the figure. The point set of cars (category 3) and dog (category 6) is far from the point sets of other categories. This is because the pixel values of these two categories of infrared objects change greatly during the movement. Therefore, the position of their point set is farther away from several other categories of infrared objects.

We compare the proposed algorithm with three image classification algorithms: convolutional neural network (CNN) [14], multi-class Support Vector Machine (SVM) from LIBSVM [54], and multi-label lazy KNN (ML-KNN) [55] in terms of operation speed and classification accuracy. The comparison of the classification results of different algorithms used in this paper is shown in Figure 6.

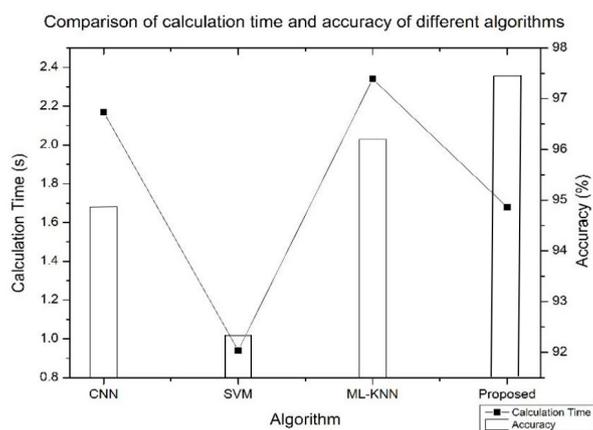


Figure 6. Comparison of calculation time and accuracy of different algorithms.

As we can see from Figure 4, CNN (Convolutional Neural Network) is not as good as the proposed algorithm because of its complex structure and a small number of data set samples, which does not allow us to train the network effectively. Although SVM (Support Vector Machine) is faster in finding classification hyperplanes in the high-dimensional space, it cannot accurately divide the infrared object image data points. As a result, the accuracy of SVM is lower than the accuracy of the proposed algorithm. The ML-KNN (Multi-label Lazy KNN) algorithm is more accurate, but it takes more time to calculate the result when compared to the algorithm proposed in this paper.

The results are summarized in Table 3 using typical classification assessment metrics [56]. Here FPR is False Positive Rate and AUC is Area under Receiver Operating Characteristic (ROC) Curve.

Table 3. Summary of performance characteristics of infrared object tracking methods. CNN: Convolutional Neural Network. SVM: Support Vector Machine. ML-KNN: Multi-label Lazy KNN.

Metric	CNN	SVM	ML-KNN	Proposed
Accuracy	94.8%	92.4%	96.2%	97.46%
FPR	0.044	0.093	0.054	0.032
Precision	0.932	0.911	0.955	0.963
Recall	0.950	0.931	0.968	0.988
F-score	0.941	0.921	0.961	0.975
AUC	0.967	0.958	0.979	0.987

4. Discussion and Final Remarks

We have developed and implemented an infrared object classification method for infrared images with mainly static backgrounds, i.e., under the condition that there are few movements in the background. For this type of image, our method achieved a high accuracy of 97.46%, which exceeded the accuracy of other methods using state-of-the-art object classifiers such as CNN, SVM or ML-KNN. The algorithm proposed in this paper can establish a high-dimensional manifold space of infrared object images and can classify different categories of infrared objects. Particularly, the proposed method can successfully perform infrared object classification even if only a small number of images are available for training. Most importantly, our method can effectively work with small data samples, on which the deep learning networks cannot be trained effectively. Finally, our experiments verify that the proposed algorithm can effectively classify different categories of infrared objects in the manifold space.

The achieved result demonstrates that our method could already be used for several applications such as infrared security cameras. Based on the main concepts used in the development of the method presented in this paper, we plan to work further on the development of new methods for infrared object tracking in images with dynamic backgrounds and cluttered object space, focusing on such

applications as autonomous driving or military applications (such as those described in [57,58]) based on the concepts and ideas which were successfully validated in this paper.

Author Contributions: Conceptualization, L.K., W.W., and R.L.; methodology, L.K., W.W., and R.L.; software, G.H., Z.Z., and L.K.; validation, G.H., Z.Z., L.K., W.W., and R.L.; formal analysis, G.H., Z.Z., L.K., W.W., and M.W.; investigation, G.H., Z.Z., L.K., W.W., R.L., and R.D.; resources, G.H., Z.Z., L.K., and R.L.; writing—original draft preparation, G.H., Z.Z., L.K., and W.W.; writing—review and editing, R.D., and M.W.; visualization, G.H., Z.Z., and L.K.; supervision, W.W., and R.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the National Natural Science Foundation of China (No. 61671222) and Development Program of Shaanxi Province (No. 2018ZDXM-GY-036) and Shaanxi Key Laboratory of Intelligent Processing for Big Energy Data (No. IPBED7).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Dara, S.; Tamma, P. Feature Extraction by Using Deep Learning: A Survey. In Proceedings of the Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 29–31 March 2018; pp. 1795–1801, doi:10.1109/ICECA.2018.8474912.
- Gabryel, M.; Damaševičius, R. The image classification with different types of image features. In Proceedings of the International Conference on Artificial Intelligence and Soft Computing ICAISC, Zakopane, Poland, 11–15 June 2017; pp. 497–506, doi:10.1007/978-3-319-59063-9_44.
- Zhou, B.; Duan, X.; Ye, D.; Wei, W.; Woźniak, M.; Damaševičius, R. Heterogeneous image matching via a novel feature describing model. *Appl. Sci.* **2019**, *9*, 4792, doi:10.3390/app9224792.
- Zhou, B.; Duan, X.; Ye, D.; Wei, W.; Woźniak, M.; Połap, D.; Damaševičius, R. Multi-level features extraction for discontinuous object tracking in remote sensing image monitoring. *Sensors* **2019**, *19*, 4855, doi:10.3390/s19224855.
- Zhou, B.; Duan, X.; Wei, W.; Ye, D.; Wozniak, M.; Damasevicius, R. An adaptive local descriptor embedding zernike moments for image matching. *IEEE Access* **2019**, *7*, 183971–183984, doi:10.1109/ACCESS.2019.2960203.
- Riaz, F.; Azad, M.A.; Arshad, J.; Imran, M.; Hassan, A.; Rehman, S. Pervasive blood pressure monitoring using Photoplethysmogram (PPG) sensor. *Future Gener. Comput. Syst.* **2019**, *98*, 120–130, doi:10.1016/j.future.2019.02.032.
- Hui, B.; Song, Z.; Fan, H.; P.; Zhong, H.; Hu, W.; Zhang, X.; Ling, J.; Su, H.; Jin, W.; Zhang, Y.; Bai, Y. A dataset for dim-small object detection and tracking of aircraft in infrared image sequences. *China Sci. Data* **2019**, *1*–12, doi:10.11922/csdata.2019.0074.zh.
- Li, Y.; Li, Z.; Zhu, Y.; Li, B.; Xiong, W.; Huang, Y. Thermal infrared small ship detection in sea clutter based on morphological reconstruction and multi-feature analysis. *Appl. Sci.* **2019**, *9*, 3786, doi:10.3390/app9183786.
- Ma, Y.; Wang, Y.; Mei, X.; Liu, C.; Dai, X.; Fan, F.; Huang, J. Visible/Infrared combined 3D reconstruction scheme based on nonrigid registration of multi-modality images with mixed features. *IEEE Access* **2019**, *7*, 19199–19211, doi:10.1109/ACCESS.2019.2895905.
- Younsi, M.; Diaf, M.; Siarry, P. Automatic multiple moving humans detection and tracking in image sequences taken from a stationary thermal infrared camera. *Expert Syst. Appl.* **2020**, *146*, 113171, doi:10.1016/j.eswa.2019.113171.
- Chen, Y.; Song, B.; Du, X.; Guizani, M. Infrared small object detection through multiple feature analysis based on visual saliency. *IEEE Access* **2019**, *7*, 38996–39004, doi:10.1109/ACCESS.2019.2906076.
- Zhang, K.; Yang, K.; Li, S.; Chen, H. A difference-based local contrast method for infrared small object detection under complex background. *IEEE Access* **2019**, *7*, 105503–105513, doi:10.1109/ACCESS.2019.2932729.
- Li, L.; Zhou, F.; Zheng, Y.; Bai, X. Reconstructed saliency for infrared pedestrian images. *IEEE Access* **2019**, *7*, 42652–42663, doi:10.1109/ACCESS.2019.2906332.
- Gao, J.; Lin, Z.; An, W. Infrared small object detection using a temporal variance and spatial patch contrast filter. *IEEE Access* **2019**, *7*, 32217–32226, doi:10.1109/ACCESS.2019.2903808.

15. Lan, X.; Ye, M.; Shao, R.; Zhong, B.; Jain, D.K.; Zhou, H. Online non-negative multi-modality feature template learning for RGB-assisted infrared tracking. *IEEE Access* **2019**, *7*, 67761–67771, doi:10.1109/ACCESS.2019.2916895.
16. Li, J.; Huo, H.; Sui, C.; Jiang, C.; Li, C. Poisson reconstruction-based fusion of infrared and visible images via saliency detection. *IEEE Access* **2019**, *7*, 20676–20688, doi:10.1109/ACCESS.2019.2897320.
17. Ma, M. Infrared pedestrian detection algorithm based on multimedia image recombination and matrix restoration. *Multimed. Tools Appl.* **2019**, 1–16, doi:10.1007/s11042-019-7444-4.
18. Qu, H.; Zhang, L.; Wu, X.; He, X.; Hu, X.; Wen, X. Multiscale object detection in infrared streetscape images based on deep learning and instance level data augmentation. *Appl. Sci.* **2019**, *9*, 565, doi:10.3390/app9030565.
19. Shen, G.; Zhu, L.; Lou, J.; Shen, S.; Liu, Z.; Tang, L. Infrared multi-pedestrian tracking in vertical view via siamese convolution network. *IEEE Access* **2019**, *7*, 42718–42725, doi:10.1109/ACCESS.2019.2892469.
20. Sun, Y.; Yang, J.; Li, M.; An, W. Infrared small-faint object detection using non-i.i.d. mixture of gaussians and flux density. *Remote Sens.* **2019**, *11*, 2831, doi:10.3390/rs11232831.
21. Sun, Y.; Yang, J.; Long, Y.; An, W. Infrared small object detection via spatial-temporal total variation regularization and weighted tensor nuclear norm. *IEEE Access* **2019**, *7*, 56667–56682, doi:10.1109/ACCESS.2019.2914281.
22. Wang, H.; Shi, M.; Li, H. Infrared dim and small object detection based on two-stage U-skip context aggregation network with a missed-detection-and-false-alarm combination loss. *Multimed. Tools Appl.* **2019**, 1–22, doi:10.1007/s11042-019-7643-z.
23. Yun, S.; Kim, S. TIR-MS: Thermal infrared mean-shift for robust pedestrian head tracking in dynamic object and background variations. *Appl. Sci.* **2019**, *9*, 3015, doi:10.3390/app9153015.
24. Zhang, P.; Wang, X.; Wang, X.; Fei, C.; Guo, Z. Infrared small object detection based on spatial-temporal enhancement using quaternion discrete cosine transform. *IEEE Access* **2019**, *7*, 54712–54723, doi:10.1109/ACCESS.2019.2912976.
25. Zhang, T.; Wu, H.; Liu, Y.; Peng, L.; Yang, C.; Peng, Z. Infrared small object detection based on non-convex optimization with lp-norm constraint. *Remote Sens.* **2019**, *11*, 559, doi:10.3390/rs11050559.
26. Zhang, Y.; Zheng, L.; Zhang, Y. Small infrared object detection via a mexican-hat distribution. *Appl. Sci.* **2019**, *9*, 5570, doi:10.3390/app9245570.
27. Zhou, F.; Wu, Y.; Dai, Y.; Wang, P. Detection of small object using Schatten 1/2 quasi-norm regularization with reweighted sparse enhancement in complex infrared scenes. *Remote Sens.* **2019**, *11*, 2058, doi:10.3390/rs11172058.
28. Zhang, K.; Li, X. Infrared small dim object detection based on region proposal. *Optik* **2019**, *182*, 961–973.
29. Deng, L.; Zhang, J.; Zhu, H. Infrared moving point object detection using a spatial-temporal filter. *Infrared Phys. Technol.* **2018**, *95*, 122–127.
30. Nie, J.; Qu, S.; Wei, Y.; Zhang, L.; Deng, L. An infrared small object detection method based on multiscale local homogeneity measure. *Infrared Phys. Technol.* **2018**, *90*, 186–194.
31. Ge, H.; Zhu, Z.; Lou, K. Tracking video target via particle filtering on manifold. *Inf. Technol. Control.* **2019**, *48*, 538–544, doi:10.5755/j01.itc.48.4.23939.
32. Zhu, J.Y.; Krähenbühl, P.; Shechtman, E.; Efros, A.A. Generative Visual Manipulation on the Natural Image Manifold. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 597–613.
33. Cong, W.; Wang, G.; Yang, Q.; Li, J.; Hsieh, J.; Lai, R. CT image reconstruction on a low dimensional manifold. *Inverse Probl. Imag.* **2019**, *13*, 449–460, doi:10.3934/ipi.2019022.
34. Luo, F.; Huang, Y.; Tu, W.; Liu, J. Local manifold sparse model for image classification. *Neurocomputing* **2019**, *382*, 162–173, doi:10.1016/j.neucom.2019.11.084.
35. Bernstein, A.; Kuleshov, A.; Yanovich, Y. Manifold Learning in Regression Tasks. In Proceedings of the International Symposium on Statistical Learning and Data Sciences, Egham, United Kingdom, 20–23 April 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 414–423.
36. Bai, S.; Bai, X.; Tian, Q. Scalable Person Re-Identification on Supervised Smoothed Manifold. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2530–2539.
37. Zhu, B.; Liu, J.Z.; Cauley, S.F.; Rosen, B.R.; Rosen, M.S. Image reconstruction by domain-transform manifold learning. *Nature* **2018**, *555*, 487.

38. Tenenbaum, J.B.; de Silva, V.; Langford, J.C. A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science* **2000**, *290*, 2319–2323.
39. Calandra, R.; Peters, J.; Rasmussen, C.E.; Deisenroth, M.P. Manifold Gaussian Processes for Regression. In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, Canada, 24–29 July 2016; pp. 3338–3345.
40. Lu, J.; Wang, G.; Deng, W.; Moulin, P.; Zhou, J. Multi-Manifold Deep Metric Learning for Image Set Classification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1137–1145.
41. Wei, D.; Shen, X.; Sun, Q.; Gao, X.; Yan, W. Prototype learning and collaborative representation using Grassmann manifolds for image set classification. *Pattern Recognit.* **2020**, *100*, 107123, doi:10.1016/j.patcog.2019.107123.
42. Xu, J.; Wang, C.; Qi, C.; Shi, C.; Xiao, B. Iterative manifold embedding layer learned by incomplete data for large-scale image retrieval. *IEEE Trans. Multimed.* **2019**, *21*, 1551–1562, doi:10.1109/TMM.2018.2883860.
43. Cui, J.; Li, S.; Xia, Q.; Hao, A.; Qin, H. Learning multi-view manifold for single image based modeling. *Comput. Gr.* **2019**, *82*, 275–285, doi:10.1016/j.cag.2019.05.030.
44. He, R.; Feng, X.; Wang, W.; Zhu, X.; Yang, C. W-LDMM: A wasserstein driven low-dimensional manifold model for noisy image restoration. *Neurocomputing* **2020**, *371*, 108–123, doi:10.1016/j.neucom.2019.08.088.
45. Liu, T.; Shi, Z.; Liu, Y. Visualization of the image geometric transformation group based on riemannian manifold. *IEEE Access* **2019**, *7*, 105531–105545, doi:10.1109/ACCESS.2019.2932412.
46. Zhao, X.; Li, Y.; Wang, H. Manifold based on neighbour mapping and its projection for remote sensing image segmentation. *Int. J. Remote Sens.* **2019**, *40*, 9304–9320, doi:10.1080/01431161.2019.1629718.
47. Liu, X.; Ma, Z.; Niu, G. Mixed region covariance discriminative learning for image classification on riemannian manifolds. *Math. Prob. Eng.* **2019**, *2019*, 1261398, doi:10.1155/2019/1261398.
48. Lu, J.; Tan, Y.P.; Wang, G. Discriminative multimanifold analysis for face recognition from a single training sample per person. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *35*, 39–51.
49. Qi, G.-J.; Luo, J. Small Data Challenges in Big Data Era: A Survey of Recent Progress on Unsupervised and Semi-Supervised Methods. *arXiv* **2019**, arXiv:preprint/1903.11260.
50. Dean, J.; Corrado, G.; Monga, R.; Chen, K.; Devin, M.; Le, Q.V.; Mao, M.Z.; Ranzato, M.A.; Senior, A.W.; Tucker, P.A.; et al. Large Scale Distributed Deep Networks. In Proceedings of the Neural Information Processing Systems NIPS, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1232–1240.
51. Akula, A.; Ghosh, R.; Kumar, S.; Sardana, H.K. Moving object detection in thermal infrared imagery using spatiotemporal information. *JOSA A* **2013**, *30*, 1492–1501.
52. Berg, A.; Ahlberg, J.; Felsberg, M. A Thermal Object Tracking Benchmark. In Proceedings of the 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Karlsruhe, Germany, 25–28 August 2015; pp. 1–6.
53. Davis, J.; Sharma, V. Background-Subtraction using Contour-based Fusion of Thermal and Visible Imagery. *Comput. Vision Image Underst.* **2007**, *106*, 162–182.
54. Chang, C.; Lin, C. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 1–27, doi:10.1145/1961189.1961199.
55. Zhang, M.-L.; Zhou, Z.-H. ML-KNN: A lazy learning approach to multi-label learning. *Pattern Recognit.* **2007**, *40*, 2038–2048, doi:10.1016/j.patcog.2006.12.019.
56. Tharwat, A. Classification assessment methods. Available online: <https://www.sciencedirect.com/science/article/pii/S2210832718301546> (accessed on 2 February 2020).
57. d’Acremont, A.; Fablet, R.; Baussard, A.; Quin, G. CNN-Based Target Recognition and Identification for Infrared Imaging in Defense Systems. *Sensors* **2019**, *19*, 2040.
58. Ivanovas, A.; Ostreika, A.; Maskeliūnas, R.; Damaševičius, R.; Połap, D.; Woźniak, M. Block Matching Based Obstacle Avoidance for Unmanned Aerial Vehicle. In Proceedings of the Artificial Intelligence and Soft Computing, ICAISC, Zakopane, Poland, 3–7 June 2018; Springer: Berlin/Heidelberg, Germany, 2018, Volume 10841, pp. 58–69, doi:10.1007/978-3-319-91253-0_6.

